TEACHER EFFECTIVENESS SYSTEMS, FRAMEWORKS AND MEASURES:

A REVIEW

Prepared by:

Prof Janet Clinton

Dr Melody Anderson

Georgia Dawson

Alexander Dawson

Scott Bolton

Rob Mason

For more information on the Centre for Program Evaluation Please visit http://education.unimelb.edu.au/cpe

TABLE OF CONTENTS

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Teacher Effectiveness Systems, Frameworks and Measures: A Review

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

x

Teacher Effectiveness Systems, Frameworks and Measures: A Review

AEDs	Allied Educators
AITSL	Australian Institute for Teaching and School Leadership
AMs	Deputy Head and Assistant Masters/Mistresses
APST	The Australian Professional Standards for Teachers
AST	Academy of Singapore Teachers
ATAR	Australian Tertiary Admission Rank
ATPDF	The Australian Teacher Performance and Development Framework
BME	Black and Ethnic Minority
CALD	Culturally and Linguistically Diverse
CLASS™	The Classroom Assessment Scoring System
CMs	Certificated Masters/Mistresses
CPE	Centre for Program Evaluation
CSET	The California Subject Exam for Teachers
CSTP	California Standards for the Teaching Profession
CTA	The California Teacher Association
DC	District of Columbia
DCPS	District of Columbia Public Schools
DOEs	The Desired Outcomes for Education
EBAs	Enterprise Bargaining Agreements
ECE	Early childhood education
ECNZ	Education Council of New Zealand
EDB	Education Bureau
EIS	Educational Institute of Scotland
ERO	Education Review Office
GCE N(A)	General Certificate of Education (Normal; Academic)
GCE N(T)	General Certificate of Education (Normal; Technical)
GCE O	General Certificate of Education (Ordinary)
GMs	Graduate Masters/Mistresses
GPA	Grade point average
GTCS	General Teaching Council for Scotland
HKCAAVQ	The Hong Kong Council for Accreditation of Academic and Vocational Qualifications
ITE	Initial Teacher Education

ITT	Initial Teacher Training
KMK	Kultusministerkonferenz
MET	Measures of Effective Teaching
MGSE	Melbourne Graduate School of Education
MOE	Ministry of Education
NAPLAN	National Assessment Program – Literacy and Numeracy
NBCs	National Board Certificated teachers
NBPTs	The National Board for Professional Teaching Standards
NCTL	The National College for Teaching and Leadership
NGO	Non-Government Organisation
NIE	National Institute of Education
NQT	Newly Qualified Teachers
NZQA	New Zealand Qualifications Authority
NZTC	New Zealand Teachers Council
OECD	Organisation for Economic Co-operation and Development
Ofsted	Office for Standards in Education
PACT	Performance Assessment of California Teachers
PARCC	Partnership for Assessment of Readiness for College and Careers
PGCE	Post Graduate Certificate in Education
PGDE	Professional Graduate Diploma in Education
PGMs	Principal Graduate Masters/Mistresses
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PRD	Professional Review and Development
PRTs	Provisionally Registered Teachers
PSLE	Primary School Leaving Examination
PT	Permitted Teacher
QTS	Qualified Teacher Status
RT	Registered Teacher
SCQF	Scottish Credit and Qualifications Framework
SGMs	Senior Graduate Masters/Mistresses
TIMSS	Trends in International Mathematics and Science Study
TPA	Teaching Performance Assessment
TPM	Teacher Performance Management
UoMC	The University of Melbourne Commercial Limited

Teacher Effectiveness Systems, Frameworks and Measures: A Review

VAM	Value-Added Modelling
VCLA	Virginia Communication and Literacy Assessment
VSE	Validated Self-Evaluation

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Evaluation	The systematic and objective assessment of an on-going or completed project, program, or policy, and its design, implementation, and results.
Effective teaching	Strategies which research and workplace knowledge suggests contribute to successful learning outcomes for students.
Formative	Evaluating student learning to provide feedback to students and devise/change teaching and learning programs.
Minority groups	Persons belonging to national or ethnic, religious, and linguistic minorities, enriching the diversity of their societies.
Observation	A formal or informal observation of teaching while it is taking place in a classroom or other learning environment. Typically conducted by fellow teachers, administrators, or instructional specialists, classroom observations are often used to provide teachers with constructive critical feedback aimed at improving their classroom management and instructional techniques.
Pre-service teachers	Students in initial teacher education programs provided by higher education institutions.
Registration	Regulatory processes for entry and continued employment in the teaching profession.
Standards	A common understanding and language for professional discussions and reflections on teacher preparation, practice, and improvement.
Summative	Evaluating student achievement of learning goals at a point in time.
Value-add model	A measure of a teacher's impact on student achievement—that is, the value he or she adds—apart from other factors that affect achievement, such as individual ability, family environment, past schooling, and the influence of peers.

Evaluating the performance and impact of teachers is an integral part of ensuring the quality and effectiveness of teaching. Any framework or model used to evaluate the qualities of teachers and teaching must prioritise the consequential impact on students. A major claim throughout this review is that the criterion of success of any evaluation of teachers and teaching should be framed in terms of this impact.

The specific aim of this project was to examine and summarise current international policies and practices regarding teacher evaluation, evaluate these systems, and subsequently inform the development of a teacher effectiveness measure for Australia. Evaluation systems from Australia; New Zealand; England; Scotland; Germany; Austria; Singapore; Hong Kong; South Korea; California, USA; Virginia, USA; Washington, USA; Washington D.C., USA; Ontario, Canada; British Columbia, Canada; and Alberta, Canada were selected for review on the basis of comparability to the Australian context. This examination and summary was achieved through a rapid synthesis of controlled research, grey literature, and policy documents relating to teacher effectiveness and evaluation procedures. Policies, practices, and procedures were summarised and represented in an evaluation crosswalk for ease of reference and clarity. All information was then further evaluated using a traffic light system, with categories chosen to reflect the extent of implementation and the levels within each system where implementation had occurred.

Systems for teacher evaluation vary from sophisticated national systems to localised, informal approaches. Systemic approaches are typically framed by national or statutory professional standards that articulate the knowledge, practice and engagement required across career stages.

Major features of effective systems include a well-developed and coherent evaluation framework, multiple evidenced based dimensions, and utilise multiple methods and tools that are implemented by a well-trained workforce of evaluators.

The quality, validity and reliability of tools currently available vary significantly across countries and contexts. An extensive review of currently available measures of teacher effectiveness is provided in this report, entailing a description and overview of each method and an analysis of evidence relating to reliability and validity. The measures include classroom observation, teaching performance portfolios, teacher interviews, performance and development interviews, peer ratings and student ratings.

The influence of the teaching and learning environment, specifically the instructional context, curriculum and assessment systems, class size, facilities, and materials, also needs to be considered by policymakers in implementing evaluation systems.

This review provides a listing of common dimensions that are utilised, in the accountability frameworks and measures of effective teachers and teaching, by the sample countries. This list is not intended to represent a model of effective teaching, but rather, to categorise a view of those dimensions commonly utilised.

Through the synthesis numerous factors associated with teacher effectiveness, influence and impact emerged that related to a specific evidence-based framework or theory, professional standards and appraisal frameworks or a specific measure or evaluation tools. These factors have been conceptually sorted and a diagrammatic model (see the figure below) that includes higher order dimensions is articulated in this report.

Teaching

Subject matter knowledge
Instructional practice skill
Pedagogical knowledge
Preparation & planning
Evaluation, assessment & feedback
Learning strategies

Teacher as a person

Communication
Mind frames
Psychosocial resources, & attitude, Cultural competency
Numeracy & literacy ready

Behaviour as a teacher

Adhere to a set of standards
Professionalism
Leadership
Accreditation & credentialing

Continuous learning

Skill & specialisation
Career progression
Subject specialisation

Influence & Impact

Student achievement
Student affect
Colleagues
Classroom climate
Parent engagement
School culture

Figure 1 Dimensions of effective teaching

It should be noted that these dimensions emerge from the sample of countries and states reviewed. The model posits that teacher effectiveness relates to the impact and influence the teacher has on the education community. While impact on the learning lives of students is core to this, this review points to the premise that teacher effectiveness, in relation to impact and influence, is multi-dimensional and that there are a number of factors that relate causally to these two outcomes.

This review is descriptive by design and provides a high-level picture of the state of play of teacher evaluation and teacher evaluation measures. This summary includes a series of high-level recommendations for government and policymakers in relation to measures of teacher effectiveness categorised as:

SYSTEM-LEVEL POLICY SETTINGS AND PRE-REQUISITES

- Standards must be evidence-based and support career progression
- The evaluation system needs to separate the formative and summative evaluation of quality teaching
- There needs to be a regulatory body in place to articulate and support evaluation, and house the results
- Evaluative judgments must be made against standards that are progressive across career

2

Teacher Effectiveness Systems, Frameworks and Measures: A Review

- Further definitional work around teacher effectiveness and impact is needed, working towards agreement on the definition of these terms

FEATURES OF HIGH QUALITY MEASURES OF TEACHER EFFECTIVENESS

RESOURCES
- Measures need to reflect progressive development
- Measures must be valid, reliable, consistent over time, and unbiased
- Multiple methods of measurement need to be available and utilised simultaneously

CONTEXT
- Measures must be adaptable to multiple contexts
- Remote and various communities must be able to utilise measures
- Multiple voices need to be taken into account, for example, student, teachers, and parents

DIMENSIONS
- Instructional practice
- Skill and knowledge
- Professionalisation-consideration
- Teacher impact on educational community
- Student learning outcomes (cognitive and non-cognitive)

APPLICATION/IMPLEMENTATION OF MEASURES OF TEACHER EFFECTIVENESS

MEASUREMENT ASSUMPTIONS AND PROCEDURES
- Defensible weighting of the various components must take place in relation to individual teachers
- Set of guidelines need to exist for the conglomeration or weighting of measures to make effective judgments
- Measures must operate across different contexts
- Defensible cut-scores for making decisions about high quality

EVALUATOR EXPERTISE
- Expert evaluators are available
- Training for evaluators is accessible
- Capacity building in the form of induction and ongoing support needs to be available for teachers and evaluators

USE OF EVALUATION RESULTS
- Evaluative results must be utilised for educational improvement of school community
- Results need to inform policy and practice within the context being evaluated
- Results must support decisions about professional growth, e.g. professional learning, certification at higher levels of the standards, etc.

EVALUATION PROCEDURES: REMOTE AND INDIGENOUS COMMUNITIES.
- Measures must be adaptable to multiple contexts
- Remote and Indigenous communities need to able to utilise measures
- Communities engaged in process and criteria for success

Teacher Effectiveness Systems, Frameworks and Measures: A Review

APPLICATION TO AN AUSTRALIAN CONTEXT

This review highlights pockets of excellence; however, no one country or state provides an exemplar of a teacher evaluation system or measures that are generally applicable to the Australian context.

The current Australian system for teacher evaluation is noted by its accreditation systems. While the system at a national level provides strong guidelines for career pathways for teachers and guidance for aspiration for quality teaching as per the standards, there is little agreement about measures or methods for evaluating the levels of quality of teaching across the country. Furthermore, the systems appear to add little value to the enhancement of teaching practice. Consequently, our understanding of quality teaching and our evidence as a nation remains anecdotal and costly. There also appears variable and seemingly, little capacity to evaluate teacher effectiveness at a local, state or national level, hence evaluation capacity building in the workforce is needed.

Therefore, in addition to taking into account these generic features of effective systems there are a number of elements that surface from this review that may be valuable to consider in relation to an Australian teacher evaluation system and measurement of teacher effectiveness:

DIMENSIONS OF TEACHER EFFECTIVENESS:

- Based on the Australian Professional Standards for Teaching
- Consider teacher effectiveness as multidimensional
- Context variable
- The model of teacher effectiveness is evidence bound

AN EVALUATION SYSTEM:

- Adheres to the evaluation standards of usability, feasibility, validity, accuracy and accountability
- Relates to an aspirational process that empowers teachers and celebrates the profession
- Ensures an evaluation workforce and encourages evaluation capacity building
- Encourages a mindset of evaluative thinking
- Encourages life long career development
- Considers value for money
- Provides summative evaluation, such as accreditation and credentialing
- Provides for continuous monitoring and improvement

THE SYSTEM CONTEXT:

- Adapts to context
- Engages Indigenous communities and leaders
- Describes appropriate guidelines for rural, remote and Indigenous areas
- Describes appropriate guidelines for remote and rural communities
- Applies at a state and territory, jurisdiction and school level

MEASUREMENTS AND RESOURCES SHOULD:

- Adhere to assessment and measurement principles
- Utilise multiple methods
- Aggregate and weight multiple dimensions in a transparent and defensible manner
- Employ multiple iterations of appraisals
- Ensure multiple uses of data that informs policy and practice and theory
- Be adapted to teaching standards/experience expectations targets and indicators of success
- Consider timing of measures and development of teacher skill as well as utilising progressive assessment

This review provides grounds for a further systematic review of the dimensions of teacher effectiveness and aligned measures that may be appropriate for Australia. Understanding the adaptation of an evidence base related to remote and Indigenous communities should be a particular focus.

By and large this document offers a description and an assessment of international evaluation tools and their respective regulatory standards, frameworks, reliability and validity. It has been designed to be descriptive and deliver a high-level picture of the state of play of teacher evaluation and measures and assessment of teacher effectiveness. It is not meant to be all encompassing. However, the researchers believe it provides a guide for future directions to understand teacher evaluation and appropriate measures of teacher effectiveness. Essentially, it demonstrates the need for further examination of evidence, policy and practice encompassing systems of teacher evaluation in Australia.

While the Australian education system performs well on international rankings of student achievement, there is room for growth. Australian students scored higher than the OECD average on measures of mathematics, scientific and reading literacy on the 2015 PISA assessment. However, the ranking of Australian students has declined on international rankings of literacy, numeracy, and science in every testing cycle since 2003 (Thomson, De Bortoli, & Buckley, 2013). This trend has continued, with the most recent 2015 PISA results showing a continued decline in the PISA measurements of mathematics, scientific and reading literacy (OECD, 2016a). While PISA rankings currently fail to acknowledge many factors that comprise a high quality education system, such as arts, sports, social engagement, and student wellbeing, this trend is somewhat concerning. In contrast, many OECD countries have demonstrated gains (Thomson et al., 2013). Indeed, only Sweden, Finland, New Zealand and Iceland have declined at a faster rate than Australia (Thomson et al., 2013). While these international fluctuations in rankings on literacy, numeracy, and science may be attributed to the large gains made in developing countries, for Australia to maintain an internationally competitive education system it is important to continually aspire to greater educational outcomes. This relative international decline has occurred in parallel with the stagnation of NAPLAN literacy and numeracy results at a national level (ACARA, 2016). Further, Australia's decline in national and international measures has been occurring in most states and is not isolated to one particular group or sector of the education community (Thomson et al., 2013). These dual declines highlight the critical need to examine Australia's education system and strive for improvement in order to remain competitive in the evolving global economy.

Teachers have been shown to elicit the greatest impact of in-school variables (Hattie, 2009). Similarly, Darling-Hammond (2000) noted that the effects of quality teaching on student outcomes are greater than those that arise from students' backgrounds. This impact may be positive or negative, with the effect of poor quality teaching being seen as debilitating and cumulative on the student outcomes (Darling-Hammond, 2000). Slater, Davies, and Burgess (2009) noted that students displayed almost a 12 month difference in academic progression when being taught by a low-performing teacher versus a high-performing teacher.

As such, teacher quality and effectiveness have become critical components of school improvement. These two terms are often used interchangeably in the literature, and there has been considerable debate regarding their definition and appropriate use. Generally, teacher effectiveness can be conceptualised as being on a continuum covering the extent to which a teacher is able to progress student outcomes. This is often measured using student achievement results and other summative tools. Teacher quality, however, refers to teacher attributes, such as capabilities, training, knowledge, or beliefs.

High-performing education systems around the world have known and demonstrated that improving the quality and effectiveness of teaching can dramatically improve student outcomes at both a school and system level. These systems, such as Finland, Singapore, and Ontario, Canada, have shown that effective teachers learn from each other, growing and developing to continually target the learning needs of the child. More effective teachers, and by extension teaching, are key to improving school education (Hanushek, Kain, O'Brien, & Rivkin, 2005; Hanushek, Kain, & Rivkin, 1998; Rockoff, 2004).

In order to promote collaboration and systemic improvement, effective systems move policy levers that seek to increase the effectiveness– not only the quality – of teaching. For example, high-performing education systems have prioritised the mentoring of teachers by highly skilled observers who provide constructive feedback about how to better tailor their teaching to the needs of individual students. These systems have also continued to make time for all teachers, regardless of their experience or level, to undertake targeted practical research and professional development in their schools, with a focus on how to promote gains in student learning. Australia has attempted to develop professional standards for teachers as a quality mechanism to attract, develop, recognise, and retain teachers. However, the best-performing systems have used these standards as a means to continually improve teaching and relate teaching performance to its impact upon student achievement.

Governments and policy makers have now realised that measuring and evaluating teacher quality and effectiveness is key to improving educational outcomes for students. However, teacher evaluation has remained a controversial issue in education. Historically, mixed empirical evidence regarding its effect on student outcomes and the impact of

various stakeholders with different agendas has limited the widespread implementation of teacher evaluation programs. Recent investigations by schools, systems, and non-government organisations have brought teacher evaluation back into the spotlight.

Models of teacher evaluation can focus on summative or formative outcomes depending on the intended use of results. Summative use of evaluations can include providing accountability or informing high-stakes promotions. Formative use of evaluations can include providing constructive feedback and direction for professional development. As there have often been multiple sources of motivation for implementing teacher evaluation processes, teacher evaluation systems often seek to evaluate for both summative and formative purposes simultaneously.

As Australia has sought to improve student achievement and ensure it has a world-class system of education, increasing teacher quality and effectiveness have become essential policy. Effective evaluation of teacher effectiveness has permitted continuous, targeted professional development, allowing teachers to reflect on and continually develop their practice. Highly visible teacher effectiveness measurement and evaluation have also provided opportunities to incentivise, recognise, and reward teaching abilities and high performance. This, in turn, may help address concerns about the attractiveness of teaching as a career choice and about the image and status of teachers as professionals in Australia, thereby improving teacher quality.

The scope of this project was to develop an initial synthesis of research that '*draws on policy and international research to explore possible measures of teacher effectiveness and/or quality*'. In particular, this project focused on methodologies employed to monitor the effectiveness of a teaching workforce and their potential application to the Australian context. To meet the needs of this objective within the time frame, this should be considered a rapid review of the literature. This rapid review will provide an overview of teacher evaluation practices and inform policy consideration on the critical dimensions of teaching.

REVIEW OBJECTIVES AND QUESTIONS

The specific objectives of this review were to summarise and examine current international policies and practices regarding teacher evaluation, to understand and evaluate practices for evaluating teachers, and subsequently to inform the development of a measure of teacher effectiveness in Australia.

In particular, this review aimed to identify and explore specific measures of the effective teacher and teaching; however specific pedagogical approaches have been excluded. The evaluation of teaching has been considered to exist along a continuum, from entry into initial teacher education, to graduate teachers, through to fully registered teachers.

Review questions

1. What are the current policies in place that inform teacher evaluation in Australia and abroad?
2. What theories underpin these policies and approaches?
3. What are the critical dimensions for evaluating teachers?
4. What are the specific tools that exist for evaluating teachers, and are they supported by the literature as valid, reliable, and generally effective?

HOW TO USE THIS DOCUMENT

The findings of this document have been divided into three sections. They are outlined as follows:

1. **Evaluation Crosswalk** – In order to quickly examine and compare the different countries' policies and practices regarding teacher evaluation, the evaluation team have constructed an explorative crosswalk of policy, dimensions of interest and specific evaluation tools. This tool can be used as a quick reference to gain an understanding of a particular site and its components, or to compare a number of different sites on policy, dimensions of interest, and specific evaluation tools.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

2. **Narrative of the Report** – The narrative of the report is a more detailed and in-depth examination of:

   a. **International Teacher Evaluation Practices –** an examination of each of the sites and the specifics of the policy, dimensions of interest and specific evaluation tools.

   b. **Methods for Teacher Evaluation –** a broad overview of some methods for evaluating teachers and associated research.

   c. **Factors Associated with Teacher Effectiveness, Influence, and Impact –** a summary of some factors associated with teacher effectiveness, influence and impact.

3. **Highlights of Best Practice** – This traffic light system is a snap shot evaluation of the different policies, dimensions of interest, and specific evaluation tools for each of the sites. The traffic light system allows for the different sites and components of the sites to be quickly and effectively compared against each other.

METHODOLOGY

RESEARCH DESIGN AND METHODS

To examine the aforementioned research questions a broad literature search was conducted. Various documents were accessed, including controlled research, grey literature, and policy documents relating to teacher effectiveness. These documents were analysed, with relevant data and information extracted for summary. Findings were then systematically summarised in order to understand the common elements and key constructs relating to teacher effectiveness as well as other contextually appropriate approaches to the measurement and evaluation of effective teachers. As distinctly different sources were required to appropriately answer these questions, the review process consisted of two components:

1) Policy Review

The first stage of this review examined the procedures in place for teacher evaluation in Australia and internationally. A semi-systematic approach was used, with researchers searching government education websites, the OECD site and relevant academic databases. A limited number of sites were reviewed to fit within the rapid timeline involved in this study. Sites selected for review were:
1. Australia
2. New Zealand
3. The United Kingdom
   a. England
   b. Scotland
4. Canada
   a. Ontario
   b. British Columbia
   c. Alberta
5. Germany
6. Austria
7. Singapore
8. Hong Kong
9. South Korea
10. The United States of America
    a. California
    b. Washington State
    c. Washington D.C.
    d. Virginia

Teacher Effectiveness Systems, Frameworks and Measures: A Review

WHY WERE THESE SITES SELECTED?

Commonwealth countries (New Zealand, The United Kingdom [England and Scotland], and Canada [Ontario, British Columbia and Alberta]) were selected for their similarity to the Australian context and their comparability of education systems. Similarly, USA, Germany and Austria were selected for review as their education systems were considered to be sufficiently compatible with Australia. As teacher evaluation processes vary substantially between different states and provinces in the USA and Canada, expert recommendation informed the inclusion of several states from both these countries known to have teacher evaluation processes in place.

Several top performing OECD countries were also selected for review: Singapore, Hong Kong, and South Korea (OECD, 2013b). These three sites were included as they are acknowledged for having comprehensive systems for teacher evaluation as well as achieving high rankings on tests of international student achievement. Findings from these countries should be viewed with historical and contextual factors in mind. It should be noted that these education systems are not entirely compatible with Australian teaching practices, and findings will not be directly transferable to the Australian system. Shanghai, Taipei, and Macao were excluded from this review, as these education systems were deemed incompatible with Australian teaching practices.

Despite their high performance in literacy, numeracy and science measures, it should be noted that several top-performing OECD countries do not have any formal teacher evaluation systems in place, including Finland and Belgium (French). Other OECD countries with no teacher evaluation system include Denmark, Norway, Iceland, and Spain (OECD, 2012).

2) Evaluation Tool Research Synthesis

The second stage of this review was to synthesise the literature regarding specific evaluation frameworks and evaluation tools relating to teacher evaluation. These are summarised in the Constructs of Interest section that follows.

A. SOURCES

To examine current teacher evaluation practices in Australia and abroad the following sources were examined:

- Official government websites, including Education Departments and other Government entities relating to education
- Government-affiliated websites, such as teaching authorities and regulatory bodies
- Non-government body websites relating to education, including teaching unions and advisory groups
- International research groups, such as the OECD

To examine specific evaluation tools used for teacher evaluation the following sources were examined:

- Randomised controlled trials
- Quasi experimental studies
- Existing reviews and synthesis
- Research-based discussion and white papers
- Qualitative research

Teacher Effectiveness Systems, Frameworks and Measures: A Review

B. TYPES OF PARTICIPANTS:

A range of groups spanning initial teacher education candidates and the teacher workforce was included. Hence a number of different organisations were included:

- Government, both national and international
- Peak Bodies
- Research Institutes
- Commercial and NGO
- Regulatory Bodies

C. CONSTRUCTS OF INTEREST

The following constructs were considered along a developmental continuum from initial teacher education candidate to lead teacher:

- Teacher Preparation
  - o Licencing and registration of teaching courses
  - o Alternate teacher preparation programs (e.g. Teach for Australia)
  - o Registration processes for new graduates
  - o Ongoing/Renewable Registration
- Course Accreditation
  - o Regulatory bodies
- Outcome Measures
  - o Student Results
    - ▪ National exams
    - ▪ OECD results (PISA, TIMSS and PIRLS)
    - ▪ Value-Added Modelling (VAM)
  - o Teacher Knowledge
    - ▪ Tests of subject content knowledge
    - ▪ Tests of pedagogical knowledge
  - o Survey Results
    - ▪ Teacher self-ratings
    - ▪ Student ratings
    - ▪ Peer ratings
    - ▪ Senior staff ratings (principals/lead teachers)
  - o Classroom Observation
    - ▪ Mentor observation ratings
    - ▪ Peer observation ratings
    - ▪ Senior staff observation ratings (principals/lead teachers)
- Evaluation Information Use
  - o Summative Use
    - ▪ Pre-service practicum assessment
    - ▪ Initial teacher certification
    - ▪ Initial and ongoing employment decisions
    - ▪ Performance pay/Salary decisions
    - ▪ Sanctions for underperformance
  - o Formative Use
    - ▪ Identifying goals for improved performance
    - ▪ Informing professional development

Teacher Effectiveness Systems, Frameworks and Measures: A Review

- Contextual Factors and Disadvantaged Communities
    - Indigenous populations
    - Children with disabilities
    - Children from Culturally and Linguistically Diverse (CALD) communities
    - Rural communities
    - Absenteeism
    - Access to education

D. INCLUSION CRITERIA

Policy documents were included if they met the following criteria:

- Related to Primary and Secondary school teachers
- Currently implemented policy documents
- English-language documents

Research related to specific evaluation tools used for teacher evaluation was included if it met the following criteria:

- Related to Primary and Secondary school teachers
- Conducted from 2006 to current
- English-language documents

E. EXCLUSION CRITERIA

To ensure efficient selection, increased quality, and precision of the review and therefore its findings, specific exclusion criteria were utilised. Policy documents were excluded based on the following criteria:

- Early childhood teaching and teachers
- Non-English literature
- Non-implemented policy

Research relating to specific evaluation tools used for teacher evaluation was excluded based on the following criteria:

- Non-Primary and Secondary school teachers (i.e. early learning and tertiary teachers)
- Research conducted prior to 2006
- Non-English publication only

F. ELECTRONIC SEARCHES

To extract policy information, the following sources were used:

- Government websites
- Peak body websites (e.g. teaching union websites)
- The OECD publication library

To obtain articles, documents, chapters, and texts relating to specific evaluation systems, methods, tools, and features, widely used online databases were used, including:

- *JSTOR*, a digital library of academic journals, books and primary sources, with full- text searches of almost 2,000 journals;
- *ERIC* (Educational Resources Information Center), an online library of education research and information, sponsored by the Institute of Education Sciences (IES) of the U.S. Department of Education;

Teacher Effectiveness Systems, Frameworks and Measures: A Review

- *PsycINFO,* from the American Psychological Association, which includes 3.7 million citations of peer-reviewed journals, books, dissertations in psychology and behavioural sciences;
- *Google Scholar*, a search engine providing peer-reviewed papers, theses, books, abstracts and articles from academic publishers.
- Communication with researchers, experts, and research institutes in the field of teacher effectiveness
- Proquest Education
- The OECD publication library

G. ETHICS

As this was a review of secondary data, ethical clearance was not required.

LIMITATIONS

It should be acknowledged that there were several inherent limitations to a rapid synthesis as a methodological approach. Given the scope and breadth of this study and the limited timeframe involved, a deep analysis was not feasible. Instead, this review is best considered a probing study that provides a broad overview of teacher evaluation as it relates to the Australian context in order to provide guidance for future comprehensive analyses. Further, while this study employed a systematic approach, the rapid pace of this review also increased the likelihood of missing relevant information. Finally, as with other systematic reviews, the researchers were limited by their own language and context. While this limited the exploration of some international documents, the researchers endeavoured to explore international research where available utilising documents in English, translated documents, and secondary sources where appropriate.

EVALUATION CROSSWALK

Table 1 - Evaluation Crosswalk of Chosen Sites

		AUSTRALIA	NEW ZEALAND	ENGLAND
STANDARDS AND THEORY	Who sets the standard?	National – Department of Education and Training (DET) in collaboration with AITSL	National – Education Council of New Zealand (ECNZ) and endorsed by the Ministry of Education	National – Department for Education (DfE)
	Do standards vary for different career levels?	Yes	Only for registering practicing teachers	No
	What theoretical framework underpins teacher standards?	Influenced by the work of John Hattie and Australian Professional Standards for Teachers (APST) in all states/territories	*Practicing Teacher Criteria* and other domains of professionalism developed by ECNZ	Influenced by the work of Sir Michael Barber and professional review
TEACHER PREPARATION	Accreditation of ITE	State/territory level authorities as well as meeting AITSL standards	National -ECNZ	National – Department for Education
	Probation/Provisional teaching status	Yes – Provisional teaching status duration varies between differing states/territories, tested against the APST standards	Yes – Induction period lasting 2-6 years depending on ability to meet the Practicing Teacher Criteria	Yes – Teachers have a one year probationary period, following which they are appraised, and receive full registration if they pass
	Registration	State/territory level authorities	National - ECNZ	National – Awarded by the National College for Teaching and Leadership, on behalf of the Secretary of State
REGISTRATION	Registering Body	State/territory level authorities	National - ECNZ	National - Awarded by the National College for Teaching and Leadership, on behalf of the Secretary of State.
	Alternate pathways to registration available	Yes. Internship programs (e.g. Teach for Australia) and various tertiary degree options.	Not found in the review literature* *no alternate pathways to registration were identified on the ECNZ website	Yes. School Direct in-school traineeship, Teach First Scheme fast-track traineeship for academic high-achievers.
	Requirements for ongoing registration	Timing and requirements differ from region to region	Renewal every three years dependent upon character checks, PD, and meeting Practicing Teacher	Not found in the review literature

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		AUSTRALIA	NEW ZEALAND	ENGLAND
			Criteria	
EVALUATION FRAMEWORK	Formally articulated framework in place?	No – Broad national and state guidelines only	Yes – Somewhat limited depth and limited enforcement	No – Broad national guidelines only
EVALUATION FRAMEWORK METHODS AND SPECIFIC EVALUATION TOOLS	Responsible body (State, National, etc.)	State/territory provide recommendations; ultimately individual schools	Individual school Board of Trustees. National guidelines are provided	Individual schools. Following Federal legislation
	Frequency of evaluation	Varies between states/territories. Typically annual	Annual	Annual
	Inspectorate	No	Yes – Conducted by the Education Review Office (ERO)	Yes – Conducted by the Office for Standards in Education, Children's Services and Skills (Ofsted)
	Student outcomes used	Yes – Learning, engagement and wellbeing are considered	Not found in the review literature	Not found in the review literature
METHODS AND SPECIFIC EVALUATION TOOLS USE OF EVALUATIONS	Observation	Yes - Implied importance, no specific rubric provided	Yes – Rubrics provided for English and Maori students	Yes – Implied importance, no specific rubric provided
	Survey ratings	Yes – Student, peer and parent measures are considered	Not found in the review literature	Not found in the review literature
	Teacher Self-evaluation	Yes – AITSL self-evaluation tool	Yes	Not found in the review literature
	Internal/External Evaluators	Internal. Generally principal	Both – Generally Principal or delegated to senior staff, but ERO also conducts informal reviews – not in depth	Both – Head Teacher in conjunction with external evaluators
	Aggregation and Weighting of Results	Not found in the review literature	Not found in the review literature	Not found in the review literature
	Validity measures	Not found in the review literature	Not found in the review literature	Not found in the review literature
	Professional development	Yes	Yes	Yes
USE OF EVALUATIONS	Student assignment	No	No	No
	Career progression	Yes – Only a small influence	Yes – Only a small influence	Yes

		SCOTLAND	ONTARIO	BRITISH COLUMBIA
STANDARDS AND THEORY	Who sets the standard?	Scotland Government, the General Teaching Council for Scotland (GTCS)	Ontario Ministry of Education and Ontario College of Teachers	British Columbia Ministry of Education

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		SCOTLAND	ONTARIO	BRITISH COLUMBIA
	Do standards vary for different career levels?	Yes – Levels include: Registration, Career-Long Professional Learning, and Leadership and Management	Not found in the review literature	Not found in the review literature
	What theoretical framework underpins teacher standards?	Standards and evaluation are driven by the *How good is our school?* publications designed by expert review	Standards informed by emergent themes from a two-year study of teacher standards from across the world	Not found in the review literature
TEACHER PREPARATION	Accreditation of ITE	Teaching programs must be accredited by the GTCS, and must be re-accredited every six years		Accreditation of teacher programs is the responsibility of The British Columbia Teaching Council
	Probation/Provisional teaching status	Yes	Transitional Certificates of Qualification and Registration are available to teachers who have not completed their education program	No – but there are several qualification certificates available
REGISTRATION	Registering Body	National - GTCS	Provincial - Ontario College of Teachers	Provincial – Teacher Regulation Branch, British Columbia Ministry of Education
	Alternate pathways to registration available	Not found in the review literature* *no information about alternate pathways was identified on the associated websites	No; however some who do not meet requirements may qualify for conditional registration, e.g. First Nations persons	Partially. Conditional teaching registration available for a maximum of 60 months while candidates participate in additional learning to gain full registration
	Requirements for ongoing registration	Not found in the review literature	Not found in the review literature	None. Full registration via a Professional Certificate does not expire
EVALUATION FRAMEWORK	Formally articulated framework in place?	No formal framework. Individual schools practice voluntary Validated Self-Evaluation; does not focus on teachers	Teacher Performance Appraisal System set by the Ministry of Education	No formal framework. Evaluations are agreed upon between the teachers' union and the school board.
	Responsible body (State, National, etc.)	Local Authorities. National recommendations for VSE and PRD provided	Provincial – Ontario Teacher Performance Appraisal System	School Districts
	Frequency of evaluation	Not found in the review literature	Beginners - Twice each year	Generally, every 3 years

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		SCOTLAND	ONTARIO	BRITISH COLUMBIA
			Experienced – Once every 5 years	
	Inspectorate	Yes – guidelines given by Education Scotland	Yes – Conducted by Principal or Senior Staff member	No
METHODS AND SPECIFIC EVALUATION TOOLS	Student outcomes used	No. Suggested use in self-evaluation process	Potential source of evidence for evaluation	No
	Observation	Suggested use in self-evaluation	Yes	Yes
	Survey ratings	Not found in the review literature	Yes – not mandatory	No
	Teacher Self-evaluation	Encouraged by GTCS	Yes – not mandatory	Yes
	Internal/External Evaluators	Internal and External – individual teachers and Education Scotland	Internal – Principal	Internal – Principal
	Aggregation and Weighting of Results	Not found in the review literature	Not found in the review literature	No – deemed satisfactory or unsatisfactory at principal's discretion
	Validity measures	Not found in the review literature	Not found in the review literature	Not found in the review literature
USE OF EVALUATIONS	Professional development	Yes	Yes	Not found in the review literature
	Student assignment	No	No	Not found in the review literature
	Career progression	Yes	Yes	Yes

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		ALBERTA	GERMANY	AUSTRIA
STANDARDS AND THEORY	Who sets the standard?	Ministry of Education	National	Each Province (German: Land).
	Do standards vary for different career levels?	Yes – interim certification versus permanent certification. Also vary depending on the school type and subject	No. But they do vary depending on the school type and subject	No. But they do vary depending on the school type and subject
	What theoretical framework underpins teacher standards?	Not found in the review literature	Not available in English	Not found in the review literature
TEACHER PREPARATION	Accreditation of ITE	Ministry of Education	Each State's Kultusministerkonferenz is responsible for teacher training	ITE content is nationally legislated
	Probation/Provisional teaching status	Interim Professional Certificate – must be held for at least two years	Yes	Yes. Approximately 2 years
REGISTRATION	Registering Body	Provincial - Ministry of Education through the Office of the Registrar	No registration is required upon completion of ITE	No registration is required upon completion of ITE
	Alternate pathways to registration available	Not found in the review literature	No	No
	Requirements for ongoing registration	None. Full registration via a Permanent Professional Certification does not expire	None. Qualification is lifelong and does not expire	None. Qualification is lifelong and does not expire
EVALUATION FRAMEWORK	Formally articulated framework in place?	No formal framework	No formal framework	No formal framework
	Responsible body (State, National, etc.)	School	School and Province	School and Province
	Frequency of evaluation	Upon request or if the Principal has reason to believe teacher not meeting standards	Very rarely, generally only in response to serious complaints	Very rarely, generally only in response to serious complaints
	Inspectorate	No	Not found in the review literature	Not found in the review literature
METHODS AND SPECIFIC EVALUATION TOOLS	Student outcomes used	Yes – not mandatory	Not found in the review literature	Yes – teacher encouraged to reflect
	Observation	Yes – not mandatory	Not found in the review literature	Very uncommon
	Survey ratings	No	Not found in the review literature	Not found in the review literature
	Teacher Self-evaluation	Not found in the review literature	Not found in the review literature	Not found in the review literature
	Internal/External Evaluators	Internal – School Principal	External, in response to complaints. Internal evaluations are encouraged, but not enforced	External, in response to complaints. Internal evaluations are encouraged, but not enforced

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		ALBERTA	GERMANY	AUSTRIA
	Aggregation and Weighting of Results	Not found in the review literature	Not found in the review literature	Not found in the review literature
	Validity measures	Not found in the review literature	Not found in the review literature	Not found in the review literature
USE OF EVALUATIONS	Professional development	Yes	Compulsory, but not linked to evaluation	Compulsory, but not linked to evaluation
	Student assignment	Not found in the review literature	Not found in the review literature	Not found in the review literature
	Career progression	Not found in the review literature	Not found in the review literature	Not found in the review literature

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		SINGAPORE	HONG KONG	SOUTH KOREA
STANDARDS AND THEORY	Who sets the standard?	Ministry of Education Singapore	Education Bureau of Hong Kong	No standards exist
	Do standards vary for different career levels?	No	No	No, but competencies vary depending on the subject and year level
	What theoretical framework underpins teacher standards?	A 2000 review from consultants, teachers, and school leaders to create new frameworks	Not found in the review literature	Not available in English
TEACHER PREPARATION	Accreditation of ITE	None	All teaching programs must be accredited by the Hong Kong Council for Accreditation of Academic and Vocational Qualifications	All teaching programs must be accredited by the Korean Educational Development Institute
	Probation/Provisional teaching status	No	Not found in the review literature	No probation period; however there are two stages of registration
REGISTRATION	Registering Body	Does not have national licensing or registration requirements.	Education Bureau of Hong Kong	Ministry of Education, Science, and Technology (MEST) – however there appear to be no registration processes beyond completion of a certified ITE program
	Alternate pathways to registration available	None	Yes – Non-Graduate Teacher Qualification Assessment (typically for mature age students)	None
	Requirements for ongoing registration	None	Not found in the review literature	None – PD is generally required by school districts
EVALUATION FRAMEWORK	Formally articulated framework in place?	Yes – Enhanced Performance Management System (EPMS)	No formal framework – Education Bureau provides guidelines however	Yes. National Performance Based Incentive Scheme as well as individual school guidelines
	Responsible body (State, National, etc.)	National	National - and individual schools	National – and individual schools
	Frequency of evaluation	Year-long process every year	Decided by schools – one or two years	Annually
	Inspectors	Internal reporting officer, usually Head of Department	No	Yes. External monitoring groups established by each provincial body
METHODS AND	Student outcomes used	Yes	Yes	Yes

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		SINGAPORE	HONG KONG	SOUTH KOREA
SPECIFIC EVALUATION TOOLS	Observation	Yes, potentially	Yes	Yes
	Survey ratings	Not found in the review literature	Not found in the review literature	Yes
	Teacher Self-evaluation	Yes	Yes	Yes
	Internal/External Evaluators	Both	Internal	Both
	Aggregation and Weighting of Results	Determined by individual schools	Not found in the review literature	Results from each evaluator are combined
	Validity measures	Not found in the review literature	Not found in the review literature	Not found in the review literature
USE OF EVALUATIONS	Professional development	Yes	Yes	Yes
	Student assignment	No	No	No
	Career progression	Yes - mandated	Yes	Yes

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		CALIFORNIA	WASHINGTON STATE	VIRGINIA
STANDARDS AND THEORY	Who sets the standard?	State – Department of Education	State – Department of Education	State – Virginia Department of Education
	Do standards vary for different career levels?	No	No	Yes – Adapted to specific subject areas
	What theoretical framework underpins teacher standards?	Not found in the review literature	Not found in the review literature	Standards rooted in empirical research with evidence provided by VDOE
TEACHER PREPARATION	Accreditation of ITE	State – Department of Education, California Commission on Teacher Credentialing	State – Washington State Achievement Council	State – VDOE
	Probation/Provisional teaching status	Yes - Preliminary Credential lasts five years, after which teachers must qualify for Clear Credential	Yes – provisional in the first three years of teaching or in a new district, and in the first year if teacher has 2+ years in Washington State	Not found in the review literature
REGISTRATION	Registering Body	Commission on Teacher Credentialing	Office of Superintendent and Public Instruction	VDOE
	Alternate pathways to registration available	Internships, or special pathways for members of the Peace Corps	Yes, available for career-switchers, district staff, and para-educators	State-certified programs, e.g. Troops to Teachers program
	Requirements for ongoing registration	Passing the Personal Fitness Questions	Completion of National Board assessment or completion of Professional Growth Plans	Licenses must be renewed every five years with testing
EVALUATION FRAMEWORK	Formally articulated framework in place?	No formal framework	State legislature outlining evaluation requirements	Minimum evaluation expectations outlined by the State
	Responsible body (State, National, etc.)	State	State	State - VDOE
	Frequency of evaluation	Yes, variable on teacher experience – not enforced	Once a year, otherwise depending on teacher experience	Probationary teachers annually; teachers under contract every five years
	Inspectors	Not found in the review literature	No – sometimes external evaluator	Principal
METHODS AND SPECIFIC EVALUATION TOOLS	Student outcomes used	Yes	Yes	Yes
	Observation	Encouraged, however it is not outlined in guidelines	Yes – frequency depends on teacher experience	Yes
	Survey ratings	May be used	Not found in the review literature	Yes
	Teacher Self-evaluation	Encouraged	Yes – not mandatory	Yes
	Internal/External	Not specified by the framework	Internal – Principal, or external –	Internal unless external evaluators

21

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		CALIFORNIA	WASHINGTON STATE	VIRGINIA
	Evaluators		designated evaluator	are needed
	Aggregation and Weighting of Results	Not found in the review literature	Yes – Student growth impact rating	Student learning accounts for 40% of evaluation; value-added data another 20%
	Validity measures	Not found in the review literature	Not found in the review literature	Not found in the review literature
	Professional development	Yes	Yes	Yes
USE OF EVALUATIONS	Student assignment	Not found in the review literature	Not found in the review literature	Not found in the review literature
	Career progression	Not found in the review literature	Not found in the review literature	Not found in the review literature

Teacher Effectiveness Systems, Frameworks and Measures: A Review

		WASHINGTON D.C
STANDARDS AND THEORY	Who sets the standard?	State – Department of Education, District of Columbia Public Schools
	Do standards vary for different career levels?	Not found in the review literature
	What theoretical framework underpins teacher standards?	Not found in the review literature
TEACHER PREPARATION	Accreditation of ITE	District of Columbia Office of the State Superintendent of Education (OSSE)
	Probation/Provisional teaching status	Not found in the review literature
REGISTRATION	Registering Body	OSSE
	Alternate pathways to registration available	Not found in the review literature
	Requirements for ongoing registration	Not found in the review literature
EVALUATION FRAMEWORK	Formally articulated framework in place?	Yes – IMPACT system
	Responsible body (State, National, etc.)	State
	Frequency of evaluation	Not found in the review literature
	Inspectors	Internal and External
METHODS AND SPECIFIC EVALUATION TOOLS	Student outcomes used	Yes
	Observation	Yes - Both school personnel and external evaluators
	Survey ratings	No
	Teacher Self-evaluation	No
	Internal/External Evaluators	Yes – Both internal and external
	Aggregation and Weighting of Results	Yes, specified by IMPACT
	Validity measures	Not found in the review literature
	Professional development	Yes
USE OF EVALUATIONS	Student assignment	No
	Career progression	Yes

23

Teacher Effectiveness Systems, Frameworks and Measures: A Review

EXECUTIVE SUMMARY

- Currently, responsibility for evaluation of teachers lies mainly with the states and territories; hence no nationally consistent process exists. All teacher evaluations occur at the school level and are mandated at the state/territorial level with influence from relevant unions in the Enterprise Bargaining Agreements (EBA) process.
- The Australian Professional Standards for Teachers (APST) used to evaluate teachers are established and prescribed nationally through the Australian Institute for Teaching and School Leadership (AITSL). The standards are presented as domains of professional knowledge, practice, and engagement expected of teachers across progressive career stages. While they do not prescribe a specific evaluation approach, they form the basis of teacher evaluation at all career levels across states and territories.
- Guidelines for carrying out evaluations are provided by the Australian Teacher Performance and Development Framework. Recommended tools include student outcomes, observations, student feedback, peer/supervisor feedback, parent feedback, and self-assessment. Tools are selected at the discretion of the individual school.
- Evaluation results are used for registration purposes, but also have influence on decisions regarding career stages or underperformance, as well as influencing school priorities regarding professional development. This makes evaluation results both summative and formative.

INTRODUCTION

In Australia it is compulsory to attend Primary and Secondary school between the ages of six and sixteen. School education (pre-tertiary) lasts for 13 years and typically involves the following stages:

- **Primary School** – Operates for seven years, starting at Foundation through to Year 6 (Ages 5-12 years)
- **Secondary School** – Operates for six years (four years for early school leavers) from Years 7 to 12
    - Junior Secondary School – Operates for four years, from Years 7 to 10 (Ages 12-16)
    - Senior Secondary School – Operates for two years, Years 11 and 12 (Ages 16-18) (Australian Government, 2016)

The Australian Bureau of Statistics reported that between 2014 and 2015 over 3.7 million students attended Australian schools (ABS, 2016a). Of these, 65.2% attended government schools and 34.8% attended non-government schools (ABS, 2016a).

Governance

Australia's education system consists of three tiers, represented as national, state or territory, and school (ABS, 2016b). Schools are divided into two broad categories: public schools (government or state schools) and private schools (Catholic schools and independent schools) (ABS, 2016b). Standards for teaching are outlined by the Australian Federal Government, in conjunction with the AITSL. Regulatory bodies within each state or territory are responsible for the accreditation of initial teacher education programs, registration of teachers, and setting standards for performance review and development. EBAs between each state government education department and its teachers, often negotiated by a representing union, strongly influence the protocols, implementation and outcomes of teacher evaluation. Although there are national and state frameworks in place for teaching standards and evaluation, individual schools are ultimately accountable for administering regular performance and development reviews of their teaching staff.

STANDARDS AND THEORY

The Australian Professional Standards for Teachers (APST) provide a guiding framework for Australian teachers and teaching (AITSL, 2013). Established by the Australian Federal Government in conjunction with AITSL and endorsed by the Ministers for Education in all states and territories, the APST were first implemented in 2013 (AITSL, 2013). The APST provide guidance for best practice for teaching professionals categorised under the three domains of professional knowledge, professional practice, and professional engagement across career stages. These domains inform teacher performance assessments and review targets. These standards are summarised in Table 2.

Table 2 - Australian Professional Standards for Teachers

Domain	Standard
Professional Knowledge	1. Know students and how they learn
	2. Know the content and how to teach it
Professional Practice	3. Plan for and implement effective teaching and learning
	4. Create and maintain supportive and safe learning environments
	5. Assess, provide feedback and report on student learning
Professional Engagement	6. Engage in professional learning
	7. Engage professionally with colleagues, parents/careers and the community

The expectations for teachers in relation to each standard vary as a function of their career stage, with teachers classified as:

- *Graduate*: the teacher has completed a nationally recognised teaching program, qualifying them for provisional registration
- *Proficient*: the teacher has achieved each of the proficiency standards, qualifying for full registration
- *Highly Accomplished*: the teacher is recognised as highly effective in their classroom practice, and also as a contributor to the school community
- *Lead*: the teacher is recognised by colleagues, parents/carers and community members as an exemplary teacher, playing a significant role in the leadership and development of other teachers

TEACHER PREPARATION

Australian teachers are credentialed at the state or territory level and are generally required to undertake a minimum of four years of tertiary study. Essential eligibility requirements include the Australian Tertiary Admission Rank (ATAR) for undergraduates and grade point average (GPA) for graduate entry applicants. Selection criteria and processes are not uniformly applied with initial teacher education providers setting program related goals and priorities to attract particular types of applicants. There are three main pathways to qualify as a teacher:

- A four-year undergraduate teacher education degree (e.g. Bachelor of Education). This is predominantly for Primary school teachers.
- A double-degree program – two-year study of concurrent degrees, one being a teacher education course. Most universities offer this option over four years (e.g. Bachelor of Arts with Bachelor of Teaching). This is predominantly for Secondary school teachers.
- An academic degree, followed by a two-year postgraduate teacher education course (Graduate Diploma in Education or Master of Teaching (Primary and/or Secondary) (Education and Training, 2013).

Secondary teachers may also qualify with a Master of Teaching (Secondary) qualification via two-year internship style programs offered by university providers, such as Teach for Australia.

In order to graduate from Initial Teacher Education (ITE) programs, students must demonstrate that they meet the APST expectations at the *Graduate* level (AITSL, 2016a). This is established via satisfactory completion of an ITE program, allowing for graduation. From the 1st of July 2016, all students enrolled in an ITE course (either undergraduate or postgraduate) are also expected to sit and pass the *Literacy and Numeracy Test for Initial Teacher Education Students* prior to graduation (ACER, 2016). Those graduating in 2016 must pass the test to be eligible for provisional registration. From 2017 onwards, passing literacy and numeracy tests will be required to graduate with a teaching degree. Government policy mandates that all teacher candidates enrolled after July 2016 must pass the test to: 1) be eligible for registration if graduating in 2016; 2) gain a teaching qualification if graduating after 2016 (The Department of Education and Training, 2016).

TEACHER REGISTRATION

There is a nationally consistent approach to registration of Australian teachers. Registration processes and practices feature a common set of elements that benefit teachers by:

- Improving the mobility of teachers throughout the nation
- Requiring the same standards and consistent processes to achieve full registration
- Ensuring that registration is part of a wider framework for teachers' career progression and professional learning guided by the Australian Professional Standards for Teachers (AITSL, 2014)

Graduate teachers apply for provisional registration via state or territory regulators. They generally have up to five years to demonstrate and document their teaching performance against the APST at the *Proficient* level in order to attain full registration. Evaluation protocols vary across the states and territories; however, they typically include combinations of standards-based demonstrations of teaching performance, evidence-based classroom inquiry, collegiate observational visits, and teacher reflection as well as panel interviews and recommendations from the principal or nominee. These evaluations evidence *Graduate* teachers as effective members of the teaching workforce and promote them to the next career stage of *Proficient Teacher*. Full registration is offered for a fixed period (maximum of 5 years) and renewable providing requirements for continued professional development and good character checks are met (AITSL, 2016f). Registration must be renewed periodically, but the timing, requirements, and process of renewal changes depending on jurisdiction and local teaching authority (AITSL, 2016f).

COURSE ACCREDITATION

Australia has a national approach to accreditation of ITE programs, using nationally agreed Standards and Procedures to ensure that teachers are prepared at a consistently high standard (AITSL, 2015). The jurisdictional teacher regulatory authority in each state or territory accredits ITE programs.

The regulatory bodies for course accreditation and teacher registration are:

- Victoria - Victorian Institute of Teaching
- New South Wales - Board of Studies, Teaching and Educational Standards
- Queensland - Queensland College of Teachers
- Australian Capital Territory - Teacher Quality Institute
- Tasmania - Teachers Registration Board of Tasmania
- South Australia – Teachers Registration Board of South Australia
- Western Australia - Teacher Registration Board of Western Australia
- Northern Territory - Teacher Registration Board of the Northern Territory

To be accredited, ITE providers are required to demonstrate that their program ensures pre-service teachers possess the knowledge and skills necessary for them to be effective classroom teachers by graduation. Recent reforms

provide clarity and rigour (AITSL, 2016e), with ITE providers required to demonstrate that they meet the nationally agreed *Standards and Procedures* articulated by AITSL in *Accreditation of Initial Teacher Education Programs in Australia* (AITSL, 2015). This includes plans for collection of evidence of the impact of their program, documented against the APST at the *Graduate* level (AITSL, 2013). This evidence forms part of a submission for annual review or audit by the respective regulatory body. While the specific demands for pre-service primary and secondary school teachers vary, all courses must align with the APST.

EVALUATION FRAMEWORK

There is no nationally enforced teacher evaluation framework in Australia. Instead, the Federal government provides recommendations for managing teacher performance and development. Using the APST as a framework, the process is outlined in the *Australian Teacher Performance and Development Framework* (AITSL, 2013). This Framework describes evaluation of teacher performance as an *ongoing cycle*, comprising three key stages:

1. *Reflection and Goal Setting*: teachers are required to establish and regularly review classroom or practice-specific goals relative to the APST, in consultation with their principal or delegate.
2. *Professional Practice and Learning*: information about teacher effectiveness is collected from a range of sources, including: student outcome data, direct observation and evidence from colleagues. This information is used to reflect upon and evaluate teacher performance.
3. *Feedback and Review*: teachers should receive regular formal and informal feedback on their performance. Formal reviews of teacher performance in terms of their development goals should occur annually, with teachers given both verbal and written feedback.

This national framework is not consistently adopted across Australia. The states and territories provide their own frameworks and recommendations for performance and development reviews that reflect the specific EBAs struck between each state or territory's schools and employed teachers. While state and territory education departments generally imply that evaluations are mandatory, only Western Australia and New South Wales have mechanisms in place to ensure that evaluations occur. Most states and territories require annual appraisals, with some specifying a time frame (e.g., in New South Wales and Australian Capital Territory they must occur in term four; in Victoria they take place in April/May). While both federal and state bodies provide recommendations for teacher evaluation, individual schools are ultimately responsible for the performance development review of their teachers.

METHODS AND METHODOLOGIES

There is currently no consistent national evaluation system. Protocols and specific tools for teacher appraisal differ between states and territories. Typically, teachers are expected to meet the standards set out in the APST, usually assessed by a member of the Principal class. The *Australian Teacher Performance and Development Framework* (ATPDF) provides National guidance for evaluation protocols (AITSL, 2013), including recommendations that evaluation data come from multiple sources. These include but are not limited to:

- Student outcomes
- Direct observation of teaching
- Student feedback
- Peer/supervisor feedback
- Parent feedback and
- Teacher self-assessment

The appropriateness of these sources is determined by the school context and teacher expertise. Student outcomes, direct observation, and evidence from colleagues are listed as the minimal number of sources.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Student Outcomes

The national framework (APTDF) describes student outcomes as learning, engagement in learning and wellbeing, and acknowledges that these can be measured in numerous ways. While it is clear from the framework this is not intended as the sole measure for evaluating teacher effectiveness, it is also clear that it is a significant determining factor.

Direct Observation of Teaching

Both national and state/territory frameworks for teacher performance and development emphasise the importance of classroom observation in teacher performance assessments. The national framework states that teachers must be observed in their classrooms by their principal, mentor or another assessor; however, specific constructs or tools to support observations are not provided. Teachers applying for certification as *Highly Accomplished* and *Lead Teachers* must be observed by trained external assessors, who are teachers, school leaders, and other individuals with the requisite expertise. Both New South Wales and the Northern Territory require classroom observations to be a part of the assessment. In New South Wales a minimum of two observations are required per evaluation period, while the Northern Territory requires a minimum of three per evaluation period.

Survey Ratings

Survey ratings from students, peers, and parents may be included in evaluations; however, these are considered non-essential elements.

Teacher Self-Assessment

Australian teachers are encouraged to regularly and systematically self-assess as reflective practitioners. AITSL supports this with a selection of online tools comprising the Teacher Toolkit (AITSL, 2016d). An exemplar from the suite is the *Reflection on Practice Tool* (AITSL, 2016c), which requires teachers to consider how their practice best aligns with focus area descriptors that are de-identified to minimise the influence of career stage or dimensional categories.

Internal/External Evaluators

The principal generally carries out school-based evaluations internally. External evaluators may be employed if there is a conflict of interest or if the teacher contests the evaluation results. External assessors are essential to the certification process for *Highly Accomplished* and *Lead* teachers.

Aggregation and Weighting of Results

Not referenced in reviewed policy documents.

USES OF EVALUATION

As previously noted, evaluation results are used for registration purposes. University-set tasks are used to evaluate whether beginning teachers meet the *Graduate* teacher standards. Once a beginning teacher graduates, they are eligible to register as a Provisionally Registered Teacher (PRT). PRTs are required to demonstrate their ability to meet the APST at a *Proficient* level in order to attain full registration, usually within a period of 2-5 years depending on the state/territory in which the teacher is registered (Victorian Institute of Teaching, 2016a). This registration process varies across the states and territories, typically involving combinations of standards-based demonstrations of teaching performance, evidence-based classroom inquiry, observational collegiate visits, and teacher reflection. Notes from the graduate teacher's principal or nominated supervisor are key inputs to this process. These registration processes deem graduate teachers proficient and effective members of the teaching workforce.

School-based evaluation results have a minor influence on decisions about career stage promotion or underperformance. In New South Wales, Victoria, and Tasmania, school-based evaluations are linked to salary decisions; however, the weight of these evaluations is generally not significant. While identifying and supporting underperforming teachers is frequently cited as a central aim of school-based evaluation in Australia, it is rarely used for this purpose. Instead, processes typically exist externally to the broader performance and development process.

All states and territories require schools to use data and feedback emanating from school-based evaluation processes to determine school priorities and professional development at the individual, school, and professional learning community levels. Schools generally retain the evaluation documentation, with the exception of NSW, where the process is more formal and evaluation documents are submitted to the education department.

CONSIDERATIONS FOR MINORITY GROUPS

Aboriginal and Torres Strait Islander students numbered 200,563 persons in 2015-2016. This comprised approximately 5.3% of Australian students (ABS, 2016a). The Australian Professional Standards for Teachers address strategies for teaching students from diverse backgrounds, with specific expectations varying as a function of career level (AITSL; 2013). Specifically, standards 1.3, 1.4, 2.3, and 2.4 illustrate the expectation of teachers working with diverse linguistic, cultural, religious, and socioeconomic backgrounds, as well as Aboriginal and Torres Strait Islander students specifically (AITSL; 2013). While these standards provide an illustration of the expectations of teachers, they provide no guidance in the consideration of such groups when evaluating teacher effectiveness.

EXECUTIVE SUMMARY

- An explicit and mandatory national framework for teacher evaluation exists in the Guidelines on Performance Management Systems. The Board of Trustees at each school implements the evaluation processes outlined in the Guidelines through the principal or nominee. There is no centralised training or evaluation of appraisers.
- New Zealand has a strong theoretical framework and strong guiding standards for teacher practice in the Practising Teacher Criteria. Evaluation is strongly linked to these standards. There is less emphasis on mechanisms to ensure that the standards used are applied rigorously, fairly, and consistently across the profession.
- At the school level methods vary for providing accurate and reliable assessment of teaching performance against the Practising Teacher Criteria. Evaluation can rely on subjective interpretation of the criteria.
- Evaluations are linked to salaries and promotion or dismissal; however, this has elements of subjectiveness in light of the evaluation processes. Teacher feedback processes are not defined.

INTRODUCTION

In New Zealand Primary or Secondary schooling is compulsory between the ages of six and sixteen; however, most children begin at five years of age. School education (pre-tertiary) lasts for 13 years and is divided into:

- **Early Learning –** Open to children aged up to 5 years old, divided into Early Childhood Education (ECE) services and kōhanga reo
- **Primary schooling –** Operates for eight years
    - Full Primary Schools: Years 1 – 6 (5-10 years old) or Years 1-8 (5–12 years old)
    - Intermediate Primary Schooling: Years 7 - 8 (11-12 years old)
- **Secondary school –** Operates for five years, for students in Years 9 - 13 (13-19 years old)
    - Further Learning – for students who leave school at 16 years old (Ministry of Education New Zealand, 2016a)

In 2015, 776,815 students attended New Zealand schools (Education Counts, 2016). Of these, 84.8% attended state schools, and 16.2% attended non-state run schools, including private and religious schools (Education Counts, 2016).

Governance

The New Zealand Ministry of Education is the national governing body for education and schooling, with local Ministry Offices governing schools in their constituency. Schools in New Zealand are categorised as:

- *State* – owned and run by the state
- *State-integrated* – state schools with special character, philosophy or religion
- *Maori Medium* schools – state schools following Te Aho Matua philosophy and practices (Education in New Zealand, 2016)

Each school is governed by an established Board of Trustees and may choose either a standard or alternative constitution. Typically the governing board is a collective representation including the principal, a staff representative, five elected parent representatives, a student representative (Year 9+), co-opted trustees, and up to four proprietor's representatives (integrated schools only) (Ministry of Education New Zealand, 2016f). Boards gain approval of their constituency from their local Ministry Office.

Standards for teaching are set by the Ministry of Education in collaboration with the Education Council of New Zealand (ECNZ; formerly the New Zealand Teachers Council [NZTC]). The ECNZ also acts as the national regulatory body, responsible for accreditation of ITE programs, certification of teachers, and setting standards for performance

and development reviews. Collective agreements between teaching unions and the Ministry of Education strongly influence the implementation, methodology and outcomes of teacher evaluation. Separate agreements are in place for primary and secondary school teachers; however, both contain language regarding the role of performance management. Although there are national frameworks in place for teaching standards and evaluation, ultimately individual schools and boards of trustees are accountable for administering regular performance and development reviews of their teaching staff (Ministry of Education New Zealand, 2016b).

STANDARDS AND THEORY

New Zealand has teacher standards for *Graduate* and *Practising* teachers, established by the ECNZ. ITE institutions are required to demonstrate how their program meets the *Graduating Teacher Standards* through program delivery and assessment (ECNZ, 2016c). Graduating teachers are expected to possess skills and knowledge in domains relating to professional knowledge, professional practice, and professional values and relationships. These domains and dimensions are summarised in Table 3.

Table 3 – New Zealand Graduating Teacher Standards

Domain	Standard
Professional Knowledge	1. Graduating teachers know what to teach 2. Graduating teachers know about learners and how they learn 3. Graduating teachers understand how contextual factors influence teaching and learning
Professional Practice	4. Graduating teachers use professional knowledge to plan for a safe, high quality teaching and learning environment 5. Graduate teachers use evidence to promote learning
Professional Values and Relationships	6. Graduating teachers develop positive relationships with learners and the members of learning communities 7. Graduating teachers are committed members of the profession

Endorsed by the Ministry of Education, the *Practising Teacher Criteria* were developed by the NZTC in 2006 for use when teachers are gaining or maintaining full certification, to provide a valid framework for reflection and to guide the professional learning and development of teachers (ECNZ, 2016e). Criteria and key indicators include professional knowledge, practices, attributes, and values that facilitate academic, social, and cultural learning in diverse education settings. The criteria and key indicators are both discrete and integrated, as summarised in Table 4.

Table 4 - New Zealand Practicing Teacher Criteria

Standard	Criteria
Professional Relationships and Professional Values	1. Establish and maintain effective professional relationships focused on learning and well-being of ākonga* 2. Demonstrate commitment to promoting the well-being of all ākonga* 3. Demonstrate commitment to bicultural partnership in Aotearoa New Zealand 4. Demonstrate commitment to ongoing professional learning and development of personal professional practice 5. Show leadership that contributes to effective teaching and learning
Professional Knowledge in Practice	6. Conceptualise, plan and implement an appropriate learning program 7. Promote a collaborative, inclusive and supportive learning

	environment
	8. Demonstrate in practice their knowledge and understanding of how ākonga* learn
	9. Respond effectively to the diverse language and cultural experiences, and the varied strengths, interests and needs of individuals and groups of ākonga*
	10. Work effectively within the bicultural context of Aotearoa New Zealand
	11. Analyse and appropriately use assessment information, which has been gathered formally and informally
	12. Use critical inquiry and problem-solving effectively in their professional practice

Note. *Ākonga is the Māori word for learner or student which is preferred by the Education Department

TEACHER PREPARATION

There are 156 approved ITE programmes in New Zealand, delivered as 80 qualifications by 25 providers (ECNZ, 2016d). Teachers can choose from a range of approved programs depending upon the sector, such as an Undergraduate degree (3 or 4 years/Primary and Secondary), Undergraduate diplomas (3 years/Early Childhood only), or a one-year graduate diploma to achieve a qualification at Level 7 on the New Zealand Qualifications Authority (NZQA) register of Quality Assured Qualifications (2016) (ECNZ, 2016g).

TEACHER REGISTRATION

After gaining a teaching qualification, graduate teachers apply for provisional certification and a current practicing certificate or authorisation, issued by the national regulatory body, the ECNZ (Education Council New Zealand, 2016b). After a period of induction and mentoring (minimum of two years, maximum of six) teachers are appraised by a mentor/professional leader against the ECNZ *Practicing Teacher Criteria* (ECNZ, 2016e). Evidence must be provided that these standards for quality teaching have been achieved for teachers in New Zealand to gain full certification. This certification must be renewed every three years and is contingent upon teachers maintaining character checks, completing *satisfactory* professional development and having been successfully appraised as meeting the *Practicing Teacher Criteria* during recent service (ECNZ, 2016f).

COURSE ACCREDITATION

The ECNZ accreditation panel approves programs in the university sector. Programs in the non-university sector require a joint panel with New Zealand Qualifications Authority (NZQA). Providers must meet the documented program requirements as outlined in Approval, Review and Monitoring Processes and Requirements for Initial Teacher Education Programmes (ECNZ, 2015) and demonstrate how they meet the Graduating Teacher Standards (ECNZ, 2016c) in program delivery and assessment. Initially programs are monitored annually then biannually after either three or four years, with written reports submitted by the monitor (ECNZ, 2015). After six years of delivery, programs undertake panel reviews mainly for minor changes and future proposals (ECNZ, 2015). While the specific requirements vary between sectors, all programs must align with the Graduating Teacher Standards as an assurance of the quality of all graduates from all ITE programs (ECNZ, 2015).

EVALUATION FRAMEWORK

New Zealand has a prescribed national framework for School-based teacher evaluation, which is set by the Secretary for Education, and outlined in the *Guidelines on Performance Management Systems* (Ministry of Education New Zealand, 2016c). This framework outlines teacher evaluation processes that are implemented by the School Boards of Trustees through the school principal or nominated evaluator (e.g. senior teacher). The *Guidelines on Performance Management Systems* link the identification of expected performance of individual teachers to both

32

Teacher Effectiveness Systems, Frameworks and Measures: A Review

the minimum requirements for teacher registration and the Professional Standards contained in the Primary (Ministry of Education New Zealand, 2016d) and Secondary Teachers Collective Employment Contracts (Ministry of Education New Zealand, 2016g). However, the standards are intended to be broad to allow the school's Boards of Trustees to have flexibility in designing evaluation systems appropriate to their school and community, within a minimum quality assurance and accountability framework (Ministry of Education New Zealand, 2010, 2016c).

In constructing and operating *Performance Management Systems*, school Boards of Trustees are required to identify both performance expectations and development objectives based on key performance areas and/or key performance responsibilities for each teacher (Ministry of Education New Zealand, 2010, 2016c). These responsibilities are tied to the criteria by which ECNZ deems a teacher 'satisfactory' (Practicing Teacher Criteria) (Ministry of Education New Zealand, 2010, 2016c). This process has a degree of oversight, with the Education Review Office (ERO) conducting school evaluations in collaboration with regional Boards of Trustees (Education Review Office, 2014). In addition to these processes, the ECNZ is also currently engaged in a multi-phase project examining the role of teacher appraisal in schools (ECNZ, 2016a).

METHODS AND METHODOLOGIES

Typically, teachers are expected to meet the standards set out by the ECNZ, usually assessed by a member of the *school principal or nominated evaluator*. The *Guidelines on Performance Management Systems* provide national guidance for evaluation protocol, including recommendations that evaluation data come from multiple sources. These include but are not limited to:

- Direct observation of teaching
- Teacher Self-assessment
- Internal or external evaluators

Student Outcomes

Not referenced in reviewed documents.

Direct Observation of Teaching

The observation of teaching is a critical component of the school based evaluation process (Ministry of Education New Zealand, 2016c). The specific responsibilities of those involved with evaluation are outlined by the Ministry of Education's performance management guidelines (Ministry of Education New Zealand, 2016c).

Survey Ratings

Not referenced in reviewed documents.

Teacher Self-Assessment

Teacher self-assessment is another critical component of school-based evaluation and is identified as a teaching responsibility in the Ministry of Education's performance management guidelines (Ministry of Education New Zealand, 2016c). No further detail, beyond the importance of including self-appraisal by the teacher, was provided in the *Guidelines on Performance Management Systems*.

Internal/External Evaluators

In most schools, teacher evaluations will be the responsibility of experienced teachers and managers (Ministry of Education New Zealand, 2010). This, however, may vary depending on the size of the school. In smaller schools the role of evaluator is likely to be the responsibility of the principal, while in larger schools other senior staff may be

Teacher Effectiveness Systems, Frameworks and Measures: A Review

appointed to the evaluator role by the principal (Ministry of Education New Zealand, 2010). No special training is required or given to evaluators.

The *Guidelines on Performance Management Systems* require that an evaluator is a professionally competent person (Ministry of Education New Zealand, 2016c). However, ensuring the quality and consistency of school-based evaluation of teachers has been an ongoing challenge in New Zealand (Ministry of Education New Zealand, 2010). No formal appraisal of evaluation is carried out. The Education Review Office (ERO) conducts regular school reviews, and has the authority to access all school documentation, including information on teacher appraisal and professional development plans (Education Review Office, 2014; Ministry of Education New Zealand, 2010). The ERO also looks at teacher appraisal documents and professional learning plans within the context of school-wide planning and decision making (Education Review Office, 2014; Ministry of Education New Zealand, 2010). Reports do not assess individual teachers but do comment on teacher quality overall and areas that need attention (Education Review Office, 2014; Ministry of Education New Zealand, 2010). There are few mechanisms to ensure that the standards used are applied rigorously, fairly, and consistently across the profession. Additionally, the inclusion of different stakeholders in this process, such as unions, teaching organisations, and parents (either directly or indirectly) in the teacher evaluation process causes challenges in determining desired outcomes and implementation processes. While the ERO has some oversight, teacher evaluation only forms a small part of their review process, and is therefore not always a priority for review.

Aggregation and Weighting of Results

Not referenced in reviewed documents.

USES OF EVALUATION

Appraisal results are used to determine recommendations in relation to the application for full registration and renewal of practicing certificates. As per the Primary Teachers' Collective Agreement, annual appraisal results guide progression on the salary scale (Ministry of Education New Zealand, 2015a). The Secondary Teachers' Collective Employment Contract contains less explicit guidelines in this regard, requiring teachers to be working towards high competence and quality, which may be evidenced by performance appraisals; however, this is not explicitly stated (Ministry of Education New Zealand, 2015b). The current scale has 14 steps and the initial placement on that scale is dependent on qualification levels. Satisfactory teacher appraisal results lead to further progression up this scale. In practice, teachers usually progress up the salary scale each year, with appraisal results only providing a small influence on this decision. If performance shortfalls have been identified, then professional development opportunities can be implemented to redress these issues. If serious concerns are identified, teachers are given a period of 10 weeks to remedy the matter of concern through appropriate assistance and guidance programs. Failing to do so may result in dismissal.

CONSIDERATIONS FOR MINORITY GROUPS

The Ministry of Education has developed *Ruia: Teacher Appraisal for Māori learners' success*, which provides guidance in the use of teacher appraisal as a tool for improving Māori student outcomes (Ministry of Education New Zealand, 2016e). *Ruia* provides a comprehensive set of appraisal guidelines, tools and exemplars for improving Māori student outcomes based on research and confirmed through exploratory work with schools that had made a difference for Māori students (Ministry of Education New Zealand, 2016h). A system of "appraisal for learning" is encouraged, whereby teachers learn about their effectiveness as it relates to student outcomes. The interactive tool (http://appraisal.ruia.educationalleaders.govt.nz/Reviewing-appraisal) helps school leaders examine their current appraisal processes to identify their schools' strengths and needs in relation to teacher appraisal that improves outcomes for Māori students. The *Ruia* framework is specifically targeted towards Māori students and does not address considerations for other minority groups.

EXECUTIVE SUMMARY

- There is no nationally enforced teacher evaluation system in England. Schools are guided by a national framework and legislation that encourage evaluation on the basis of teaching standards; however evaluations are conceived and implemented exclusively at a school level. School head teachers are expected to perform teacher evaluations; however, this may be delegated to any qualified teacher.
- Teaching standards are set at a national level in the form of the Teachers' Standards, and represent minimum expected practice organised around two aspects: teaching and personal/professional conduct. The teaching standards are intended to drive aspects of appraisal content but appraisals must also relate to the specific context and role of the teacher.
- The national framework and legislation suggests that teacher evaluation methods should include student outcomes, direct observation, and peer ratings.
- Teacher evaluations can be used for summative purposes, including career progression and occasionally dismissal. Results are used formatively in the form of written feedback and instructions to develop certain needed areas of teaching practice. The weighting of teacher evaluations is likely very low, as a systematic approach is not implemented between schools.

INTRODUCTION

Schooling in England is compulsory from primary school through to the end of secondary school (Department for Education, 2016). Schooling in England is organised into:

- **Primary School –** Operates for seven years, which students attend between the ages of 5-11 (occasionally separated into infant and junior schools)
- **Secondary –** Operates for six years, which students attend between the ages of 11-16
- **Further Education –** Students attend between 16-18 (delivered by schools, sixth form colleges, further and higher education colleges and private training providers)

In 2015, 8.4 million students were enrolled in English schools (Department for Education, 2015), with approximately 7% of these students attending independent schools (Independent Schools Council, 2016). The proportion of students at private schools increases with age, with 14% of students over 16 attending a private school, compared to 3% of four-year-old students (Independent Schools Council, 2016).

Governance

The United Kingdom Department for Education oversees education in England, with the Office for Standards in Education (Ofsted) monitoring educational services in public schools and the Independent Schools Inspectorate responsible for private schools (Office for Standards in Education, 2016). Local government authorities are responsible for ensuring that national legislation is adhered to, and have some authority in establishing educational policies at the local level. Although schools must follow national regulations concerning teacher evaluation, these are very general, and they are largely independent in conducting appraisals.

STANDARDS AND THEORY

Introduced by Sir Michael Barber, the English education system has a set of standards with a particular emphasis on the importance of teacher effectiveness in determining students' learning outcomes (Barber & Mourshed, 2007). The existing Teachers' Standards came into place in 2011, following a government commissioned report evaluating the previous standards (Department for Education, 2011a). The standards outline the minimum practice expected of teachers awarded Qualified Teacher Status (QTS), and are organised into two parts, summarised in Table 5 (Department for Education, 2011b):

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Table 5 - English Teachers' Standards

Standard	Criteria
Teaching	1. Set high expectations which inspire, motivate and challenge pupils 2. Promote good progress and outcomes by pupils 3. Demonstrate good subject and curriculum knowledge 4. Plan and teach well-structured lessons 5. Adapt teaching to respond to the strengths and needs of all pupils 6. Make accurate and productive use of assessment 7. Manage behaviour effectively to ensure a good and safe learning environment 8. Fulfil wider professional responsibilities
Personal and Professional Conduct	9. Teachers must uphold public trust in the professions and maintain high standards of ethics and behaviour, within and outside the school. 10. Teachers must have proper and professional regard for the ethos, policies, and practices of the school in which they teach and maintain high standards in their own attendance and punctuality. 11. Teachers must have an understanding of, and always act within, the statutory frameworks, which set out their professional duties and responsibilities.

TEACHER PREPARATION

To teach in England, teachers are required to gain Qualified Teacher Status (QTS) (Department for Education, 2014). There are several pathways available for obtaining QTS, including university-led training or school-led training programs (Department for Education, 2016b). Universities and colleges offer either three or four year undergraduate Bachelor of Education programs, or at the graduate level a one year Post Graduate Certificate in Education (PGCE). School Direct in-school training is an alternate pathway through partnership between schools and accredited Initial Teacher Training providers (Department for Education, 2016a). The Teach First Scheme is an option for high achieving qualified applicants to teaching (TeachFirst, 2016).

TEACHER REGISTRATION

Graduates are awarded QTS by the National College for Teaching and Leadership (NCTL) as the competent authority (NCTL, 2016). In order to become a registered teacher, graduates, referred to as Newly Qualified Teachers (NQT) are assessed during a one-year induction period, during which they are mentored at the school level and provided with professional development by the local education authority (Department for Education, 2012). The revised DfE Teachers' Standards (2012) are used by head teachers and appraisers to assess NQTs completing their induction period, and thereafter at a level consistent with their relevant role and career stage (Department for Education, 2012).

COURSE ACCREDITATION

Initial Teacher Training providers are required to design and deliver programs that are consistent with the revised DfE Teachers' Standards (2011). Trainee Teachers are assessed against these standards to ensure that they meet the expected level prior to being awarded QTS (Department for Education, 2013).

The NCTL is the regulatory body responsible for accrediting initial teacher training (ITT) providers, who must demonstrate a ranking of *at least good* to Ofsted against the Secretary of State's ITT criteria (Office for Standards in Education, 2015). Ofsted conducts inspections of all providers and programs for quality assurance. These inspections typically occur twice annually across a six-year cycle (Office for Standards in Education, 2015).

Teacher Effectiveness Systems, Frameworks and Measures: A Review

EVALUATION FRAMEWORK

Individual teacher evaluation in England is guided by *The Education (School Teachers' Appraisal) (England) Regulations 2012* (The National Archives, 2012). The legislation encourages teachers to be evaluated in relation to the standards above, as well as more targeted objectives that relate to their specific role, experience and school context. All of the evaluation criteria must be discussed and agreed upon by the evaluator and the staff member at the beginning of the appraisal period. Teachers are evaluated by the head teacher or nominated by the head teacher. These formal evaluations are conducted annually; however, teachers should be informally evaluated more regularly, and provided with ongoing feedback and support (Department for Education, 2012). It should be noted that the evaluation guidelines are broad and allow schools to establish their own systems of evaluation. Due to this, there is a high degree of variance of teacher evaluations between and within schools.

This evaluation system is different from the evaluation of schools and Head teachers through the Ofsted body. School evaluation involves the direct observation of a random assortment of teachers by both the external inspector and someone from within the school, most commonly the head teacher (Ofsted, 2015). The inspector also observes the Head Teacher (or other) giving feedback, with the school's leadership and guidance also considered part of the overall evaluation. While this is an attempt to evaluate teaching and learning in the school more broadly, it also acts to analyse the contribution that leadership specifically provides for the school culture.

METHODS AND METHODOLOGIES

The Education (School Teachers' Appraisal) (England) Regulations 2012 provides little direction as to the specific methods for evaluating teachers (The National Archives, 2012). Evaluators are given discretion regarding the selection of evaluation tools, determined by the school context and teacher expertise. Methods suggested as a minimum recommendation for evaluation include:

- Student outcomes
- Direct classroom observation
- Peer feedback

Student Outcomes

Not referenced in reviewed documents.

Direct Observation of Teaching

While evaluators are encouraged to observe teachers in their classrooms, there are only very general guidelines as to how observations should be executed. Regular teacher observations are required, with the amount and type of observation determined by individual teacher and school needs. Formal observations are prescribed along with informal 'drop-ins' by those in leadership positions to ensure that high standards are implemented and maintained (Department for Education, 2012).

Survey Ratings

Not referenced in reviewed documents.

Teacher Self-Assessment

Not referenced in reviewed documents.

Internal/External Evaluators

The Department of Education model teacher appraisal policy suggests that the head teacher has the responsibility of deciding who appraises teachers in their school, alongside an external evaluator (Department for Education, 2012). An external evaluator also appraises the head teacher. Similarly, the *Education (School Teachers' Appraisal) (England) Regulations 2012* act states that the head teacher must appraise teachers within their school (The National Archives, 2012).

Aggregation and Weighting of Results

Not referenced in reviewed documents.

USES OF EVALUATION

The purpose of England's teacher evaluation system is both formative and summative in nature. Evaluation outcomes are used to hold teachers accountable through linkages to career progression and penalties, the latter leading to dismissal in some cases. In instances where there are concerns about a teacher's performance, the Head Teacher is required to hold a formal meeting to identify areas of challenge for the teacher and establish an improvement plan. Teachers should then be monitored, and given high levels of support to improve their teaching. However, failure to improve within a set period of time may lead to dismissal.

Following the evaluation, all teachers should be provided with a written report outlining their professional development needs and specific instructions of how these can best be met. There is also a strong emphasis on teachers using their appraisal feedback to improve and develop their teaching practices.

CONSIDERATIONS FOR MINORITY GROUPS

In England there are some concerns regarding under-representation of Black and Ethnic Minority (BME) groups in teaching, with calls since the early 2000's to ensure greater participation (Hargreaves, 2011). The limitation of diversity in the teaching population in England does not mirror the student population with one in four children belonging to BME groups (National Archives, 2009). Furthermore, BME teachers are often disproportionately located in low socioeconomic regions and dealing with complex social issues within the classroom. Despite these issues, there are no specific aspects of the teacher evaluation system that cater for the context that BME teachers work within.

EXECUTIVE SUMMARY

- Scotland does not have a nationally enforced framework for teacher evaluation. Schools are expected to take responsibility for their own educational quality, including teacher evaluation processes. Schools and teachers are encouraged to engage in Validated Self-Evaluation (VSE); however, this is voluntary and not enforced at any level.
- There are national guidelines for Professional Review and Development, which involve regular professional learning and self-evaluation against the standards for teaching, which are set at a national level. Scottish teachers must provide evidence of these activities as part of their mandatory reaccreditation process every five years.
- There are no specific or mandatory tools for evaluation, and no information was found regarding commonly used methods of appraisal. Self-assessment is encouraged.
- Evaluation results appear to mainly serve the attainment or maintenance of registration.

INTRODUCTION

Public education in Scotland is free from 5 to 19 years old, and it is compulsory to attend school until the age of 16. School education is organised as follows:

- **Primary School** – Operates for seven years, from P1 to P7
- **Secondary School** – Operates six years, from S1 to S6

The Scottish education framework is separate from the rest of the United Kingdom (England, Wales and Northern Ireland). Scotland follows the *Curriculum for Excellence* for primary and secondary schools (Education Scotland, 2016a). This was a significant educational reform, introduced in 2010-2011, allowing teachers and students more freedom over the content and delivery of education in Scottish classrooms (Education Scotland, 2016a). While there are still three core subjects that must be taught (health and wellbeing, literacy, and numeracy), schools are free to select the content they wish to teach and determine the best way of connecting material for consolidating learning (Education Scotland, 2016a).

As of 2015, 673,087 children were enrolled in Scottish primary and secondary schools (The Scottish Government, 2015). Of these, approximately 4.3% attended independent schools (Scottish Council of Independent Schools, 2016). No current statistics were available regarding the enrolment of children in Catholic schools; however, this may be estimated at roughly 15% based on 2013 statistics (The Scottish Government, 2013).

Governance

Scotland's public education system can be divided into three tiers of governance, summarised as National, Local Authorities, and School. At a National level, the General Teaching Council for Scotland (GTCS) sets teacher standards, while the Scottish government sets school curriculum and provides guidelines for Professional Review and Development for teachers. Thirty-two Local Authorities (or councils) in Scotland act as Education Authorities, who own and operate all schools within their council boundaries (GTCS, 2016e). Consistent with the Scottish approach to improving schools, responsibility for teacher evaluation ultimately lies with the individual school and the individual teacher. Teaching trade unions, such as the Educational Institute of Scotland (EIS), which represents roughly 80% of teachers in Scotland, also play a large role in monitoring and guiding the implementation of evaluation practices via collective bargaining agreements (EIS, 2015).

STANDARDS AND THEORY

The Scottish education system has a set of professional standards for teachers, set by the GTCS. Different standards are offered for Registration (GTCS, 2012c), Career-Long Professional Learning (i.e. registered teachers; GTCS, 2012a),

39

Teacher Effectiveness Systems, Frameworks and Measures: A Review

and Leadership and Management (i.e. middle and head teachers; GTCS, 2012b). While the specific standards differ by career level, the domains of Professional Values and Personal Commitment, Professional Knowledge and Understanding, and Professional Skills and Abilities underpin all. In addition to these domains, the Professional Standards for Leadership and Management include Strategic Knowledge and Interpersonal Skills and Abilities (GTCS, 2012b). Specific professional actions for Middle Leaders and Head Teachers are also outlined (GTCS, 2012b). The standards are designed for a range of uses including individual reflection and self-evaluation, gatekeeping, and guiding registration, in addition to providing support for school staff in performance development processes. In this respect, the standards are not dissimilar to the current Australian standards set by AITSL.

TEACHER PREPARATION

There are two pathways into teaching in Scotland. Prospective primary or secondary teachers may undertake either a four-year undergraduate program or a one-year Professional Graduate Diploma in Education (PGDE) collectively offered as Initial Teacher Education (ITE) programs by eight universities across Scotland (Scottish Government: Teacher Education, 2015). ITE programs typically include curricular studies, professional studies, educational studies (history and educational theory) and professional placements (Teach in Scotland, 2016).

Entry requirements for all ITE programs are set by the GTCS and articulated in the *Scottish Credit and Qualifications Framework* (SCQF, 2016). Each ITE Provider may set additional requirements, however the common essential elements are:

- A National Qualification Course award in English at SCQF Level 6 (Higher Grade English or an accepted alternative)
- A National Qualification Course award in Mathematics at SCQF Level 5 (National 5 Mathematics or an accepted alternative)
- An appropriate level of literacy and numeracy
- Competence in the use of Information and Communications Technology;
- References evidencing experience, interests and personal characteristics desirable in a teacher (Teach in Scotland, 2016)

Student teachers require a provisional registration after completing their ITE, after which they can apply for full registration. The Standard for Provisional Registration (SPR) describes what is expected of teachers at the end of their ITE before they seek their provisional registration (GTCS, 2016g). Once held, student teachers move towards attaining the Standard for Full Registration (SFR), which sets out expectations for professional qualities and capabilities (GTCS, 2016h).

TEACHER REGISTRATION

Teachers in Scotland are legally required to be registered with the GTCS (GTCS, 2016d). Completion of a recognised SCQF Level 9 (i.e. a Bachelor's degree) or higher teaching degree is required to register (GTCS, 2016d). This can include either the completion of an accredited four-year undergraduate programme or the completion of an accredited one-year PGDE.

Beyond normal registration, teachers can apply for an additional registration that covers the teaching of students three to 18 years of age with additional support needs (GTCS, 2016a). Further, a Professional Registration can be sought which provides specific registration for a teacher to teach in another subject area (GTCS, 2016b).

COURSE ACCREDITATION

Teaching programs must be accredited by the GTCS, and must be re-accredited every six years (GTCS, 2006; GTCS, 2013). Programs are evaluated to determine whether they sufficiently address the specific knowledge, skills and attributes that graduates are expected to possess in relation to the *Professional Standards for Registration* (GTCS, 2012c). Programs can also be accredited conditionally, meaning that they can only continue to educate prospective

teachers if they meet certain standards (GTCS, 2006). In this instance programs are under more scrutiny, and evaluated more regularly (GTCS, 2006).

EVALUATION FRAMEWORK

There is no nationally enforced teacher evaluation framework in Scotland. Instead Scotland's education system is built on the premise that schools must take responsibility for the quality of education they provide, and should be committed to continuous improvement. As a whole, Local Authorities, and the schools therein, are expected to engage in a process of Validated Self-Evaluation (VSE). This is a voluntary process that aims to support and challenge the work of education authorities to improve the quality of provision and outcomes for learners. Schools collaborate with Education Scotland to review overall school quality. Compliance with this process appears to be variable, with only a handful of Local Authority VSE reports published on the Education Scotland website over the past few years (Education Scotland, 2016d). While this process does not explicitly target teachers for evaluation, they are considered important to the evaluation.

In addition to the VSE process, Education Scotland provides national guidance on Professional Review and Development (PRD) (GTCS, 2016c). This guidance includes information about the Professional Update process whereby teachers engage in regular professional learning, self-evaluate against the Standards, and provide evidence of these activities every five years as part of the re-accreditation process for individual teachers (Education Scotland, 2016c). Hence the evaluation framework is built around a statutory requirement for registration. The PRD process is underpinned by the following set of principles:

1. Teachers at all levels are entitled and responsible for participating in PRD
2. PRD should positively impact professional learning and the outcomes for students
3. PRD should be an ongoing process
4. PRD should be based on evidence-based self-evaluation
5. PRD should involve professional dialogue
6. PRD should be recorded and weighed in terms of evidence of impact

Nonetheless, despite this guiding evaluation framework, pragmatics regarding frequency of the PRD process are not specified, suggesting that this process sits within the domain of individual schools and school staff.

METHODS AND METHODOLOGIES

Scotland does not have a systematic review process for teachers, and instead encourages teachers to engage in ongoing self-evaluation. While the PRD framework provides some guidance regarding professional learning and self-evaluation, specific tools for evaluation are not outlined by Education Scotland or GTCS (GTCS, 2016c). GTCS does provide examples of evidence for teachers to use in their process of self-evaluation for their professional learning, which include analysis of pupil work, teacher talk, reflections on professional dialogues with peers, parents, and colleagues, and analysis of survey ratings (GTCS, 2016i). However, these suggestions are only recommended for the self-evaluation process.

Student Outcomes

Teachers in Scotland are encouraged through the GTCS to understand and build evidence of their impact on student learning through a continuous cycle of professional learning. Student achievement and learning is seen as one of a range of sources to support the self-evaluation process. No specific information was found regarding the use of explicit student achievement measures on a more formal evaluation process for teachers.

Direct Observation of Teaching

Observation is mentioned by GTCS only as one method for bolstering self-evaluation and professional learning (GTCS, 2016c).

41

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Survey Ratings

Not referenced in reviewed documents.

Teacher Self-Assessment

Teacher self-evaluation is encouraged by GTCS as aforementioned. Education Scotland sets out general guidelines and suggestions for teacher to assist in the self-evaluation process, but these are not enforced (GTCS, 2016f).

Internal/External Evaluators

As previously mentioned, professional development processes are the responsibility of the individual teacher within the context of their school. Education Scotland and GTCS provide a framework for professional learning to guide these activities for teachers.

Scotland has a central inspectorate, operating under Education Scotland, charged with evaluating the effectiveness of each local authority in providing quality education, and supporting schools in improving quality. In undertaking specific school inspections, Education Scotland canvasses the views of students and parents with regards to specific quality indicators of staff. For example, a question from the student questionnaire states: "Staff talk to me regularly about how to improve my learning". These views and perceptions are, however, at the aggregate level, rather than at an individual teacher level.

Aggregation and Weighting of Results

Not referenced in reviewed documents.

USES OF EVALUATION

Overall, the use of evaluation for teachers in the Scottish Education system appears to be largely perfunctory, serving to obtain or maintain registration. The GTCS does however encourage self-evaluation for individuals supported by the provision of a range of materials to guide the teacher in understanding their skills and areas for development (GTCS, 2012a). GTCS encourages an evaluative culture for the Scottish teaching workforce that is demonstrated in the requirement to participate in professional development (known as professional update) to maintain registration (Education Scotland, 2016c). Furthermore, there is also an emphasis on professional learning as part of self-evaluation. However, this appears to only be encouraged, rather than enforced beyond stating, "All teachers have an entitlement to and responsibility for their own professional learning" (Education Scotland, 2016b). Similarly, Head Teachers are asked to provide ongoing feedback; however, this is not mandated (Education Scotland, 2016b).

CONSIDERATIONS FOR MINORITY GROUPS

Scotland has an additional registration category for teachers engaged with students with Additional Support needs as previously mentioned.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

EXECUTIVE SUMMARY

- There are no regular or reoccurring teacher evaluations in Germany, and formal evaluations are mostly carried out in response to serious complaints or for decisions regarding employment such as promotion. This is attributed to the rigorous training German teachers receive.
- Evaluations are conducted by an external inspector employed by the Länder (States), who relies solely on their professional judgement, as there is no framework or set criterion for evaluation.
- Teaching standards in Germany are set at a national level, but are not available in English.

INTRODUCTION

While there is some variance between the Länder (States), there are typically six different types of schools in Germany, with education compulsory for children aged 6 to 18 (OECD, 2009):

- **Primary School** (The Grundschule) – Operates for five years, for students aged 6-10, although there are some variances between Länder.
- **Secondary School –** Operates for nine years, for students ages 10-18. Secondary school is divided into: Gymnasium; Realschule (middle school); Hauptschule (high school); Gesamtschule (comprehensive school); and Förderschulen/Sonderschulen (special schools), with each of the schools varying in their educational focus and curriculum.

Governance

The education system in Germany is the shared responsibility of the German Federal Ministry of Education and Research (BMBF) and the relevant body in each of the 16 Länder. The Länder are largely responsible for policies relating to teacher evaluation, as such regulations and practices for teacher appraisals vary throughout the country.

STANDARDS AND THEORY

Teaching standards in Germany are set at a national level. Full articulation of these standards was not available in English, and will therefore not be included in this review.

TEACHER PREPARATION

In Germany, entry to university programs is determined by applicants' performance in their final secondary school examination, the Abitur. Teacher applicants are selected from high achievers with results in the top third.

German teachers have the longest pre-service teacher training among all 64 of the PISA 2012 countries (OECD, 2013). Teacher training courses are offered at universities and teacher training colleges, with different degree qualifications needed for each of the schools. Most of their programs have two phases. A typically three-and-a-half-year academic phase concludes with a State Examination after which they are typically awarded a degree (European Agency for Special needs and Inclusive education, 2015). This is followed by a second phase of preparatory service in schools, where students undertake pedagogical studies and ministry provided training seminars, teaching at a reduced salary. A second State Examination, including classroom observations, performance assessments, an oral examination and essay, is held at the end of the second phase (European Agency for Special needs and Inclusive education, 2015). The relevant body in each Land sets examinations, which also vary depending on the school type, subject and year level. The Ministries of Education and Cultural Affairs, or Kultusministerkonferenz (KMK), also set national regulations for the practicum with some provision for universities and colleges to vary content and structure.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

TEACHER REGISTRATION

At the successful completion of both phases of their teacher education program, graduates are qualified to enter the profession as probationary teachers. This probationary period lasts for two years, after which their appointment is life-long.

COURSE ACCREDITATION

The KMK in each Land regulates teacher training and preparation, and sets standards for courses and examination (EU, 2015). In recent years, the KMK has provided additional funding to support PISA studies and reformation of the German academic system consistent with the Bologna Declaration, allowing for consistent training of teachers between nations. The KMK and the Länder have a brokered agreement that stipulates the structure, content, and duration of teacher education programs and general examination requirements.

EVALUATION FRAMEWORK

There are no regular or reoccurring teacher evaluations in Germany, and most teacher evaluations are carried out in response to complaints or decisions regarding employment status (e.g., promotion to principal) (OECD, 2014a).These teacher evaluations are typically conducted by an external inspector employed by the Länder. However, there is no framework or set criterion for evaluation, which relies solely on the evaluator's professional judgement (OECD, 2014a). The lack of emphasis on teacher evaluation in Germany is presumably an outcome of the intense initial teacher training, with the OECD (2004) claiming there is the impression that teachers have the "authority to practise largely autonomously for their entire careers".

Germany has a highly structured legal framework for external school evaluation, with evaluations conducted regularly by the relevant body in each Land (Figazzolo, 2013; OECD, 2014a). However, these school evaluations do not involve formative or summative teacher appraisals.

METHODS AND METHODOLOGIES

Specific evaluation methods or methodologies were not identified in English documents.

Student Outcomes

Not referenced in reviewed documents.

Direct Observation of Teaching

Not referenced in reviewed documents.

Survey ratings

Not referenced in reviewed documents.

Teacher Self-Assessment

Not referenced in reviewed documents.

Internal/External Evaluators

Not referenced in reviewed documents.

Aggregation and Weighting of Results

Not referenced in reviewed documents.

USES OF EVALUATION

Teacher evaluations in Germany predominantly occur if teachers are being assessed due to inability to complete their job satisfactorily (Figazzolo, 2013). Therefore teacher evaluations appear to have little impact on career progression, which is largely determined by years in service (OECD, 2014). While dismissal may result from a poor evaluation, German teachers have a high level of job security and termination is rare.

CONSIDERATIONS FOR MINORITY GROUPS

Not referenced in reviewed documents.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

EXECUTIVE SUMMARY

- There is no systematic teacher evaluation system in Austria. Teacher evaluation is almost exclusively limited to circumstances where there has been a serious complaint against a teacher. In such cases, regional and national inspectors are responsible for evaluations in response to serious complaints about a teacher, and also for evaluating the heads of school. No information regarding training or selection of inspectors was available in English.
- Standards for expected learning outcomes are set by each Land (Province) and allow for teachers to use their own methods and techniques.
- Teacher preparation is rigorous, with appointments requiring checking of personal and professional qualifications as well as a counseling interview. Prospective teachers must undergo either a university-based program of 4.5 years, or a three-year program followed by additional in-service training courses.
- There is no formative evaluation process and appraisals are not linked to professional development.
- Austrian teachers report feeling that their improvements in their teaching will not be recognised, and that school principals fail to identify effective or under-performing teachers.

INTRODUCTION

The Austrian educational system consists of several school types and stages, with compulsory education lasting nine years (Bundesministerium für Bildung und Frauen, 2014):

- **Primary Education** – Operates for four years, and is compulsory. Students attend from ages 6-10 years
- **Lower Secondary Education** – Operates for five years, and is compulsory. Students attend from ages 10-14 years.
- **Upper Secondary** – It is compulsory for students attend at least one year of upper secondary education, typically at 15 years old. Following this, compulsory education is complete, however students may continue their schooling via the following:
  - o Academic Upper Secondary school – a continuation of Upper Secondary school leading to tertiary studies
  - o Pre-Vocational and Vocational Schooling – education focused on gaining trade-based skills

Private schools represent approximately 10% of Austrian schools (OECD, 2012).

Governance

Education in Austria is the joint responsibility of the National Government and the nine Länder (Provincial Governments). The Federal Ministry for Education and Women's Affairs (BMBF, 2015) is responsible for developing educational legislation and principles, which are enforced across the Länder. However, each Land is given a high degree of autonomy in educational governance, and for creating teacher evaluation policies (OECD, 2012). Furthermore, key stakeholders in the educational system (i.e., students, parents, teachers, and teacher unions) are given representation and power in decision making (Delannoy, McKenzie, Wolter, & van der Ree, 2003). Schools generally carry out teacher appraisals, however, there are no explicit frameworks for evaluations. Regional and national inspectors are responsible for evaluations in response to serious complaints about a teacher, and also for evaluating the heads of school (OECD, 2012).

STANDARDS AND THEORY

Standards in Austria are set by each Land, and allow for teachers and school leaders to use their own methods and techniques to accomplish learning outcomes. These standards vary depending on grade level, subject area and

school type. They outline the students' expected learning outcomes, as opposed to the teaching practices (OECD, 2012). English translation of these standards is not yet publicly available.

TEACHER PREPARATION

Austria has very recently introduced new teacher training programs, aiming to reorganise and standardise teacher training across the country, effective 2015-2017 (European Union, 2015). By regulation, Study Commissions set the entrance requirements for teaching applicants that generally include checking of personal and professional qualifications and a counselling interview.

Austria has a two-tiered system for teacher education (European Agency, 2016b). Compulsory School teachers (primary, lower secondary, special school and pre-vocational school) graduate from a three-year program in a higher education college in the area of general education and are awarded a Bachelor of Education (European Agency,2016a). They are then required to undertake additional in-service training courses offered by the teacher training college or at the school level. Lower and senior secondary school teachers undertake a university-based program of 4.5 years.

TEACHER REGISTRATION

Teachers have a probationary period (around 2 years), and following a successful evaluation will generally have a permanent position (OECD, 2012). There is no registration.

COURSE ACCREDITATION

In Austria, national laws and decrees define through legislation the organisational and operational requirements for University provided teacher education programs. Legislature for teacher training colleges specifically details program duration, syllabus, examination regulations and certification (OECD, 2013).

The content of initial teacher education programs in Austria is decided autonomously by the Study Commissions in accordance with the *Regulation on University Curricula* (Hochschul-Curricula-Verordnung (European Union, 2015). The Ministry of Education make recommendations and provide advice about course content, whilst teacher educators' autonomy and participation depends upon the culture and tradition of the teacher training college.

EVALUATION FRAMEWORK

Austria has no systematic process of teacher evaluation, as evaluations only rarely occur (OECD, 2012). In circumstances where there has been a serious complaint against a teacher a summative evaluation will be carried out by someone employed by the provincial educational office (Delannoy et al., 2003). These inspections do not follow a set criteria or framework, and "tend to have no consequences" (OECD, 2012, pp.22). Furthermore, there is no formative evaluation process (Delannoy et al., 2003), and appraisals are not linked to professional development.

METHODS AND METHODOLOGIES

Specific evaluation methods or methodologies were identified in English documents.

Student Outcomes

While there is a strong emphasis on standardised testing in Austria (with students participating in 'low stakes' examinations each year), this data is largely used for school-based evaluations, and not for teacher evaluations. Teachers are encouraged to reflect upon their students results; however, this is in no way enforced.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Direct Observation of Teaching.

Austrian teachers are given a high level of autonomy, and direct observations of teaching are very uncommon (OECD, 2012), acting as the "exception rather than the rule" (Delannoy, et al., 2003, p.23). Furthermore, head teachers and school administrators are given little training in how to conduct observations, and teacher evaluations more generally (OECD, 2012).

Survey ratings

Not referenced in reviewed documents.

Teacher Self-Assessment

Not referenced in reviewed documents.

Internal/External Evaluators

Not referenced in reviewed documents.

Aggregation and Weighting of Results

Not referenced in reviewed documents.

USES OF EVALUATION

As Austria has no systematic process of teacher evaluation, and as evaluations only occur in cases of poor performance, evaluations appear to have little impact on Austrian teachers' career outcomes. Evaluations are not linked to salary increases, which instead follow tenure-based advancements (OECD, 2012). There is concern over the role evaluation may play in teacher career progression in Austria. Indeed, Austrian teachers appear to feel under-recognised, with the OECD (2009) reporting that 85% of Austrian teachers believed that improvements in their teaching will not be recognised, and 90% reporting that the highly effective teachers in their school received no acknowledgement or reward (OECD, 2009). Furthermore, over 50% of Austrian teachers claimed that their principal does not use appropriate methods to identify effective or under-performing teachers (OECD, 2009). While teachers are required to participate in professional development programs, these are not linked to evaluation (Delannoy et al., 2003).

CONSIDERATIONS FOR MINORITY GROUPS

Not referenced in reviewed documents.

EXECUTIVE SUMMARY

- Singapore has a highly centralised education system under the Ministry of Education Singapore (MOE), with a centralised teacher evaluation program known as the Enhanced Performance Management System (EPMS). Teachers receive considerable support to encourage ongoing growth and development as a professional.
- The Ministry of Education prescribes a set of Graduand Teacher Competencies across 7 core competencies: nurturing the whole child, providing quality learning of child, providing quality learning of child in co-curricular activities, cultivating knowledge, winning hearts and minds, working with others, and knowing self and others. Teacher competencies are prescribed for all teachers. These competencies guide teacher evaluation procedures.
- As the national government is the sole provider of teacher education in the country, no course accreditation or teacher registration is required.
- Evaluation tools include student outcomes, student wellbeing, observation, survey ratings, and self-assessment.
- The EPMS is both summative and formative in nature. Teacher evaluation does inform hiring decisions, annual target setting, career pathways and annual bonuses. More saliently, as a year-long process, teacher evaluation emphasises professional development and growth.

INTRODUCTION

Singapore consists of public and some private sector schools that are categorised as:

- **Primary School** – Operates for six years, from Primary 1 to Primary 6

- **Secondary School** – Operates for four to five years, from Secondary 1 to Secondary 4

- **Post-Secondary School** – Includes educational institutions, such as Junior College, Polytechnic, and Institute of Technical Education. The varying streams within the post-secondary school category last for differing durations and lead to different career pathways (Ministry of Education, 2015)

Within the Singaporean system, education is compulsory for 'school-age children', or those between the age of six and fifteen (MOE, 2015; Ministry of Education Singapore, 2016b). At the primary school level there is little variation: children either attend primary school or enroll in special education. At the secondary school level and above, the MOE provides a range of institutions. Students can attend a private or public school, with the option of choosing academic or technical pathways (MOE, 2016a). In this respect, the Singaporean Education system caters specifically to learning abilities, and consistently tests students at the primary and secondary level to find the appropriate institution for each student. There are several national examinations which are developed and conducted by the Singapore Examinations Assessment Board (SEAB), a statutory board under the MOE. The national examinations undertaken by enrolled students include:

- The Primary School Leaving Examination (PSLE) – a national examination administered by the MOE for all students towards the end of the 6th year of primary school, prior to entering secondary school

- The General Certificate of Education Normal Level examinations (GCE N) - a series of national exams set for secondary school students after the first four years of secondary education, with several levels outlined below.
    - GCE N(T) – an exam taken by 'Normal Technical' students, leading to admission into an Institute of Technical Education
    - GCE N(A) – and exam undertaken by 'Normal Academic' students, leading to higher secondary education within the 'Normal Academic' stream

- The General Certificate of Education Ordinary Level examination (GCE O) – a national exam conducted with after four years of express, or five years of normal academic secondary education(Ministry of Education, 2015).

At the post-secondary level, there is a large variety of organisations and institutions available to graduating students with diverse interests. Universities, polytechnics and arts institutions are all viable options for post-secondary students (Ministry of Education, 2015).

Governance

Education in Singapore is closely overseen by the MOE. While schools may be granted 'independent' or 'autonomous' status, this is uncommon, with roughly 8% of schools in Singapore operating under these titles (Tan & Dimmock, 2014). Typically, the more high-performing the school, the more likely it is to be granted greater autonomy (National Institute of Education Singapore, 2016; Tan & Dimmock, 2014). This largely centralised education system can be conceptualised as consisting of three components: the MOE, the National Institute of Education (NIE), and individual schools (Ministry of Education, 2015; Tan & Dimmock, 2014).

The MOE exerts a large degree of control via the formulation and implementation of educational policies relating to curriculum, pedagogy, and assessment (Ministry of Education, 2016). The MOE also oversees the management and development of government-funded schools, the Institute of Technical Education, polytechnics, and universities (Ministry of Education, 2016). Using data gathered annually from schools, as well as internal government data, the MOE forms strategies for meeting the needs of schools (The Consortium for Policy Research in Education, 2007). These plans include short-term strategies, such as replacing retiring teachers, as well as long-term strategies, including planning new schools (CPRE, 2007). All employed teachers, including prospective teachers enrolled in ITE in Singapore at the NIE, are employed directly by the MOE (CPRE, 2007). While there has been some shift towards school autonomy, the MOE retains a large degree of control over schools through directive policies and the legal ability to approve, vary, and revoke governing board constitutions in public and independent schools in Singapore (Singapore Government, 2013; Tan & Dimmock, 2014).

The NIE is the sole ITE provider in Singapore (Ministry of Education Singapore, 2016a; Tan & Dimmock, 2014). While the NIE is an autonomous institute of the Nanyang Technological University, it shares a symbiotic relationship with the MOE. All prospective teachers hired by the MOE who do not meet the pre-requisite pedagogical training are expected to enrol in teacher training through the NIE (CPRE, 2007). The NIE collaborates with the MOE in defining academic qualification standards for teachers, and has some input in determining the interview and selection criteria for prospective teachers (CPRE, 2007). Further, the NIE is responsible for conducting educational research, largely funded by the MOE (Tan & Dimmock, 2014). This includes assisting in the design of curriculum and other critical education system components, which are then promptly reflected in course content (CPRE, 2007). As the NIE has strong ties to the MOE, ITE and research goals closely align with MOE policy goals (NIE, Academy of Singapore Teachers, 2016a; 2016; Tan & Dimmock, 2014).

Individual schools are divided into clusters of 12-14 schools, and overseen by cluster superintendents who represent the MOE (Tan & Dimmock, 2014). MOE staff are required to participate in regular evaluations referred to as the Enhanced Performance Management System (EPMS). Their immediate supervisor evaluates staff members using the framework outlined by the MOE.

STANDARDS AND THEORY

The Singaporean education system, and in particular expected competencies for students and teachers, emerged following a fundamental review of the education system in 2000. This included engaging with educational consultants, school leaders, and teachers to formulate new frameworks for guiding education professionals in Singapore. One framework developed by the MOE is a set of student-focussed 'competencies' that outline the attributes an individual should have gained once they have completed formal education (Ministry of Education Singapore, 2016c). The Desired Outcomes for Education (DOEs) are the culmination of 'competencies' from each

level, and establish a common purpose amongst all educators in Singapore (Ministry of Education Singapore, 2016c). The DOEs note that a Singaporean graduate should be:

- A **confident person** who has a strong sense of right and wrong, is adaptable and resilient, knows himself, is discerning in judgement, thinks independently and critically, and communicates effectively;
- A **self-directed learner** who takes responsibility for his own learning, who questions, reflects, and perseveres in the pursuit of learning;
- An **active contributor** who is able to work effectively in teams, exercises initiative, takes calculated risks, is innovative, and strives for excellence; and,
- A **concerned citizen** who is rooted to Singapore, has a strong civic consciousness, is informed, and takes an active role in bettering the lives of others around him (MOES, 2016c).

The MOEs Framework for 21st Century Competencies build on the DOEs further, outlining specific values expected in Singaporean students, including:

- Self-Awareness
- Self-Management
- Social Awareness
- Relationship Management
- Responsible Decision Making

Teachers are considered integral to promoting these attributes in Singaporean students. As such, the DOEs are strongly linked to expectations relating to teacher performance (MOE, 2016b). The Academy of Singapore Teachers (AST), a branch of the MOE, outlines the expected competencies of teachers in the Teacher Growth Model (Academy of Singapore Teachers, 2016b). These competencies directly mirror the DOEs for students in Singapore, utilising the same four competencies.

The MOE and NIE articulate expected standards for graduated teachers via the Graduand Teacher Competencies, outlined in Table 6 (National Institute of Education, n.d.-a). These competencies apply to all graduating teachers as well as currently working teachers, and guide teacher evaluation procedures.

Table 6 - The Graduand Teacher Competencies in Singapore

Performance Dimensions	Core Competencies

Professional Practice	1. Nurturing the whole child
	2. Providing quality learning of child
	3. Providing quality learning of child in CCA
	4. Cultivating Knowledge:
	   i. With subject mastery
	   ii. With reflective thinking
	   iii. With analytic thinking
	   iv. With initiative
	   v. With creative teaching
	   vi. With a future focus

Leadership & Management	5. Winning Hearts and Minds
	   i. Understanding the Environment
	   ii. Developing Others
	6. Working with Others
	   i. Partnering Parents
	   ii. Working in Teams

Personal Effectiveness	7. Knowing Self and Others

Teacher Effectiveness Systems, Frameworks and Measures: A Review

	i.	Tuning into self
	ii.	Exercising personal integrity and legal responsibilities
	iii.	Understanding and respecting others
	iv.	Resilience and adaptability

TEACHER PREPARATION

The NIE is the sole teacher education institute in Singapore (Center on International Education Benchmarking, 2016). In partnership with the MOE, NIE's initial teacher education programs prepare teachers for all Singaporean government schools (Primary, Secondary and Junior Colleges). ITE programs include the following:

- Diploma in Education (2 years, Primary/Secondary)
- Bachelor of Science (Education)/Bachelor of Arts (Education) (4 years, Primary and Secondary)
- Postgraduate Diploma in Education (PGDE) (1 year, Primary/Secondary/Junior College) (CIEB, 2016)

The Bachelor of Education (3 years part-time + half year full-time) program is designed for primary teachers seeking a degree-level qualification. Singapore recruits prospective candidates from the top third of secondary school graduates (CIEB, 2016). The number of applicants is limited by demand, and set by the MOE. After being shortlisted against NIE's university admission criteria, applicants undertake a faculty interview, followed by a short teaching allocation as an untrained teacher as a further suitability assessment (CIEB, 2016). The MOE provides all student teachers tuition waivers and a monthly salary or stipend.

Building from the significant education revisions in 2000, the NIE launched the Programme Review and Enhancement (PRE) initiative in 2008. This initiative resulted in significant reform in ITE, including the development of the Model of Teacher Education for the 21st century (TE21). The TE21 provides six recommendations for teacher education in Singapore, including (National Institute of Education, 2010):

1. **Re-envisioning the Values, Skills, and Knowledge (V^{3}SK) Model**
   The V^{3}SK Framework represents the underpinning philosophy of teacher education at NIE (National Institute of Education, 2010, n.d.-b). Within the V^{3}SK are the expected values, skills, and knowledge teachers are expected to maintain in Singaporean schools. This includes three value paradigms: (i) Learner-centred; (ii) Teacher Identity; and (iii) Service to the Profession and Community, as well as skills and knowledge relating to the teaching profession. These domains are reflected in the ITE programs provided by the NIE.
2. **Implementation of the Graduand Teacher Competencies**
   See previous section for more information.
3. **Strengthening the Theory-Practice Nexus**
   This initiative involves various initiatives for bridging the gap between educational theory and practice, such as mentoring programs, enhanced practicum and opportunities for reflection.
4. **Extended Pedagogical Repertoire**
   This includes refinements to the NIE program and frameworks for ongoing refinement.
5. **Assessment Framework for 21st Century Teaching & Learning**
   The role of assessments and examination of teacher portfolios is highlighted as a component of ITE to continue to develop.
6. **Enhancing Pathways for PD**
   This includes opening further pathways to higher degrees, including:
   - 2-year full time PGDE-Master's programme
   - 4+1 accelerated Bachelor-Master's programme for the top 10% of the cohort

Teacher Effectiveness Systems, Frameworks and Measures: A Review

THE TEACHING TRACK SYSTEM

Upon completion of ITE programs, teachers in Singapore enter the MOE's Teaching Track system. This system has three career tracks, including Teaching, Leadership, and Senior Specialist. Each member of the teaching profession starts as a classroom teacher, and teachers are encouraged to pursue higher positions. Promotions along each track are determined via specific selection criteria, EPMS processes, and recruitment decisions.

PROFESSIONAL DEVELOPMENT

All teachers in Singapore are entitled to 100 hours of Professional Development (PD) annually via the *Professional Development Packages and Leave Scheme*. Under this scheme teachers may participate in PD during school hours with resources provided for relief teachers. The MOE also provides funding for scholarships and study leave, including local and international institutions (Tan & Dimmock, 2014). The MOE also offers opportunities for teachers to continue their education, while the *Teachers' Work Attachment* program allows teachers to partner with an external organisation to expand their ideology and perspectives on education (MOES, 2016a).

TEACHER REGISTRATION

Singapore does not have national licencing or registration requirements for graduate teachers. Due to the close relationship between the MOE and NIE in setting standards for teachers, it is expected that the NIE ITE programs will directly reflect these standards. As such, teachers are not required to complete any additional examinations or demonstrate competence beyond their NIE training.

Permission to teach is only necessary when a private school operator seeks to employ a teacher. These teachers must provide evidence of qualifications in excess of the minimum requirement and subject specific knowledge. Application and registration fees are paid to the relevant MOE department for a letter of permission to be provided.

There are no initial or ongoing regulatory requirements for Singapore teachers. As employees of the MOE, teachers are required to meet set expectations within the EPMS, and also expected to engage in regular professional development, which is funded by the MOE.

COURSE ACCREDITATION

Due to the close collaboration between the NIE and the MOE there is no external accreditation body or program. NIE undergo stringent self-evaluation to ensure coherence, accountability, and 21st century relevance.

EVALUATION FRAMEWORK

Singapore has a nationally orchestrated and implemented evaluation system for teachers, referred to as the Enhanced Performance Management System (EPMS). EPMS is a component of Singapore's Education Service Professional Development and Career Plan (Edu-Pac), which features three components: a career path, recognition through monetary rewards, and an evaluation system. This is a performance-linked method of evaluating teacher effectiveness that applies to teachers in all subjects and levels. This system focuses on enhancing underlying characteristics, or competencies, in order to lead to higher levels of performance (Steiner, 2010).

The EPMS has two primary aspects (Tan & Dimmock, 2014). Firstly it defines the skills, knowledge, and competencies required for each teaching track at each level. Secondly, the EPMS strives to be a formative process, supporting the development of all teachers. To achieve these purposes, the EPMS utilises the following processes:

- **Performance Planning**
  At the start of each year teachers engage in self-assessment and develop goals for teaching, instructional innovations and improvements at the school. Teachers also set goals for professional development and

Teacher Effectiveness Systems, Frameworks and Measures: A Review

personal development, and meet with their reporting officer who is usually the Head of Department for a discussion about target setting and performance benchmarks.

- **Performance Coaching**
Throughout the year teachers engage in performance coaching where their reporting officer meets with them to discuss progress and share needs, and to provide guidance, feedback and support. Teachers also engage in a formal mid-year review with their reporting officer.
- **Performance Evaluation**
At the end of the year, teachers engage in a formal review. This consists of an appraisal interview conducted by their reporting officer, as well as a review of actual performance versus planned performance. Teachers receive a summative performance grade, which affects the annual performance bonus received for the year's work. During this phase of the performance evaluation decisions regarding promotions are also considered. This decision is based on teachers' current estimated potential (CEP), which is determined in consultation with senior staff members who have worked with the teacher, and is based on observations, discussions with the teacher, evidence of portfolio, and knowledge of the teacher's contribution to the school and community.

Teachers are evaluated based on outcomes and processes. Outcomes relating to teacher performance are captured through Key Result Areas (KRA), which describe the broad areas of work expected of a teacher. The KRAs for the Teaching Track are:

- The holistic development of students through:
    - Quality learning of students
    - Pastoral care and well-being of students
    - Co-curricular activities
- Contribution to the school
- Collaboration with parents
- Professional development

The knowledge and skills expected of a teacher complements the KRAs and they are:

- Knowledge
    - Teaching Area – the content and curriculum knowledge that teachers must know to teach in the classroom
    - Psychology – the knowledge of child development that teachers must know to maximise pupil potential.
    - Developments in the field of education – the knowledge in other areas of education that mould a complete educator
    - Education Policies – the rationale and the philosophy that sets the direction and focus for teachers to carry out their tasks.
- Skills
    - Teaching Pedagogy – the pedagogic techniques and approaches that teachers must practise to teach in the classroom.

On these domains teachers are evaluated based on a 4-point scale, ranging from 'Not Observed', to 'Developing', to 'Competent' and 'Exceeding'.

The processes of teaching are evaluated using the Teacher Competency Model, which consists of 13 domains, summarised in Table 7. These competencies are considered to represent underlying characteristics connected to high performance as well as the behaviours and attributes that are linked to long-term achievement and success.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Table 7 - Singapore's EPMS Teacher Competency Model

Domain	Competencies Examined
Core Competency	1. Nurturing the Whole Child
Cultivating Knowledge	2. Subject Mastery 3. Analytical Thinking 4. Initiative 5. Teaching Creatively
Winning Hearts and Minds	6. Understanding the Environment 7. Developing Others
Working With Others	8. Partnering Parents 9. Working in Teams
Knowing Self and Others	10. Tuning into Self 11. Personal Integrity 12. Understanding Others 13. Respecting Others

While teachers are expected to engage in self-evaluation based on all 13 competencies, only 9 of these competencies (Nurturing the Whole Child and the competencies related to Cultivating Knowledge, Winning Hearts and Minds and Working with Others) are used in formal evaluations conducted by the reporting officer. Each competency contains different levels of ability, with more senior teachers expected to perform at higher levels than less senior teachers. Each level contains behavioural indicators to guide reporting officers in their grading of a teacher's performance.

METHODS AND METHODOLOGIES

Student Outcomes

As well as student academic outcomes, the EPMS values more holistic student outcomes, including student wellbeing and co-curricular activities (Steiner, 2010). Results from the various national examinations as well as other measures of student achievement form part of a teacher's overall evaluation results. The specific weighting of student achievement results on evaluation results is not set by the MOE.

Direct Observation of Teaching

Direct observation is referenced as one of the potential tools used by evaluators in determining the performance of teachers. Specific observation methods or rubrics for classroom observation were not identified in the reviewed documents.

Survey Ratings

Not referenced in reviewed documents.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Teacher Self-Assessment

Teachers primarily engage in self-assessment during the performance planning stage of EPMS, which occurs at the start of the year. The EPMS system also encourages ongoing self-evaluation and response to feedback from coaching.

Internal/External Evaluators

The teacher's immediate supervisor, typically a head teacher, conducts teacher evaluations through the EPMS. Additionally, summative evaluations conducted at mid-year and at the end of the year must be co-signed by a second evaluator (Tan & Dimmock, 2014).

Aggregation and Weighting of Results

The MOE does not set requirements for the weighting of individual factors, such as student achievement, on overall teacher evaluation results. Individual schools do set internal expectations regarding the weighting of evaluation measures, leading to differences in the weighting of specific evaluation components between schools (Steiner, 2010).

USES OF EVALUATION

The EPMS is both summative and formative in nature. It is a year-long process involving planning, coaching, and evaluation, and as such involves a high degree of collaboration conducive to the continued development of teachers (Tan & Dimmock, 2014).

Performance is recognised using monetary and non-monetary rewards. Since the transformation of the performance evaluation system in 2001, the connection between a teacher's performance, their pay, and their opportunities for promotion has significantly increased (Steiner, 2010). Performance grades are used to determine annual increases in salary, as well as bonuses. These bonuses range from one to three months of the teacher's annual salary (Tan & Dimmock, 2014).

Additionally, by using a three-track system in which classroom teachers can achieve higher levels in their profession, the Ministry of Education has implemented a close relationship between professional status and performance. Within the Teacher Track System, individuals must meet specific criteria in order to advance to the next stage of their chosen track. This enables a pattern of continuous professional development and reward (Ministry of Education Singapore, 2016a).

CONSIDERATIONS FOR MINORITY GROUPS

Not referenced in reviewed documents.

EXECUTIVE SUMMARY

- No central teacher evaluation framework is present in Hong Kong; however, guidelines for developing evaluation systems for teachers within schools are present. Individual schools are responsible for implementing teacher evaluation programs, known as Teacher Performance Management.
- No teaching standards are articulated; however, a list of Expected Evidence of Performance Indicators is provided to guide teacher evaluation.
- Some guidance is provided regarding specific methodologies and methods for evaluating teachers; however, evaluation is still ultimately the responsibility of the school. Tools recommended by the Education Bureau include student outcomes, observation, teacher portfolios, self-assessment, survey ratings, and records of classroom events such as non-teaching duties and extra-curricular activities.
- Teacher evaluation is mostly formative, intended to increase accountability and staff motivation and drive professional development. Results may also be used in a summative sense to determine promotion or pay rises. Use of evaluation results for other outcomes is left up to the school's discretion.

INTRODUCTION

Pre-tertiary schooling in Hong Kong is divided into:

- **Kindergarten** – for children aged 3-5
- **Primary School** – starting at six years old and lasting for six years, from P1 to P6
- **Junior Secondary School** – three years starting from the end of P6, from JS1 to JS3
- **Senior Secondary School** - three years starting from the end of JS3, from SS1 to SS3

Students in Hong Kong are entitled to 12 years of publicly provided education, starting from primary school through to the end of secondary school (Education Bureau, 2016a). It is mandatory for appropriately aged children to attend primary and secondary school; however, kindergarten is not mandatory (Education Bureau, 2016a). Children may leave school at the end of JS3, with full subvention provided for full-time courses run by the Vocational Training Council (Education Bureau, 2016a). Classes are taught in Cantonese and English, with individual schools choosing how much of each language is used as a medium for instruction (Education Bureau, 2016a). A national exam, the Hong Kong Diploma of Secondary Education Examination, is completed at the end of secondary school (Education Bureau, 2016a). As of September 2015, approximately 273,006 children were enrolled in primary schools, while 282,525 were enrolled in secondary schools in Hong Kong (Education Bureau, 2016a).

Governance

Hong Kong is a Special Administrative Region of China, maintaining an independent system of education from the Chinese education system (Education Bureau, 2016c). The Education Bureau of Hong Kong is responsible for teacher qualification/registration (Education Bureau, 2016c). Individual schools are responsible for performing teacher evaluation, referred to as Teacher Performance Management (TPM), based on guidelines provided by the Education Bureau (Education Bureau, 2003).

STANDARDS AND THEORY

Although they are not explicitly set out as standards, the Education Bureau does provide a list of Expected Evidence of Performance indicators as part of their TPM guidelines (Education Bureau, 2003). These are intended as a guide for schools to use and adjust in order to meet the specific needs of their school context. The Expected Evidence of Performance indicators are not considered an exhaustive list of the qualities expected in teachers (Education Bureau, 2003). Domains listed include:

57

Teacher Effectiveness Systems, Frameworks and Measures: A Review

- Curriculum Planning and Organisation
- Teaching Planning Strategies and Skills
- Knowledge and Attitude
- Assessment Planning and Implementation
- Support for Student Development
- Links with parents and external organisations
- Interpersonal Relationships

TEACHER PREPARATION

To become qualified, teachers in Hong Kong must undertake an official teacher-training program or apply through the Non-Graduate Teacher Qualifications Assessment (typically intended for mature aged candidates) undertaken after a period of school based in-service training (see Registration section below: Permitted Teacher).

Official Teacher Training is offered at a Comprehensive University, the Open University of Hong Kong, or the Institute of Education. Admission requirements are set by the institutions, but are increasingly influenced by guidelines and standards produced by the Education Bureau. Applicants typically undergo an assessment of their content knowledge and are interviewed to ascertain their fluency in both English and Cantonese. Hong Kong's primary and secondary teachers who complete official teacher training hold a Teacher's Certificate or Post-graduate Diploma/Certificate in Education, with the majority holding a bachelor degree or higher.

TEACHER REGISTRATION

Teachers must be registered with the Education Bureau to teach in Hong Kong (Education Bureau, 2016c). There are two forms of registration in Hong Kong – Permitted Teacher (PT) and Registered Teacher (RT) (Education Bureau, 2016c). Individuals wishing to teach who hold academic qualifications without official teacher training must apply to be a PT. The minimum academic qualification required of a PT teaching in a school providing primary, secondary, or post-secondary education is an associate degree, a higher diploma or equivalent (Education Bureau, 2016b). Individuals who hold a teacher qualification (e.g. a local Teacher's Certificate or Post-graduate Diploma/Certificate in Education) may apply to be a RT (Education Bureau, 2016b). Private schools offering non-formal curriculum may provide some exemptions.

COURSE ACCREDITATION

The Hong Kong Council for Accreditation of Academic and Vocational Qualifications (HKCAAVQ) is an independent statutory body which provides authoritative advice on academic standards of degree programs in higher education institutions in Hong Kong (HKCAAVQ, 2016). HKCAAVQ is responsible for upholding the Academic and Vocational Qualifications Ordinance (The Government of the Hong Kong Special Administrative Region, 2007), and ensuring that higher education courses in Hong Kong meet expected standards (HKCAAVQ, 2016). However, there are nine "self-accrediting" institutions which are exempt from this process, including: City University of Hong Kong, Hong Kong Baptist University, Lingnan University, The Chinese University of Hong Kong, The Education University of Hong Kong (applicable to teacher education programs only), The Hong Kong Polytechnic University, The Hong Kong University of Science and Technology, The University of Hong Kong, and The Open University of Hong Kong (The Government of the Hong Kong Special Administrative Region, 2016). As such, many of the ITE courses in Hong Kong are asked to submit voluntarily to the standards upheld by HKCAAVQ, rather than submitting to a formal review process (The Government of the Hong Kong Special Administrative Region, 2007).

EVALUATION FRAMEWORK

While there is no mandated teacher evaluation framework in Hong Kong, the Education Bureau provides guidelines for developing evaluation systems of teachers within schools in the TPM (Education Bureau, 2003). It is the responsibility of individual schools to implement teacher evaluation programs (Education Bureau, 2003). The goals of TPM are outlined as:

58

1. *Increasing accountability* of teachers in order to determine whether teachers are satisfactorily reaching performance targets, ensure that teachers are performing the duties expected of them, identify underperforming teachers, and provide guidance for underperforming teachers
2. *Increasing staff motivation* by recognising and rewarding high quality teaching and providing constructive feedback to teachers
3. *Driving professional* development by assisting teachers in identifying areas for improvement, providing guidance, counselling, and training

METHODS AND METHODOLOGIES

The Education Bureau encourages schools to consider the purpose of TPM, whether for professional development, summative appraisal, or both (Education Bureau, 2003). It is expected that the purpose of the evaluation will inform the desired criteria, and in turn the design of the school's TPM program (Education Bureau, 2003). The appraisal cycle is decided by schools, and may be one year or a continuous period of two years (Education Bureau, 2003). The frequency of TPM also depends on whether the teacher being appraised is a new teacher in their probation period, or an experienced teacher (Education Bureau, 2003). Methods recommended by the Education Bureau for teacher evaluation include:

- Student outcomes
- Direct classroom observation
- Teacher portfolios
- Self-Evaluation
- Scrutinising Schemes of Work, Lesson Plans and Marking of Exercises / Examination Papers
- Formal or informal interviews and discussions
- Daily observation
- Stakeholder survey, e.g. student surveys, parent surveys and teacher surveys
- Record of non-teaching duties
- Record of students' award and punishment
- Record of extra-curricular activities

Student Outcomes

While student outcomes are cited as a method for teacher evaluation, the Education Bureau does not provide specific guidance for how this should be included in teacher evaluation systems (Education Bureau, 2003). It should also be noted that Hong Kong has a strong culture of transparency in regard to student outcomes, with school results often published in newspapers and other media outlets. As such, it may reasonably be expected that student outcomes are an important consideration in the evaluation of teachers in Hong Kong.

Direct Observation of Teaching

The observation process described by the Education Bureau is quite transparent, suggesting that observers should discuss the objectives and focus of the observation with teachers before the observation (Education Bureau, 2003). Feedback immediately following the observation is also encouraged, with observers instructed to focus on job performance, rather than personality, and alternate teaching methods (Education Bureau, 2003). In determining the observation rubric the Education Bureau provides little direction beyond stating that subject panels may be used to develop appropriate indicators based on specific subjects and school needs (Education Bureau, 2003). It is recommended that observations take place several times each year (Education Bureau, 2003).

Survey Ratings

Not referenced in reviewed documents.

Portfolio

No specific guidance is provided regarding how portfolios should be examined by appraisers; however, the Education Bureau suggests that portfolios be included in order to "demonstrate a teacher's accomplishments over time and across a variety of experiences" (Education Bureau, 2003, p. 10). Potential portfolio materials for examination include: scholarly writings, action research, logs of professional development activities, structured reflection on practice, and teaching artefacts (such as assignments, worksheets, students' work and project directions; Education Bureau, 2003).

Teacher Self-Assessment

While self-appraisal is considered a valid source for teacher evaluation by the Education Bureau, it is deemed a supplementary source for appraisers (Education Bureau, 2003). Another stated purpose of teacher self-assessment is to encourage input from teachers in the appraisal process and to share the responsibility of self-improvement and personal growth with appraisers (Education Bureau, 2003). Teachers are encouraged to reflect on their work, student outcomes, and professional development needs (Education Bureau, 2003).

Other

Other sources of information for teacher appraisals noted by the Education Bureau include:
- Scrutinising Schemes of Work, Lesson Plans and Marking of Exercises / Examination Papers
- Formal or informal interviews and discussions
- Daily observation
- Stakeholder survey, e.g. student surveys, parent surveys and teacher surveys
- Record of non-teaching duties
- Record of students' award and punishment
- Record of extra-curricular activities

Internal/External Evaluators

Evaluations are usually conducted by the teacher's direct supervisor (Education Bureau, 2003). In a primary school, Certificated Masters/Mistresses (CMs) are appraised by the Deputy Head and Assistant Masters/Mistresses (AMs) who are in turn appraised by the head teacher (Education Bureau, 2003). In a secondary school, Graduate Masters/Mistresses (GMs) are appraised by Senior Graduate Masters/Mistresses (SGMs), who are in turn appraised by Principal Graduate Masters/Mistresses (PGMs) or the principal (Education Bureau, 2003).

Aggregation and Weighting of Results

Not referenced in reviewed documents.

USES OF EVALUATION

The stated intention of TPM by the Education Bureau is to promote accountability, staff motivation and professional development (Education Bureau, 2003). Evaluation results may be used to determine staff promotion, incremental pay rises or disciplinary procedures (Education Bureau, 2003). However, the Education Bureau allows schools to decide how to link evaluation procedures to these other outcomes (Education Bureau, 2003). Schools can consider termination of a teacher if no improvement is seen after a period of time given for intervention, follow up, and issuing of warnings (Education Bureau, 2003).

CONSIDERATIONS FOR MINORITY GROUPS

Not referenced in reviewed documents.

EXECUTIVE SUMMARY

- Korea has several systems in place for teacher evaluation, with nationally set evaluation schemes, executed by local education authorities and individual schools. Evaluators have a degree of freedom in determining evaluation methods.
- No standards for teaching are in place to guide evaluations; however, teaching-specific criteria exist and vary depending on subject and year level taught. Teachers must take a national registration exam covering educational pedagogy, content knowledge, essay, and interview. Results determine recruitment and assignment.
- Tools for evaluation include direct observation, student and parent surveys, and peer ratings. Some aggregation of results occurs.
- Use of evaluation is both summative and formative and feeds into performance-based incentives such as salary increase, planning of training and support, determining promotion outcomes, and informing professional development.

INTRODUCTION

There are three categories of schools in Korea: National, Public and Private. National schools receive their funding from the national government. Local government funds public schools and private organisations fund the private schools. The number of private schools is relatively high (Clark & Park, 2013), with 4,454,259 students in private schools, compared to 889,640 and 6,123,680 students in national and public schools respectively (Ministry of Education Korea, 2015). Education is for 12 years and is divided into:

- **Primary school** which students attend from ages 6-12
- **Middle school** which students attend from ages 12-15
- **High school** which students attend from ages 15-18

Korean students often score highly on international assessments such as PISA, in which they ranked 5th out of 65 participating countries in 2012 (OECD, 2013b). However, this is often attributed to high engagement in private tuition, with 77% of Korean students attending private after school tutoring (OECD, 2011), as opposed to effective teaching and schooling (Clark & Park, 2013).

Governance

There are three levels of Educational Governance in Korea; National, Provincial/Metropolitan, and Local. The Ministry of Education, Science and Technology (MEST) oversees education at a national level, and is responsible for developing most policies, including those relating to teacher evaluation (Korean Educational Development Institute, 2010). The primary function of provincial and local educational offices is to enforce these national guidelines. For example, although local agencies authorise teacher promotions (Kim & Han, 2002), they must follow the procedures set by the Ministry of Education. Relatedly, schools and school principals are responsible for conducting evaluations, but must follow the national framework of teacher evaluation (Korean Educational Development Institute, 2010)

STANDARDS AND THEORY

South Korea does not have explicit standards for teaching, and instead has specific teaching criteria which vary depending on subject and year level (Centre of Study for Policies and Practices in Education, 2013). It should be noted that these documents are not available in English.

TEACHER PREPARATION

Teaching is a high status profession in Korea, and Korean teachers have the highest relative earnings in the OECD (OECD, 2014b). Gaining a place in a teacher training course is highly competitive, with only the top 5% of high school graduates eligible for entrance into elementary teacher training colleges (NCEE, 2016).

Teachers can be trained at universities or teachers' colleges, undertaking programs that involve four years of coursework. Primary teachers generally attend the specialty colleges and secondary teachers the universities. Coursework typically includes curriculum content, pedagogy, and practicum components. Candidates are awarded a bachelor's degree on successful completion. After completing an approved degree, teaching graduates must pass a national employment examination before they can be assigned a teaching position. The exam has two components, with only the top performing graduates progressing to the second phase. The first component is a written test that is made up of educational pedagogy (30%) and subject-specific content (70%). The second component is an essay and short interview. Those who studied at a general university must also pass an additional pedagogical examination to gain authorisation (Korean Educational Development Institute, 2010). Universities or teachers' colleges are the only pathway for teacher training (Educational Testing Service, 2003).

The results of these assessments are used to determine teacher recruitment and assignment, which for national and public schools is the responsibility of provincial and local government agencies (Korean Educational Development Institute, 2010). While private schools have the opportunity to use other indicators (e.g., interviewing candidates), they generally rely solely on examinations for hiring new teachers (Carnoy, Brodziak, Luschei, Beteille, & Loyalka, 2009).

TEACHER REGISTRATION

MEST licences teachers and determines the requirements for ITE programs and certification. Teacher certifications are then issued by registered universities and are effectively a license to teach. Once certification has been issued, teachers are required to complete examinations prior to employment. It is commonplace to assign teachers tenure at the beginning of their teaching career.

There are two stages of registration, with all teachers beginning as 'Grade-2' and only able to progress to 'Grade-1' after they have been working for three years and have participated in the necessary in-service teacher training courses (Korean Educational Development Institute, 2010). Specifically, teachers must spend two weeks in *pre-employment training*, which includes things like case studies and instruction in classroom management, and six months in *post-employment training*, which involves instructional guidance and classroom supervision (Kim & Han, 2002). The training occurs simultaneously with their teaching, and teachers are still recognised as being 'grade 1' teachers during this time, as there is no probationary period.

There are no official requirements to maintain certification; however, school districts generally require teachers to engage in ongoing PD to maintain their position.

COURSE ACCREDITATION

The Korean Educational Development Institute accredits all teaching programs (Korean Educational Development Institute, 2010). Courses are scored on a five-point scale, with higher-scoring institutions eligible for additional government funding, while low-scoring programs may face reductions in their student quotas. Courses that have been deemed unsatisfactory may be discontinued if they fail to improve. Appraisals occur every three to five years (Korean Educational Development Institute, 2010).

EVALUATION FRAMEWORK

Teacher evaluation in South Korea is complex and involves several evaluation systems. The *Performance Based Incentive Scheme* is a nationally implemented teacher evaluation model, conducted by local education body

evaluators. This evaluation model is used to determine performance incentives and career progression. At an individual school level, the *Teacher Appraisal for Performance* and the *Teacher Appraisal for Professional Development* are implemented (Korean Educational Development Institute, 2010). These models are more flexible and allow evaluators more freedom in determining appropriate evaluation methods and methodologies.

All teachers in schools governed by the *Elementary and Secondary Education* Act 2009 (Ministry of Education, Science and Technolog, 2009) who have taught in their current school for more than two months must be evaluated. Evaluations must occur once per year as per MEST guidelines. It is recommended that evaluations be conducted at least three months before the end of the school year to allow sufficient time for the participants to use reported results to plan for appropriate professional development for the following year.

METHODS AND METHODOLOGIES

The performance indicators for teaching vary depending on the type of teacher evaluation. Specific methods and tools are required when conducting the *Teacher Appraisal for Performance* and the *Performance Based Incentive Schemes*. However, the methods that are used in the *Teacher Appraisal for Professional Development* are much more flexible (Korean Educational Development Institute, 2010).

Student Outcomes

Although students undergo considerable standardised testing in Korea, these results are mostly used for whole school evaluations and not for individual teacher evaluations. Nonetheless, teachers and schools are encouraged to use this information for self-evaluation and reflection, particularly when determining pathways for professional development.

Direct Observation of Teaching

The principal undertakes classroom observation of teachers, as articulated in *Teacher Appraisal for Performance Development* (Korean Educational Development Institute, 2010). While an observation rubric is been acknowledged in some reports (Korean Educational Development Institute, 2010; Kim & Han, 2002), this information is not available in English.

Survey Ratings

Students and parents are surveyed on their satisfaction as part of the *Teacher Appraisal for Professional Development* (Korean Educational Development Institute, 2010).

Peer ratings are used as part of the *Teacher Appraisal for Professional Development*. Peer evaluation groups comprise a minimum of three teachers (Korean Educational Development Institute, 2010).

Teacher Self-Assessment

Not referenced in reviewed policy documents.

Aggregation and Weighting of Results

Only the weighting of results for the *Teacher Appraisal for Performance* are given, which is composed of 40% Principal + 30% Vice Principal + 30% peer (Korean Educational Development Institute, 2010).

Evaluator Training

Training is mandatory for school principals, vice principals, and those teachers who function as evaluators (i.e., those who are part of the *Teacher Appraisal for Performance*, with the government offering in-service training programs). Furthermore, teachers, parents, and students (who function as evaluators in the *Teacher Appraisal for Professional Development*) are offered training by individual schools and local educational authorities, but participation is not obligatory (Korean Educational Development Institute, 2010).

Internal/External Evaluators

Internal evaluators carry out *Teacher Appraisal for Performance* and *Teacher Appraisal for Professional Development* evaluations. Local educational bodies determine *Performance Based Incentive Schemes* and associated career outcomes in accordance with national criteria (Korean Educational Development Institute, 2010).

USES OF EVALUATION

The Korean teacher evaluation system is used for both summative and formative purposes. The *Performance Based Incentive Scheme* is summative in its aim to identify and financially reward effective teachers. In contrast, the *Teacher Appraisal for Professional Development* is formative, as its primary goal is to provide teachers with feedback and opportunities for improvement. Other uses of the evaluation system are:

- Provision of performance-based incentives, whereby teachers receive salary increases based on their performance appraisal
- Providing information from the professional development appraisal, to make a plan to provide the teacher with training and support (undertaken by the school head and local school office of education)
- Determining promotion outcomes. Promotion is based on the combined results of the evaluation; however, results are not weighted equally. In order to be considered for a promotion, a teacher must gain 100 points on the *Teacher Appraisal for Work Performance*; 70 points on teaching experience and 30 points on development (how points are allocated is not discussed)

CONSIDERATIONS FOR MINORITY GROUPS

Not referenced in reviewed documents.

EDUCATION GOVERNANCE IN THE UNITED STATES OF AMERICA

In the United States the federal government has little influence over the functioning of schools, beyond setting broad expectations regarding education quality and occasionally implementing broad policy initiatives, such as the Race to the Top (RTT), the 'Every Student Succeeds Act' and Teacher Incentive Fund (TIF) grants (U.S. Department of Education, 2012). These initiatives were intended to provide incentives for improvement in teacher performance (U.S. Department of Education, 2016). Beyond the federal government, there are four tiers of governance and accountability in regard to education, which are more integral to the design and implementation of education in the United States. These include: state, local, district and school.

Each state has an overarching Department of Education. The state is broken up into school districts, in which each district has authority over testing and curricula, so long as it aligns with the state and federal expectations for education. Typically, states establish state-wide benchmarks for education that are then delivered through school districts. Education delivery is not consistent between states, and there is a large variation amongst states regarding subject delivery and professional development (Corsi-Bunker, 2016). Other state responsibilities include the accreditation of teaching courses, registration of teachers and setting standards for performance review and development. The National Board issues advanced certification, with over 112,000 teachers nationally having achieved this formal recognition of accomplished teaching. Furthermore, each state is responsible for setting the standards and guidelines for teacher and school evaluation. Schools are responsible for conducting teacher appraisals, typically conducted by the school principal.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

EXECUTIVE SUMMARY

- There is no uniform or coordinated teacher evaluation system in the State of California. While legislature exists compelling schools to perform teacher evaluations, the specific nature of these evaluations is largely open to the discretion of individual school districts, and there is little accountability. Ultimately all evaluations of teachers occur at the school level and vary significantly across the state.
- The California Teachers Association (CTA) evaluation framework provides some guidance on evaluation, however school participation is voluntary. Guidelines regarding the selection of measures, use of multiple measures, and training of evaluators are provided. Specific measures are not provided.
- Teacher standards are set at a state level, with the California Standards for the Teaching Profession providing clear standards for teachers. These standards inform the implementation of the CTA evaluation framework and are intended to influence professional development.
- The CTA framework recommends measures such as student outcomes, observations, survey ratings, classroom artefacts, and contributions to the profession and community. Self-assessment is encouraged.
- The CTA framework aims to be both formative and summative in nature. The framework highlights teacher strengths/weaknesses in combination with the identification of feedback and support. Results may also be used to make personnel decisions. Specific information regarding how results may be used to inform feedback is not provided.

INTRODUCTION

In the State of California it is compulsory for children, aged 6-18, to attend school. The Californian education system is divided into:

- **Elementary School** – Operates for eight years, beginning at either Kindergarten or Prep and continues through to either Year 5 or 6.
- **Middle School** – Operates for three or four years, from Years 7-10 or Years 8-10.
- **Senior secondary school** – Operates for two years, from Year 11-12.

There are 10,041 schools in California (2014-2015). This consists of 5,826 elementary schools, 1,337 high schools, 1,179 charter schools, and 2,432 private schools (Education Data Partnership, 2016).

The student population in California was 6,235,520 students in 2014-2015, having grown by approximately 10,000 students since 2010-2011. The average class size in 2011-2012, when this data for California was last collected, was 24 pupils (Education Data Partnership, 2016).

Governance

See Education Governance in the United States of America, p.65.

STANDARDS AND THEORY

The California Standards for the Teaching Profession (CSTP; CTC, 2009) highlight what is considered by the State to be best practice for teachers. The CSTP are also intended to influence professional development and continual learning for teachers (CTC, 2009). The Standards include:

Table 8 - California Standards for the Teaching Profession

Standard	Criteria
Engaging and Supporting All Students in Learning	1. Using knowledge of students to engage them in learning 2. Connecting learning to students' prior knowledge, backgrounds, life experiences, and interests 3. Connecting subject matter to meaningful, real-life contexts 4. Using a variety of instructional strategies, resources, and technologies to meet students' diverse learning needs 5. Promoting critical thinking through inquiry, problem solving, and reflection 6. Monitoring student learning and adjusting instruction while teaching
Creating and Maintaining Effective Environments for Student Learning	1. Promoting social development and responsibility within a caring community where each student is treated fairly and respectfully 2. Creating physical or virtual learning environments that promote student learning, reflect diversity, and encourage constructive and productive interactions among students 3. Establishing and maintaining learning environments that are physically, intellectually, and emotionally safe 4. Creating a rigorous learning environment with high expectations and appropriate support for all students 5. Developing, communicating, and maintaining high standards for individual and group behaviour 6. Employing classroom routines, procedures, norms, and supports for positive behaviour to ensure a climate in which all students can learn 7. Using instructional time to optimize learning
Understanding and Organizing Subject Matter for Student Learning	1. Demonstrating knowledge of subject matter, academic content standards, and curriculum frameworks 2. Applying knowledge of student development and proficiencies to ensure student understanding of subject matter 3. Organizing curriculum to facilitate student understanding of the subject matter 4. Utilizing instructional strategies that are appropriate to the subject matter 5. Using and adapting resources, technologies, and standards-aligned instructional materials, including adopted materials, to make subject matter accessible to all students 6. Addressing the needs of English learners and students with special needs to provide equitable access to the content
Planning Instruction and Designing Learning Experiences for All Students	1. Using knowledge of students' academic readiness, language proficiency, cultural background, and individual development to plan instruction 2. Establishing and articulating goals for student learning 3. Developing and sequencing long-term and short-term instructional plans to support student learning

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Standard	Criteria
	4. Planning instruction that incorporates appropriate strategies to meet the learning needs of all students 5. Adapting instructional plans and curricular materials to meet the assessed learning needs of all students
Assessing Students for Learning	1. Applying knowledge of the purposes, characteristics, and uses of different types of assessments 2. Collecting and analysing assessment data from a variety of sources to inform instruction 3. Reviewing data, both individually and with colleagues, to monitor student learning 4. Using assessment data to establish learning goals and to plan, differentiate, and modify instruction 5. Involving all students in self-assessment, goal setting, and monitoring progress 6. Using available technologies to assist in assessment, analysis, and communication of student learning 7. Using assessment information to share timely and comprehensible feedback with students and their families
Developing as a Professional Educator	1. Reflecting on teaching practice in support of student learning 2. Establishing professional goals and engaging in continuous and purposeful professional growth and development 3. Collaborating with colleagues and the broader professional community to support teacher and student learning 4. Working with families to support student learning 5. Engaging local communities in support of the instructional program 6. Managing professional responsibilities to maintain motivation and commitment to all students 7. Demonstrating professional responsibility, integrity, and ethical conduct

TEACHER PREPARATION

In California, prospective teachers must gain an accredited Bachelor degree in education or complete an undergraduate course in select subjects followed by a post-graduate degree in teaching or an approved course. Teaching programs operate in various settings, including colleges, universities, community colleges, online schools, K-12 school districts, and certain community or regional service centres.

To teach elementary school in California, teachers must have the multiple subject teaching credential that authorises teaching in preschool, kindergarten, grades 1-12, and classes for adults (Commission on Teacher Credentialing, 2016e). This allows for pre-service teachers to be prepared for the multi-class environment of elementary schools. To teach middle or secondary school in California, teachers must have a single subject teaching credential. Those wishing to teach special education students must hold an Education Specialist Instruction Credential. All teacher candidates must pass:

- All basic skills assessment tests
- The California Subject Exam for Teachers (CSET)

Teacher Effectiveness Systems, Frameworks and Measures: A Review

- Developing English language skills (reading competence assessment)

Furthermore, California requires through legislation all teacher preparation programs to include a *Teaching Performance Assessment* (TPA) to ensure that teacher candidates have the knowledge, skills and capabilities expected of a beginning teacher in California public schools (CTC, 2016f). The CTC has developed the CalTPA to exemplify the *Teaching Performance Expectations* (TPEs) (CTC, 2016a) for beginning teachers based on the CSTP.

Alternate pathways exist as internships or for members of the Peace Corps. University intern programs allow individuals to be the teacher of record at employing schools whilst completing a professional preparation program. District intern programs are administered by employing school districts and do not necessarily involve university coursework. Peace Corps teaching experience is accepted in lieu of a professional teacher preparation program credential (CTC, 2016b).

Novice teachers are issued with a Preliminary Credential, which lasts five years. To continue teaching after this period, they must qualify for a Clear Credential by completing one of two General Education programs, or by being certified by the National Board of Teaching Standards (CTC, 2016e).

TEACHER REGISTRATION

After completing an approved preparation program, the ITE provider formally recommends the new educator to the California Commission on Teacher Credentialing (CTC) for credentialing (CTC, 2016). The educator creates an online profile and responds to Personal Fitness Questions in application for a preliminary or Level 1 teaching credential (valid for 5 years and non-renewable) (CTC, 2016c).

Additional academic requirements and professional development associated with educators' initial preparation programs that assure an ongoing capacity to meet the needs of P-12 students must be completed in order to upgrade to the Clear or Level II Credential. National Board Certification is achieved via application to the National Board for Professional Teaching Standards (NBPTS, 2016a). Clear Credentials and Level II Credentials must be renewed every five years. Renewal requires Personal Fitness Questions to be answered (CTC, 2016c).

Life credentials (issued pre-1985) are still valid unless revoked for cause and do not need to be renewed (CTC, 2016d).

COURSE ACCREDITATION

The CTC accredits ITE programs against the *Accreditation Framework: Educator Preparation in California* (CTC, 2007). The framework addresses the accreditation of colleges, universities and local education agencies for state certification and professional practice in Californian schools, ensuring accountability and high quality effective programs that meet common, national and state program standards.

EVALUATION FRAMEWORK

There is no uniform or coordinated teacher evaluation system in the State of California. Teacher evaluation is driven by the *Stull Act* (California State Government, 1971) which was enacted to ensure that untenured and tenured teachers with unsatisfactory performance ratings receive regular evaluations. Further, the *Stull Act* requires teacher evaluation frameworks to include student growth in measuring teacher effectiveness, although school districts have the flexibility to develop their own evaluation models. However, recent research suggests that compliance with the Stull Act is not consistent, and that teacher evaluation systems often fail to comply with state law (Sandoval, 2015).

The California Department of Education provides information regarding the design and implementation of teacher evaluation systems; however, there is little guidance regarding how to use these resources (California Department of Education, 2016).

69

While there is no consistently implemented teacher evaluation framework, and no state-level guidelines for the construction and application of teacher evaluation systems in California, the California Teacher Association (CTA) provides some guidance in the evaluation of teachers. The CTA evaluation framework serves as a guide for school districts to conduct appraisals, but use of this model is not mandatory.

METHODS AND METHODOLOGIES

The CTA framework assesses teachers in relation to the CSTP. The CTA framework suggests that the frequency and goals of evaluation should vary based on the teacher's level of experience (California Teachers Association, 2012). For example, it is recommended that beginning teachers should be evaluated annually, with appraisals focusing on mentoring and support. The CTA framework suggests that experienced teachers should be evaluated every 3-5 years, with an emphasis on continued growth.

The CTA framework encourages the use of multiple measures to inform teacher evaluation outcomes. Examples of measures recommended in the CTA framework include but are not limited to:

- Student outcomes
- Direct observation of teaching
  - Post-observation dialogues
- Survey ratings
- Contributions to the profession
- Contributions to the school and school community
- Instructional logs
- Lesson plans and analysis of student work
- Teacher set objectives and goals (CTA, 2012)

Student Outcomes

Evidence of student learning is discussed in the CTA evaluation framework as an essential, however complex, component of teacher evaluation (California Teachers Association, 2012). Standardised test scores are not recommended for summative evaluations of teachers in the CTA evaluation framework. Rather, the framework suggests that standardised test results should be used to inform teachers' instructional decisions.

As previously discussed, according to the Stull Act (1971) teacher evaluations are legally required to include student growth as an indicator of teacher effectiveness.

Direct Observation of Teaching

Observations are encouraged in the CTA evaluation framework; however, no rubric or clear observation guidelines are provided. Evaluators are also encouraged to engage in 'post evaluation dialogues', to clarify the teacher's intentions and actions (California Teachers Association, 2012).

Survey Ratings

While included as a potential data source, the CTA framework recommends caution when using survey ratings in teacher evaluation. In particular, the framework highlights test construction considerations, such as sample size, reliability, validity, bias on certain questions, and how the method of dissemination can impact results. The CTA framework states that survey ratings should not be used for summative evaluations, and should be used primarily for formative evaluations and guiding teaching practice.

Teacher Self-Assessment

While the CTA Evaluation Framework encourages teachers to engage in reflection during the evaluation process, teacher self-assessment is not discussed as a method for evaluating teachers.

Other Measures

- **Contributions to the profession:** Evidence that the teacher has contributed to education more generally. Examples include formal awards and recognitions, publications, presentations and leadership roles.
- **Contributions to school and school community:** Evidence that the teacher gives value beyond their classroom. For example, professional support to colleagues, leadership roles, organising additional learning programs for students
- **Instructional logs**: Described as "detailed records of teaching', these may be included in the appraisal.
- **Lesson Plans and Student work**: This involves looking at student work and artefacts (e.g., student presentations) and considering how they contribute to learning and align with the standards. Teachers can also explain how their lessons aim to develop knowledge and skills (CTA, 2012).

Internal/External Evaluators

The CTA framework does not specify whether evaluators should be internal or external to the school. The framework does recommend that evaluators receive training in evaluation procedures and specific evaluation tools. Further, regular calibration is recommended to ensure inter-rater consistency in evaluation results.

Aggregation and Weighting of Results

No information regarding the aggregation or weighting of results is provided in the CTA evaluation framework.

USES OF EVALUATION

No specific guidance regarding the use of evaluation data is provided by the California Department of Education. The CTA framework aims to be both formative and summative in nature. The CTA framework notes that formative assessments may be used to provide feedback on how teachers may improve their practice to promote student learning, and to guide what types of professional development opportunities may enhance their practice (CTA, 2012). The framework also notes that summative evaluations may be used to make decisions on an educator's performance that inform personnel decisions. More specific information regarding the use of evaluation information was not provided.

CONSIDERATIONS FOR MINORITY GROUPS

The evaluation framework does not include considerations for minority groups and cultural competency.

EXECUTIVE SUMMARY

- Washington, D.C. utilises a highly structured and prescriptive teacher evaluation model referred to as IMPACT. This evaluation system is compulsory for all public school teachers teaching Grades 1-12. All teacher evaluations occur at a school level and are mandated by the state.
- Teacher standards are set at a state level by the District of Columbia Public Schools. These standards are used to inform teaching practice evaluation procedures and in particular, classroom observations.
- Measures utilised in the IMPACT model are compulsory and include: direct classroom observation, student achievement, commitment to the school community, and core professionalism. All of these measures must be completed annually for each teacher.
- For teachers with students completing state standardised examinations (the Partnership for Assessment of Readiness for College and Careers (PARCC), completed by students in Grades 3-12), Individual Value-Added Student Achievement Data (IVA) is calculated. These results are typically factored in to IMPACT results.
- Measures are aggregated to form a final IMPACT score. The weighting of individual components varies based on the availability of PARCC results and the type of teacher being evaluated.
- School-based evaluations are intended to identify and support underperforming teachers and are linked to salary decisions and career progression. As such, results are largely summative at the school level.

INTRODUCTION

In Washington, D.C. it is compulsory to attend primary and secondary school between the ages of five and eighteen. School education (pre-tertiary) lasts for 12 years and is most commonly divided into:

- **Elementary School** – Operates for eight years, from ages 4-11 (K-Grade 5)
- **Middle School and Junior High** – Operates for three years, from ages 12-14 (Grade 6-8)
- **Senior secondary school** – Operates for four years, from ages 15-18/19 (Grade 9-12)

There are some variations to these categorisations.

There are 772 schools in the DC area: 115 Public Charter Schools, 126 Public District and 531 Private Schools (Great Schools, 2016).

Governance

See Education Governance in the United States of America, p.65.

STANDARDS AND THEORY

The Teaching and Learning Framework is set by the District of Columbia Public Schools, and guides teaching practice and evaluation (District of Columbia Public Schools, 2015c). The Teaching and Learning Framework encompasses three components, including: the planning, teaching, and performance review and development, outlined in more detail in Table 9.

Table 9 - The District of Columbia Public Schools Teaching and Learning Framework

Standard	Criteria
Plan	1. Develop annual student achievement goals 2. Create standards based unit plans and assessments 3. Create objective driven lesson plans
Teach	1. Lead well-organized, objective-driven lessons 2. Explain content clearly 3. Engage students at all learning levels in accessible and challenging work 4. Provide students multiple ways to move towards mastery 5. Check for student understanding 6. Respond to student understanding 7. Develop higher-level understanding through effective questioning 8. Maximize instructional time 9. Build a supportive, learning-focused classroom community
Increase effectiveness	1. Assess student progress 2. Track and analyse student progress data 3. Improve proactive and re-teach in response to data

TEACHER PREPARATION

In Washington D.C. teachers are required to complete a Bachelor's Degree in Education or a relevant content area and a postgraduate education degree (e.g. Master of Education), as well as standardised tests (OSSE, 2016). A general credit hour requirement for all teaching certificates is a minimum 48 semester hours in a program of general or liberal education, with additional requirements for particular subject or grade level certifications and a practicum component. Teacher candidates in Washington D.C. must achieve pass results on the Praxis Core Academic Skills Test. Content examined in licensure exams includes:

- Basic skills exam (reading, writing and mathematics)
- Content knowledge exam (teaching area)
- Pedagogy/Principles of learning and teaching exam for the applicable subject area/grade level

Many colleges use Praxis I Pre-Professional Skills Test scores to evaluate candidates' eligibility for entry to their teacher preparation programs.

TEACHER REGISTRATION

Prospective teachers are required to register with Office of the State Superintendent of Education (OSSE) in order to teach within the Washington D.C. area (OSSE, 2016c). Washington, D.C. issues a Regular I license (2 years, non-renewable) to beginning teachers who have a Bachelor's degree and have completed or are enrolled in a teacher preparation program and have passed the Praxis I exam and Praxis II content exam, and a Regular II (4-year renewable) license to teachers who have completed all DC education requirements and passed the Praxis II Pedagogy exam (Office of the State Superintendent of Education, 2014).

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Teachers from any state approved teacher preparation program are recognised by Washington D.C. as meeting the basic requirements as part of an Interstate Reciprocity Program (Office of the State Superintendent of Education, 2016b).

COURSE ACCREDITATION

ITE programs in the United States are required to be accredited with a relevant state body. In Washington, D.C., ITE accreditation is performed by the District of Columbia, Office of the State Superintendent of Education (OSSE, 2016a). To become accredited, ITE providers are required to demonstrate delivery of industry-recognized standards in child development, classroom management, and content knowledge. Additionally, programs must demonstrate objective and verifiable standards that must be successfully completed to qualify a candidate for the Regular Teaching Credential. Specifics regarding this process are not defined in the documents reviewed.

EVALUATION FRAMEWORK

In Washington, D.C., public school teachers are evaluated using a high-stakes, incentive-based evaluation system called IMPACT (District of Columbia Public Schools, 2016a). This evaluation system has three purposes: to outline clear performance expectations, provide clear feedback, and to ensure that every teacher has a plan for improving their teaching performance (District of Columbia Public Schools, 2016a; Headden, 2011). Strict guidelines are provided regarding the methods and methodologies utilised to evaluate teachers (District of Columbia Public Schools, 2015a, 2015b).

METHODS AND METHODOLOGIES

The measures used in the IMPACT model are strictly prescribed (District of Columbia Public Schools, 2016a). Multiple measures of teacher effectiveness are outlined in the IMPACT manuals (District of Columbia Public Schools, 2015a, 2015b), including:

- Direct classroom observation
- Student outcomes
  - Individual Value-Added scores, and/or
  - Other tests of student achievement
- Commitment to the School Community (CSC)
- Core Professionalism (CP)

Results from these measures are aggregated to form an overall rating of teacher effectiveness (District of Columbia Public Schools, 2016a; Headden, 2011). The weighting of each measure varies depending on the type of teacher being evaluated (District of Columbia Public Schools, 2015a, 2015b). Teachers are categorised into one of seven groups, with the majority of grades 1 to 12 teachers categorised as either:

- Group 1: general education teachers with individual value-added student achievement data (classes with standardised state tests)
- Group 2: grade 1-12 general education teachers with no standardised tests (District of Columbia Public Schools, 2016a; Headden, 2011)

Student Outcomes

Student achievement, referred to as Teacher-Assessed Student Achievement Data (TAS), is utilised in the IMPACT model as an indicator of teacher effectiveness (District of Columbia Public Schools, 2015a, 2015b). This includes results on rigorous assessments of student achievement, which is not defined in the IMPACT guides (District of Columbia Public Schools, 2015a, 2015b).

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Additionally, for teachers whose students complete the Partnership for Assessment of Readiness for College and Careers (PARCC) standardised assessment (select subjects in grades 3-12), Individual Value-Added Student Achievement Data (IVA) is calculated (District of Columbia Public Schools, 2016b, 2016d). This provides a score whereby expected test results are compared to actual results. While these results typically contribute to IMPACT results, IVA has not been included in the previous two IMPACT evaluation cycles (District of Columbia Public Schools, 2014a, 2015a).

Direct Observation of Teaching

The Teaching and Learning Framework (TLF) is used to guide direct observation of teaching. The IMPACT guides provide comprehensive rubrics, based on the TLF, with observers scoring teachers on a four-point scale for effectiveness ranging from highly effective to ineffective. Teachers are assessed on their capacity to:

- Lead well-organised objective-driven lessons
- Explain content clearly
- Engage students at all learning levels
- Provide students multiple ways to move to mastery
- Check for student understanding
- Respond to student understanding
- Develop higher-level understanding
- Maximise instructional time
- Build a supportive, learning-focused environment

Teachers are required to be formally observed several times a year; however, informal observations also occur, primarily for learning and development. The frequency of observations varies based on the teacher's level of expertise and role within the school, ranging from 1-6 observations per year. Observations are typically unannounced and typically last for approximately 30 minutes.

Internal school personnel conduct observations (usually the school principal or senior teacher) as well as external evaluators. There is no information on the proportion of evaluations conducted internally versus externally.

Survey Ratings

Survey ratings are not used to inform teacher evaluations.

Teacher Self-Assessment

Teacher self-assessments are not used to inform teacher evaluations.

Other Measures

Commitment to the School Community (CSC) is measured as part of teachers' IMPACT evaluation. Teachers are assessed twice per year on five domains, including:

- Support of the local school initiatives
- Support special education and English Language Learner programs
- High expectations (i.e. promoting high academic and behavioural expectations through innovative teaching practice)
- Partnership with families
- Instructional collaboration

Teacher Effectiveness Systems, Frameworks and Measures: A Review

The evaluator (usually the school principal or senior teacher) rates these areas on a four-point scale. An extended rubric expands on the requirements for each scoring level of the CSC domains, providing further detail for evaluators. Scores from both time points in the year are averaged and form part of a teacher's overall IMPACT score.

Core Professionalism (CP) is also evaluated as part of the IMPACT process. Teachers are assessed twice per year on four domains, including:

- Attendance
- On-time arrival
- Ability to adhere to policies and procedures
- Respect

The evaluator (usually the school principal or senior teacher) rates these areas on a three-point scale, ranging from 'meets standard' to 'significantly below standard'. An extended rubric expands on the requirements for each scoring level of the CSC domains, providing further detail for evaluators. CP ratings do not form part of their overall IMPACT score. However, a rating of slightly below on any of the CP areas results in a ten point deduction, while a rating of significantly below incurs a twenty point reduction in a teacher's overall IMPACT score.

Internal/External Evaluators

Both internal and external evaluators assess teachers. Administrators within the school (i.e., principals or assistant principals) and master educators (third party external evaluators) conduct formal observation. A master educator is described as an expert practitioner in a particular content area who will serve as an impartial observer (District of Columbia Public Schools, 2016c).

Aggregation and Weighting of Results

The weighting of results in the IMPACT system vary from year to year, however in the previous two evaluation cycles (2014-2015 and 2015-2016) the weighting of results for both Group 1 and Group 2 has been (District of Columbia Public Schools, 2014a, 2014b, 2015a, 2015b):
- 75% TLF (observation)
- 15% TAS (student achievement)
- 10% CSC

For years where IVA has been available, IMPACT results for Group 1 are generally calculated as follows:
- 35% IVA
- 40% TLF (observation)
- 15% TAS (student achievement)
- 10% CSC

USE OF EVALUATION

IMPACT results are used for both formative and summative purposes. Results inform decisions regarding teacher employment, including decisions regarding the dismissal and promotion of teachers. Additionally, teachers who receive high IMPACT results are also eligible to receive monetary bonuses of up to $25,000 per year.

Results also help schools in understanding the areas of need with regard to professional development and the types of professional development. The data also helps to inform DCPS's plans around catering support to certain school areas or communities where there is certain need. DCPS provides resources and both online and in person professional development support.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

CONSIDERATIONS FOR MINORITY GROUPS

The IMPACT Framework does not explicitly state culturally responsive or disability responsive teaching practice as criteria for teacher effectiveness.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

EXECUTIVE SUMMARY

- Teacher evaluation is a legal requirement in Virginia. While certain features (e.g. principal training in evaluation and the inclusion of student achievement as a measure of teacher effectiveness) are legally required, the design and implementation of teacher evaluation processes is ultimately the responsibility of individual schools.
- The Virginia Department of Education has set a guiding framework for evaluation known as the Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers, which sets out recommendations for evaluation procedure and method and assists principals in the design and implementation of teacher evaluation frameworks.
- The Virginia Standards for the Professional Practice of Teachers outline the expectations for teachers in Virginia. These are strongly evidence based and directly linked to teacher evaluation targets.
- Evaluation tools are recommended by the Guidelines and include student outcomes, observations, student surveys, teacher portfolios, performance artefacts, and self-assessment.
- School-based evaluations are intended to identify and support underperforming teachers and are linked to promotion decisions. Evaluation results are summative with regards to informing promotion and career decisions, but formative with regards to providing feedback for professional development.

INTRODUCTION

In the state of Virginia, it is compulsory to attend school between the ages of five to eighteen. Schooling is typically divided into:

- **Elementary School** – Operates for six years from Grades K-5 for children aged 4-11 years
- **Middle School and Junior High** – Operates for three years from Grades 6-8 for children aged 12-14 years
- **Senior secondary school** – Operates for four years from Grades 9-12 for children aged 15-19 years

As of 2014 there were 1,564,646 school aged children within the State of Virginia (Group, 2014). In the 2015-2016 school year there were 1856 local schools registered in the Virginia area (Virginia Department of Education, 2016c). These comprise:

- 33 Preschools
- 1156 elementary schools
- 305 middle schools
- 52 combined
- 250 Alternative education Centres

Governance

See Education Governance in the United States of America, p.65.

The Department of Education in Virginia sets standards for teacher professional practice. The Board of Education is in charge of setting state-wide curriculum standards and guidelines, school accreditation and teacher education and performance requirements.

STANDARDS AND THEORY

The Virginia Standards for the Professional Practice of Teachers (SPPT) outline the expectations for teachers in Virginia (Virginia Department of Education, 2011c). These standards are strongly rooted in empirical research, with

the (Virginia Department of Education, 2011b) providing evidence and justification for inclusion of each of the performance standards. The Virginia SPPT focus on teachers' ability to provide an effective learning space that promotes student outcomes. In addition to broad standards for all teachers, the Virginia SPPT also contain specific standards for multiple subject areas. Each standard contains performance indicators that may guide senior staff in evaluating teacher performance. The standards for all teachers include:

Table 10 - The Virginia Department of Education Standards

Standard	Criteria
Professional Knowledge	Teachers demonstrate an understanding of the curriculum, subject content, and the developmental needs of students by providing relevant learning experiences.
Instruction Planning	Teachers plan using the Virginia Standards of Learning, the school's curriculum, effective strategies, resources, and data to meet the needs of all students.
Instructional Delivery	Teachers effectively engage students in learning by using a variety of instructional strategies in order to meet individual learning needs.
Assessment of and for Student Learning	Teachers systematically gather, analyse, and use all relevant data to measure student academic progress, guide instructional content and delivery methods, and provide timely feedback to both students and parents throughout the school year.
Learning Environment	Teachers use resources, routines, and procedures to provide a respectful, positive, safe, student-centred environment that is conducive to learning.
Professionalism	Teachers maintain a commitment to professional ethics, communicate effectively, and take responsibility for and participate in professional growth that results in enhanced student learning.
Student Academic Progress	The work of the teacher results in acceptable, measurable and appropriate student academic progress.

TEACHER PREPARATION

The state of Virginia requires teachers to complete a state-approved education preparation program, offered at more than thirty public and private colleges and universities (Virginia Department of Education, 2016b). A minimum bachelor level is required for certification. Alternative pathways to licensure are offered to career switchers who may undertake state-certified programs (Virginia Department of Education, 2016e). Virginia's Troops to Teachers program offers stipends to interested military personnel to enrol in an approved program in return for a three-year commitment in high-need school divisions (Virginia Department of Education, 2016e).

Teacher education programs in Virginia typically involve combined academic studies and fieldwork, which may include an internship. The state mandates subject specific requisite skills and a minimum 300-375 hours field experience, 150 or more supervised (Teacher.org, 2016).

TEACHER REGISTRATION

Teaching licenses are required for educators in all public schools in Virginia, and most private schools. Teachers are able to gain a license once they have completed a relevant bachelor degree, have passed the teacher knowledge tests, and satisfied safety criteria (e.g., first aid); these need to be renewed every five years (Virginia Department of Education, 2016f). The Virginia Department of Education (VDOE) is responsible for granting teaching licenses (VDOE, 2016d).

There are also alternate routes to licensure, mostly for those who hold a Bachelor degree but not a teaching degree. Eligibility is dependent on previous experience, the subject they wish to teach, and the need for teachers in that specific area. There are also options for a fast-track teaching program.

Teacher knowledge tests are part of the licences criteria. Assessments vary depending on the subject areas they may teach, but the main tests required for initial license are:

- Academic skills assessment (Virginia Communication and Literacy Assessment)
- Content Assessment
- Reading assessment (for elementary and special education teachers)

Teachers from other states may have to undergo additional testing due to the lack of standardisation of higher education within the country.

Teachers must participate in professional development to retain their registration. There are seven options for renewal, and each is given a certain number of points (depending on its difficulty, number of hours, etc.), and teachers must get a total of 180 points.

1. College credit
2. Professional conference
3. Curriculum Development
4. Publication of Article
5. Publication of Book
6. Mentorship/Supervision
7. Educational Project
8. Professional Development Activities (VDOE, 2016f)

COURSE ACCREDITATION

Teacher education in the United States varies by program content and quality. The State of Virginia requires all education programs to be approved by the VDOE. Furthermore, teachers who wish to teach within the state, but who have not been educated at an education institution with the state of Virginia, must first have their education program approved as fitting the state requirements for teacher education before they can take up a position as a teacher in Virginia.

The state requires that prospective teachers must have completed an undergraduate degree and also complete two state mandated tests to gauge teacher competency. These tests are the Passing a basic skills test, Virginia Communication and Literacy Assessment (VCLA), and subject competence assessments. Subject competence tests involve Academic Skills Assessment, Content Assessment, and Reading Assessment (for Elementary and Special Education teachers).

Teachers from other states may have to undergo additional testing due to the lack of standardisation of higher education within the country.

EVALUATION FRAMEWORK

Teacher evaluation is a legal requirement for all teachers in Virginia (Virginia Department of Education, 2011a). Principals are required to be trained in and subsequently engage in the evaluation of their teaching workforce. While the statutory requirements do not specify the exact methods by which teachers should be evaluated, teacher evaluations are legally required to, at a minimum, consist of measures of teacher skills and knowledge as well as student academic progress. To assist in the implementation of teacher evaluation systems, Virginia receives partial funding from the Federal government. However, schools are still required to commit a large amount of resources in the implementation of teacher evaluation systems.

To assist schools in the design and implementation of teacher evaluation systems, the VDOE has constructed a guiding framework, referred to as the Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers, hereafter referred to as the Guidelines (Virginia Department of Education, 2011a). This framework was developed in conjunction with the Virginia Standards for the Professional Practice of Teachers, and is similarly evidence-based. The Guidelines recommend that probationary teachers undergo annual evaluations, while teachers under contract undergo evaluations every three years.

METHODS AND METHODOLOGIES

Principals typically conduct teacher evaluations and are expected to utilise a variety of data gathering methods. The VDOE provides principals with training in teacher evaluation procedures (Virginia Department of Education, 2016a). Recommendations for specific methods and methodologies are also provided in the Guidelines (Virginia Department of Education, 2011a). The Guidelines recommend collecting data from multiple sources, which include, but are not limited to:

- Student outcomes
- Observations
  - Formal
  - Informal
- Student Surveys
- Portfolios
- Performance Artefacts
- Self-Evaluation

Student Outcomes

It is a legal requirement that student achievement data are included in teacher evaluation results in Virginia (Virginia Department of Education, 2011a). However, no specific student achievement measures are mandated in any legislative documents allowing schools some flexibility in the selection and use of appropriate measures of student achievement (Virginia Department of Education, 2011a).

The Guidelines recommend that measures of student achievement should be selected with the specific type of teacher in mind (Virginia Department of Education, 2011a). For example, the student achievement measures used for a 4th Grade Mathematics teacher should be different to the student achievement measures used by a 7th Grade Art teacher. Where available, however, the Guidelines recommend the inclusion of data from standardised tests (Virginia Department of Education, 2011a). The Guidelines also recommend that student academic progress account for 40 per cent of a teacher's summative evaluation (Virginia Department of Education, 2011a).

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Direct Observation of Teaching

The Guidelines recommend use of both formal and informal observations in evaluating teachers (Virginia Department of Education, 2011a). Formal observations are defined as being highly structured, 30-40 minute direct classroom observations (Virginia Department of Education, 2011a). Informal observations are also encouraged, and may be less structured (Virginia Department of Education, 2011a). As described by the Guidelines, formal observations serve as both summative and formative tools, while informal observations are generally only recommended for formative use (Virginia Department of Education, 2011a).

Sample observation rubric forms are provided in the Guidelines for both formal and informal observations (Virginia Department of Education, 2011a). These example forms prompt observers to consider each of the principles of the Virginia Standards for the Professional Practice and note specific examples of each that are observed (Virginia Department of Education, 2011a, 2011c).

The Guidelines discourage walk-throughs – brief 3-5 minute visits to multiple classrooms (Virginia Department of Education, 2011a). The Guidelines warn that even with multiple walk-throughs, an accurate representation of teacher practice will be difficult to attain (Virginia Department of Education, 2011a).

Survey Ratings

The Guidelines recommend the collection of student survey information, primarily to inform formative evaluations (Virginia Department of Education, 2011a). Examples of student surveys, as well as survey summary forms are provided in the Guidelines (Virginia Department of Education, 2011a). There are four different versions of the student survey (Grades 1-2, 3-5, 6-8, and 9-12) designed to reflect developmental differences in students' ability to provide useful feedback to their teacher (Virginia Department of Education, 2011a).

Teacher Self-Assessment

Teacher self-evaluation is strongly recommended in the Guidelines (Virginia Department of Education, 2011a). A sample teacher self-evaluation form is provided in the Guidelines (Virginia Department of Education, 2011a). This example form prompts teachers to consider each of the principles of the Virginia Standards for the Professional Practice and reflect on areas of strength, areas requiring improvement, and strategies for improving performance (Virginia Department of Education, 2011a, 2011c).

Internal/External Evaluators

Both the Guidelines and Virginia law specifically reference the role of the school principal in performing teacher evaluations (Virginia Department of Education, 2011a). As such, teacher evaluation is typically considered an internal school-based process in Virginia. Under certain conditions, such as when a conflict of interest is present or when there has been a serious complaint, third-party evaluators may be used; however, this is not common (Virginia Department of Education, 2011a). An emphasis is also placed on training of evaluators, with training of principals in evaluation procedures required by law (Virginia Department of Education, 2011a, 2012).

Aggregation and Weighting of Results

The Guidelines suggest that student learning, as determined by multiple measures of academic achievement, should account for 40% of a teacher's evaluation result (Virginia Department of Education, 2011a). It is also recommended in the Guidelines that value-added data account for at least 20% of evaluation results where such data is available (Virginia Department of Education, 2011a). Finally, the Guidelines recommend that another 20% (or half the weighting of measures of student learning) should be measured using one or more alternative measures, as outlined previously (Virginia Department of Education, 2011a).

USES OF EVALUATION

The teacher evaluation process in Virginia is both summative and formative in nature. The resulting grades from teacher evaluation can inform decisions regarding promotion or dismissal (Virginia Department of Education, 2011a). Following the evaluation, teachers are classified as: (1) Exemplary; (2) Proficient; (3) Developing/Needing Improvement; (4) Unacceptable, based on their performance relative to the teaching standards (Virginia Department of Education, 2011a, 2011c).

If a teacher's performance does not meet the expectations established by the school, the Guidelines recommend placing the teacher on a Performance Improvement Plan (Virginia Department of Education, 2011a). These programs aim to support teachers by providing them with targeted support, supervision and resources (Virginia Department of Education, 2011a). The evaluator must provide the teacher with the areas of concern in writing, and work with them to improve (Virginia Department of Education, 2011a). If sufficient improvement is achieved, the teacher is no longer required to remain on the plan (Virginia Department of Education, 2011a). If there are no significant gains in their performance, they must remain on the plan (Virginia Department of Education, 2011a).

The Guidelines place a strong emphasis on performance feedback (Virginia Department of Education, 2011a). Teachers receive either written or verbal feedback from the evaluator following their evaluation (Virginia Department of Education, 2011a). This is used to inform professional development (Virginia Department of Education, 2011a).

CONSIDERATIONS FOR MINORITY GROUPS

The Guidelines do not explicitly include demographic specific criteria (Virginia Department of Education, 2011a). They do, however, include 'the promotion of cultural sensitivity' as an element of gauging whether a teacher provided a positive classroom space for learning (Virginia Department of Education, 2011a).

EXECUTIVE SUMMARY

- There are broad state-based systematic evaluation processes for teacher quality and effectiveness across Washington State. Local school districts are required by the state to establish and implement evaluation programs. Legislature outlines what is required of evaluation but leaves methods to the discretion of the school district.
- State teacher standards are directly related to evaluation requirements. The standards used to evaluate teachers change with the school district, as there are three optional frameworks to choose from. These include the *CEL 5D+™ Teacher Evaluation Rubric 2.0, Danielson's Framework for Teaching, or The Marzano Teacher Evaluation Model.*
- Student growth and observation are tools required by state legislature. Other recommended tools include self-assessment, student perceptions, and peer perceptions.
- There is little to no systematic training or evaluation of appraisers. There is also no inter-rater reliability or moderation in this process. There is a need to build the capacity of school leaders as appraisers for effective implementation of evaluation in order to improve the quality and effectiveness of teaching.
- Evaluation results are both summative and formative. Results are used to determine whether teacher performance is satisfactory, as well as to identify and target professional development.

INTRODUCTION

In Washington it is compulsory to attend primary and secondary school between the ages of eight to eighteen (Office of Superintendent of Public Instruction, 2016a). School education (pre-tertiary) lasts for a minimum of 10 years and is divided into:

- **Elementary School** – Operates for six years from Grades K-5 for children aged 4-11 years
- **Middle School and Junior High** – Operates for three years from Grades 6-8 for children aged 12-14 years
- **Senior secondary school** – Operates for four years from Grades 9-12 for children aged 15-19 years

The Office of Superintendent of Public Instruction (2016) reports that 1,088,959 students are enrolled in schools in the state of Washington.

In Washington State, there are 9 Educational Service Districts (ESDs) that oversee a total of 295 school districts. The state partially funds the evaluation initiative, putting funds into the training of teachers in the program (Office of Superintendent of Public Instruction, 2015).

Governance

See Education Governance in the United States of America, p.65.

The Washington State Legislature has determined that it is the state's responsibility to create a coherent and effective accountability framework for schools (Washington State Legislature, 2010). The legislature has assigned the state board of education responsibility and oversight for creating an accountability framework (Washington State Legislature, 2010). Each common school district board of directors has the responsibility for setting policies regarding teacher evaluation, including performance criteria and evaluation processes (Washington State Legislature, 2010).

Washington State has a set of clearly defined standards and criteria for performance and accountability outlined in the education reform bill 'E2SSB 6696', passed by the Washington State legislature (Washington State Legislature, 1991e). Minimum teacher evaluation criteria as described in the Washington State legislature (Washington State Legislature, 1991e) are:

Table 11 - The Washington State Education Standards

Standard	Description
Instructional skill	Being able to competently design and teach day-to-day classes, demonstrating adequate knowledge to do so.
Classroom management	The teacher is able to manage their classroom environment.
Professional preparation and scholarship	The teacher displays theoretical education knowledge as well as a commitment to the profession.
Effort toward improvement when needed	Demonstrates professional growth where needed.
The handling of student discipline and attendance problems	The certificated classroom teacher demonstrates the ability to manage the non-instructional, human dynamics in the educational setting.
Interest in teaching pupils	The teacher actively applies themselves and is committed to the learning of all of their students.
Knowledge of subject matter	Is appropriately knowledgeable and skilled in the subject areas that they are teaching.

TEACHER PREPARATION

Washington State requires a minimum of a four-year bachelor's degree and completion of a state-approved teacher preparation program, available through more than twenty accredited colleges and state approved colleges and universities (Teach.com, 2015).   Alternative pathways, which tend to be shorter, are available for career switchers who hold a bachelor's degree, district staff employed on conditional certificates, and para-educators with Associate's degrees. Pre-residency clearances are required for student teachers to undertake field experience in Washington State's schools (OSPI, 2016f).

Applicants for Washington State's approved teacher preparation programs must have an acceptable Grade Point Average and pass a basic skills test.

TEACHER REGISTRATION

On completion of all degree requirements, candidates are eligible to apply for full residency certification, which is the first level regular teaching certificate issued by the State of Washington Office of Superintendent of Public

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Instruction. The state requires candidates for residency to pass a specific Washington State basic skills test (WEST-B) and content knowledge test (WEST-E) (OSPI, 2016i). The Residency Certificate is valid for two years and may be reissued up to five years (OSPI, 2016g). The second level of certification is the Professional Certificate, which is issued to teachers who also complete a Pro-Teach Portfolio or achieve advanced credentialing through National Board certification (OSPI, 2016h).

Registration is required in order to teach in the state of Washington. Applicants are required to pass a skills test in order to gain entrance into a teacher training program; which is also a requirement following the completion of their studies in order to gain teacher certification (OSPI, 2016i). This includes testing of mathematics skills as well as reading and writing.

There are two levels of certification:

- Residency
- Teacher's certificate- that a prospective teacher may apply for after their first year

The residency teacher certificate is issued to first time and out of state teachers. Official transcripts and documentation must be submitted by foreign-trained and out of state applicants in order to verify their credentials (OSPI, 2016i). Certification in Washington involves first gaining a residency certificate which remains valid until they are reported as being 1.5 FTE or with more years of experience (OSPI, 2016i).

Residency Teaching Certificates may be renewed twice and last for two years upon renewal. Renewal requires completion of the online ProTeach assessment or of the National Board for Professional Teaching Standards (NBTS) Assessment (OSPI, 2016g). Professional Teacher Certificates require completion of annual Professional Growth Plans, which relate to standards and career benchmarks (OSPI, 2016b).

Teachers are considered provisional in the first three years of teaching, in the first three years of teaching in a new school district, and for one year if the teacher has two or more years of experience teaching in Washington State (Washington State Legislature, 1991c).

COURSE ACCREDITATION

Teachers in Washington State are required to have a tertiary degree in order to be considered for teacher registration. All teacher education programs must be first approved by the Washington State Achievement Council (OSPI, 2016d). The State offers different pathways to teacher accreditation which includes State approved tertiary education programs and alternative route programs (OSPI, 2016i).

EVALUATION FRAMEWORK

In Washington State, local school districts are legally required to establish and implement evaluation programs for classroom teachers (Washington State Legislature, 1991d). The intent of teacher evaluation programs, as outlined in Washington State legislature, is to: identify areas for improvement, assist teachers in making improvements, and identify teachers with unsatisfactory performance requiring remediation (Washington State Legislature, 1991g). Each district must report annually to the Superintendent of Public Instruction with data regarding: evaluation criteria and rubrics, descriptions of each rating, and the number of staff in each rating (Washington State Legislature, 2010).

State legislature has determined that the criteria for evaluation must focus on (Washington State Legislature, 2010):

1. Centring instruction on high expectations for student achievement
2. Demonstrating effective teaching practices
3. Recognizing individual student learning needs and developing strategies to address those needs
4. Providing clear and intentional focus on subject matter content and curriculum
5. Fostering and managing a safe, positive learning environment
6. Using multiple student data elements to modify instruction and improve student learning
7. Communicating and collaborating with parents and school community

Teacher Effectiveness Systems, Frameworks and Measures: A Review

8. Exhibiting collaborative and collegial practices focused on improving instructional practice and student learning

While the legislature serves to outline the broad requirements of teacher evaluation programs, details regarding specific methods and methodologies are left largely open to school district discretion (Washington State Legislature, 2010).

In response to Washington State legislature, the Office of the Superintendent of Public Instruction (OSPI) has developed the Teacher/Principal Evaluation Program (TPEP), which guides school districts in the design and implementation of teacher evaluation systems (Office of Superintendent of Public Instruction, 2016d). School districts are required to choose from three evaluation frameworks: the CEL 5D+ Teacher Evaluation Rubric 2.0, the Danielson Framework for Teaching, and the Marzano Teacher Evaluation Model (Office of Superintendent of Public Instruction, 2016c). These three models were selected as they were considered most compatible with the state teacher evaluation criteria (Office of Superintendent of Public Instruction, 2015). School districts may make amendments to their chosen framework to better suit the educational context of their schools, however all changes must first be cleared with the OSPI (Office of Superintendent of Public Instruction, 2016d). The OSPI also provides a web-based platform, eVAL, for recording and managing any evaluation data for both the teachers and evaluators (Washington State Teacher/Principal Evaluation project, 2016b). Use of this platform is not mandatory.

The assessment criteria examined within each of the OSPI endorsed evaluation models are outlined in Table 10 (Danielson, 2013b; Learning Sciences International: Learning and Performance Management, 2016; Washington State Teacher/Principal Evaluation Project, 2016a).

Table 12 - Washington State Teacher Evaluation Models

Model	Criteria
CEL 5D+™ Teacher Evaluation Rubric 2.0	1. Purpose: Setting a clear, meaningful course for student learning 2. Student engagement: Encouraging substantive, intellectual thinking 3. Curriculum and pedagogy: Ensuring that instruction challenges and supports all students 4. Assessment for student learning: Using ongoing assessment to shape and individualise instruction 5. Classroom environment and culture: Creating classrooms that maximise opportunities for learning and engagement
Danielson's Framework for Teaching	1. Planning and Preparation 2. The Classroom Environment 3. Instruction 6. Professional Responsibilities
The Marzano Teacher Evaluation Model	1. Classroom strategies and behaviour 2. Planning and Preparing 3. Reflecting on Teaching 4. Collegiality and Professionalism

The frequency and specific methods for teacher evaluation vary somewhat based on the teacher's level of experience. Annual evaluations are typically required of teachers. After four years of satisfactory evaluations, in which a teacher receives one of the top two ratings, teachers are eligible for a short-form evaluation (Washington State Legislature, 2010). The short form of evaluation includes either a thirty minute observation during the school year with a written summary or a final annual written evaluation based on the criteria in subsection (1) or (2) of this section and based on at least two observation periods during the school year totalling at least sixty minutes without a written summary of such observations being prepared (Washington State Legislature, 2010).

Teacher Effectiveness Systems, Frameworks and Measures: A Review

State legislature provides some guidance regarding the specific methods for teacher evaluation. According to Washington State legislature, teacher evaluation programs must at least include (Washington State Legislature, 1991i, 2010):

- Student Outcomes (Student Growth), and
- Observation

On each measure, teachers are required to be evaluated using a four-level rating system, with performance described along a continuum that indicates the extent to which the criteria have been met or exceeded (Washington State Legislature, 2010). The Washington State legislature describes these levels as: 1) Unsatisfactory, 2) Basic, 3) Proficient, and 4) Distinguished (Washington State Legislature, 1991m).

The school district selects the guiding framework for evaluation of specific teaching methods and methodologies. As these frameworks are discussed later in this document, this section will primarily focus on the specific legal requirements for evaluation as outlined by the Washington State legislature and TPEP framework.

The TPEP framework emphasizes the use of multiple evidence points in determining overall teacher evaluation outcomes. While school districts may differ in their preferred method of appraisal, they must all adhere to the evaluation criteria as described in the state legislature on teacher evaluation (Washington State Legislature, 1991f, 2010). The school context and the level of teacher expertise determines the appropriateness of these sources. Other methods recommended by the TPEP framework and state legislature for teacher evaluation include (Office of Superintendent of Public Instruction, 2016d):

- Observation
- Student growth measures
    - Classroom-based assessments
- Student perceptions (Washington State Legislature, 1991a)
- Input from building staff

Student Outcomes

Student outcomes are an important measure of teacher performance for the Washington teacher evaluation framework. Student growth measures (i.e. the change in student achievement between two points in time) are legally required to be included in teacher evaluations if data is available and relevant to the teacher and subject matter (Washington State Legislature, 1991a). Selection of appropriate outcome measures is the responsibility of the school district, and may include classroom-based, school-based, district-based, and state-based tools (Washington State Legislature, 1991a). More than one measure of student growth data must be used in scoring the student growth rubrics (Washington State Legislature, 2010).

The OSPI provides rubrics for summarising and evaluating student growth data. These rubrics provide guidance regarding the use of student growth measures in evaluating teachers and descriptors of performance reflective of each of the four levels of the Washington State rating system (Office of Superintendent of Public Instruction, 2014).

Direct Observation of Teaching

Direct observation of teaching is a legal requirement of teacher evaluation programs in Washington State. It is also one of the main components of the teacher evaluation frameworks endorsed by the OSPI.

The frequency of teacher observation varies based on a teacher's level of experience. Washington State legislature states that teachers who are new to the profession must be observed by their school district for a minimum of 30 minutes within the first ninety days of their employment (Washington State Legislature, 1991h). Provisional teachers are required to be observed at least three times in the performance of their daily teaching tasks each year,

Teacher Effectiveness Systems, Frameworks and Measures: A Review

for no less than a total of 90 minutes per year (Washington State Legislature, 2010; Washington State Teacher/Principal Evaluation Project, 2016c). Teachers who have progressed past provisional status are required to be observed twice per year. At least one of these observations must last for a minimum of 30 minutes, with a total observation time of at least 60 minutes per year (Washington State Legislature, 1991a). Finally, as previously mentioned, after four years of successful evaluations, teachers may be evaluated using a short-form evaluation (Washington State Legislature, 1991h, 2010). Results of observation must be promptly documented, with notes provided to the teacher within three days (Washington State Legislature, 1991a, 2010).

Teacher Self-Assessment

Self-assessment is an optional component of teacher evaluation in Washington State. While the TPEP framework and Washington State legislature do not explicitly recommend the use of teacher self-assessments in evaluating teachers, the eVal platform constructed and recommended by the OSPI contains several sections for goal setting and self-evaluation for teachers. Teachers may use the eVal online platform to look at evaluation criteria and to upload evidence that may support their performance as it relates to each evaluation criteria. Teachers can choose to share this information with their school principal (who is the acting evaluator) or to keep the information private (Education Service District 112, 2016; Washington State Teacher/Principal Evaluation project, 2016b).

Other

Student and peer perception data is noted as a potential source for informing teacher evaluation (Washington State Legislature, 1991a). No specific measures of student or peer perception are referenced in the reviewed documents.

Internal/External Evaluators

Washington State legislature requires evaluations to be conducted by the school principal or a designated evaluator (Washington State Legislature, 2010). School districts are required to offer initial and ongoing training for evaluators in evaluation procedures as well as the district's chosen teacher evaluation framework (Washington State Legislature, 1991b). Through ongoing training, it is expected that rater consistency is maximised, increasing the reliability of evaluations (Office of Superintendent of Public Instruction, 2016e). Failure of a school principal, or designated evaluator, to provide evaluation or supervision of certified classroom teachers may serve as sufficient cause for non-renewal of the evaluator's contract or discharge (Washington State Legislature, 2010).

Aggregation and Weighting of Results

The OSPI provides scoring bands for overall summative evaluations of teachers. Teachers are allocated a score of 1-4 on each of the comprehensive evaluation criteria outlined by the Washington State legislature. These scores are added to form a preliminary summative evaluation result. Following this, student growth impact ratings are calculated. Results from the comprehensive evaluation are combined with the student growth impact ratings to form an overall summative rating for each teacher (Walla Walla Valley Education Association, 2015; Washington State Legislature, 1991k).

USES OF EVALUATION

Evaluation results in Washington State are used for both summative and formative purposes. According to the WAC 392-191-045 legislature, evaluation results should be used (Washington State Legislature, 1991j):

"(1) To acknowledge, recognize, and encourage excellence in professional performance.

(2) To document the satisfactory performance by an employee of his/her assigned duties.

(3) To identify discrete areas according to the criteria included on the evaluation instrument in which the employee may need improvement.

89

Teacher Effectiveness Systems, Frameworks and Measures: A Review

(4) To document performance by an employee judged unsatisfactory based on the district evaluation criteria."

Evaluation results are used for summative purposes to determine whether teacher performance is satisfactory. Teachers whose evaluation results are deemed distinguished must be recognised and/or rewarded according to Washington State legislature (Washington State Legislature, 1991l). The precise nature of the recognition or rewards offered by school districts is not clear. Teachers whose evaluation results are deemed not satisfactory are notified in writing of the specific areas requiring improvement in addition to an improvement plan. A probationary period of sixty school days is then established. The evaluator must then meet with the employee at least twice monthly to supervise and make a written evaluation of the progress, if any, made by the employee. Failure to improve may result in re-assignment or paid leave. Continued failure to improve may also provide grounds for dismissal (Washington State Legislature, 2010).

Formative use of results is also required, with evaluations used to target appropriate professional development for each individual teacher (Washington State Legislature, 2010). Teachers are permitted to attach comments to their reports, and suggestions for improvement should be provided where appropriate (Washington State Legislature, 1991i).

CONSIDERATIONS FOR MINORITY GROUPS

The OSPI provides some guidance for evaluating teachers in: alternative learning environments (online learning, juvenile justice settings, and home school), pre-school and kindergarten, and special education settings. These guidelines generally encourage school districts to use the full instructional framework and evaluation rubric for most alternate school settings. In rare cases, the guidelines note that appropriate changes to evaluation frameworks may be made at the school district's discretion. No specific alternate measures or methodologies for evaluating teachers in alternate settings are provided by the OSPI guidelines (Office of Superintendent of Public Instruction, n.d.).

EXECUTIVE SUMMARY

- The Ministry of Education in the form of the Teacher Performance Appraisal System (TPA) provides a national framework for teacher evaluation. Implementation of this framework is the responsibility of each school's principal; however, this can be delegated to a nominated senior teacher.
- The Standards of Practice for the Teaching Profession set by the Ontario College of Teachers describes what is expected of teachers. These standards provide structure for teacher evaluation. The structure of teacher evaluations varies based on the teacher's level of experience. This includes the specific evaluation methods used and frequency of evaluations.
- All evaluations must feature a pre-observation meeting, a classroom observation, a post-observation meeting, and a summative report that includes a rating of the teacher's overall performance. Student outcomes and survey ratings may also be used at the evaluator's discretion. Self-assessment is suggested.
- The evaluation process is both summative and formative in nature. Feedback is provided to teachers in written and verbal form. No guidance is provided regarding follow up to evaluations post-feedback.

INTRODUCTION

School in Ontario is compulsory for children aged six to sixteen and is separated into (Ontario Ministry of Education, 2002):

- **Primary school** – Operates for eight years from Grade 1-8, which children attend from 6 to 13
- **Secondary school –** Operates for five years Grades 9-12; which children attend from 14-18

There are four types of publicly funded schools: English, English Catholic, French, and French Catholic, which are attended by approximately 1.4 million students (Ontario Ministry of Education, 2016). Private schools account for 5.1% of Ontario's schools (Clemens, Palacios, Loyer, & Fathers, 2014).

Governance

Education in Canada is the responsibility of each Province and there is no federal body or agency in charge of Education. The (Ontario) Ministry of Education is responsible for governing education in Ontario, while the Ontario College of Teachers regulates teacher registration. Private schools must also comply with regulations from the Ministry of Education; however, their teachers do not have to be members of the Ontario College of Teachers, although they commonly are.

STANDARDS AND THEORY

The Standards of Practice for the Teaching Profession are set by Ontario College of Teachers (2016) and describe the expectations of teachers. They were established in 1999, following a two- year review process which involved research, consultation, and analysis of teaching practices and ideals (Ontario College of Teachers, 2016b). This review process examined standards from countries across the world and identified salient themes that would also be applicable to Ontario. There are 16 competencies, organised into five domains, outlined in Table 13.

Table 13 - Ontario Standards of Practice for the Teaching Profession

Domain	Competency
Commitment to Students and Student Learning	Teachers demonstrate commitment to the well-being and development of all pupils.
	Teachers are dedicated in their efforts to teach and support pupil learning and achievement.
	Teachers treat all pupils equitably and with respect.
	Teachers provide an environment for learning that encourages pupils to be problem solvers, decision makers, lifelong learners, and contributing members of a changing society.
Professional Knowledge	Teachers know their subject matter, the Ontario curriculum, and education related legislation.
	Teachers know a variety of effective teaching and assessment practices.
	Teachers know a variety of effective classroom management strategies.
	Teachers know how pupils learn and the factors that influence pupil learning and achievement.
Teaching Practice	Teachers use their professional knowledge and understanding of pupils, curriculum, legislation, teaching practices, and classroom management strategies to promote the learning and achievement of their pupils.
	Teachers communicate effectively with pupils, parents, and colleagues.
	Teachers conduct ongoing assessment of pupils' progress, evaluate their achievement, and report results to pupils and their parents regularly.
	Teachers adapt and refine their teaching practices through continuous learning and reflection, using a variety of sources and resources.
	Teachers use appropriate technology in their teaching practices and related professional responsibilities.
Leadership and Community	Teachers collaborate with other teachers and school colleagues to create and sustain learning communities in their classrooms and in their schools.
	Teachers work with professionals, parents, and members of the community to enhance pupil learning, pupil achievement, and school programs.
Ongoing Professional Learning	Teachers engage in ongoing professional learning and apply it to improve their teaching practices.

In addition to the *Standards of Practice for the Teaching Profession* there are four *Ethical Standards for the Teaching Profession* outlined by the Ontario College of Teachers, including: (1) Care, (2) Respect, (3) Trust, and (4) Integrity (Ontario College of Teachers, 2016c).

Teacher Effectiveness Systems, Frameworks and Measures: A Review

TEACHER PREPARATION

Ontario teachers must complete a minimum three-year post-secondary degree from an approved institution. In September 2015 teacher education programs were increased to four semesters at Ontario's faculties of education and practicum was doubled from 40 days to 80 as a minimum requirement. Along with core elements there is a sharpened focus on mental health and wellbeing, diversity, and digital pedagogies (Ontario College of Teachers, 2015).

TEACHER REGISTRATION

Teachers require either a General or Transitional Certificate of Qualification and Registration from the Ontario College of Teachers to work in a publicly funded school (Ontario Ministry of Education, 2016). The Certificate is an annual license. Transitional Certificates apply to teachers who have not yet completed their teacher education program, after which they are eligible for a General Certificate (Ontario College of Teachers, 2016e). Transitional Certificates are valid for six years and may be extended.

Pre-requisites for certification for Ontario teachers are:

- Completed a (minimum) three year post-secondary degree from an approved institution
- Completed a four semester teacher education program
- Paid the registration and membership fees

Individuals who do not fully meet all these requirements may be eligible for conditional certification (Ontario College of Teachers, 2016d).

COURSE ACCREDITATION

The Ontario College of Teaching is responsible for accrediting teaching programs, and ensures that these courses give teachers the skills to meet the above standards (Ontario College of Teachers, 2016a).

EVALUATION FRAMEWORK

Ontario's Teacher Performance Appraisal System (TPA) is set by the Ministry of Education, and applies to all teachers working in publicly funded schools. Principals are responsible for conducting evaluations of all teachers working at their school; however, this may be delegated to another senior staff member when appropriate (Ontario Ministry of Education, 2010). The evaluation process varies based on the teacher's career level, with beginner teachers having to successfully pass the New Teacher Induction Program (NTIP) before they can be considered "experienced" teachers. The NTIP involves teachers being appraised on six of the criteria twice each year, and they must receive satisfactory ratings before they are able to progress. Experienced teachers are evaluated on all of the criteria once every five years.

METHODS AND METHODOLOGIES

The manner in which teachers are appraised varies based on teacher experience; despite this, all appraisals must include:

- A pre-observation meeting
- A classroom observation
- A post-observation meeting
- A summative report that includes a rating of the teachers overall performance (Ontario Ministry of Education, 2010)

The evaluator may then select other methods they feel are appropriate; however, little guidance is given as to what constitutes an appropriate indicator.

Student Outcomes

Student outcomes are listed as a potential source of evidence in the Ontario TPA manual (Ontario Ministry of Education, 2010), but the precise use of results is not defined.

Direct Observation of Teaching

The Ontario TPA manual states that prior to directly observing teaching, the principal must have a pre-observation meeting with the teacher, where clear expectations and procedures for the observation are established and explained (Ontario Ministry of Education, 2010). While it is mandatory that principals observe teachers in their instructional setting and evaluate them in relation to the teaching standards, there is little information regarding how they determine if teachers are meeting the standards.

Survey Ratings

Input from students and parents is suggested and teachers are encouraged to reflect upon the feedback they receive (Ontario Ministry of Education, 2010). However, it is not mandatory and teachers are themselves responsible for collecting this information.

Teacher Self-Assessment

Self-evaluation is suggested (Ontario Ministry of Education, 2010), but not mandatory.

Internal/External Evaluators

Teachers are evaluated internally, generally by their principal (Ontario Ministry of Education, 2010). In some circumstances an evaluator from outside the school may be used, but this is uncommon.

Aggregation and Weighting of Results

Not referenced in reviewed documents.

USES OF EVALUATION

The evaluation process is both summative and formative in nature. All teachers are given written and verbal feedback following the evaluation. This feedback should include recommendations for their ongoing development. Teachers who receive two consecutive unsatisfactory ratings will be put on review status, and will be evaluated again by the school board. Termination may result if the teachers fail the additional evaluation, although this is not guaranteed, as each school board makes decisions regarding termination on a 'case by case' basis (Ontario Ministry of Education, 2010).

CONSIDERATIONS FOR MINORITY GROUPS

Not referenced in reviewed documents.

EXECUTIVE SUMMARY

- There is no systematic or consistent framework for teacher evaluation in British Columbia. Evaluations are a common feature of collective agreements between teacher unions and schools; however, there is a large degree of variation in the requirements and outcomes following evaluations between each school district. Generally, principals conduct evaluations, with a high degree of control over the tools and methods.
- The Standards for Teaching are set out by the British Columbia Teaching Council. The role of these standards in teacher evaluation is not clear.
- Direct observation and self-assessment are mentioned in some teacher union-school agreements as tools for evaluation; however clear guidelines are not given and the choice of evaluation tools is left up to each agreement.
- Evaluations appear to be formative in nature, with evaluators providing written and verbal feedback. Some evaluation results may be seen as summative, with teachers given a rating of satisfactory or unsatisfactory. There are no explicit guidelines regarding the use of evaluation results following feedback to the teacher.

INTRODUCTION

There are approximately 1600 public schools and 350 private schools in British Columbia (British Columbia Teacher Federation, 2012), organised across school sectors, as:

- **Elementary school**, grades 1 to 6
- **Junior High,** grades 7 to 9
- **Senior High** school, grades 10 to 12

Governance

The British Columbia Ministry of Education (BCMOE) is the provincial department responsible for education. The Ministry established the British Columbia Teachers Council, which is responsible for teaching standards, appraisal, and preparation (British Columbia Teachers Council, 2012).

STANDARDS AND THEORY

Standards for Teaching are set by The British Columbia Teachers Council (2012), and state the expected behaviour and capabilities of teachers:

- Educators value and care for all students and act in their best interest
- Educators are role models who act ethically and honestly
- Educators understand and apply knowledge of student growth and development
- Educators value the involvement and support of parents, guardians, families and communities in schools
- Educators implement effective practices in areas of classroom management, planning instruction, assessment, evaluation and reporting
- Educators have a broad knowledge base and understand the subject areas that they teach
- Educators engage in career-long learning
- Educators contribute to the profession

TEACHER PREPARATION

To gain entry to a teacher training program in British Columbia (BC) candidates must have completed the BC Grade 12 diploma or equivalent (BCMOE, 2016b). Prospective teachers are required by the BCMOE to complete four years

95

Teacher Effectiveness Systems, Frameworks and Measures: A Review

of post-secondary studies and a degree or equivalent (BCMOE, 2016b) with approved coursework and teacher education components. The BC teacher education program must be approved by the British Columbia Teachers' Council to lead to certification to teach in the BC elementary and/or secondary public school system.

Academic studies must include coursework in English or French literature and composition. Elementary school teachers must complete general coursework in BC elementary curriculum areas, including Canadian studies, mathematics and select fields of science (BCMOE, 2016b). Secondary teachers must include coursework selected from a range of curriculum subject areas (BCMOE, 2016b). The teacher education component consists of education courses and classroom teaching experiences equivalent to 1.5 – 2 years of full-time study (BCMOE, 2016b). These programs must include prescribed studies in human development, educational foundations, curriculum, pedagogy, assessment and special needs (BCMOE, 2016b).

British Columbian teachers must have advanced proficiency in English, or French for those planning to teach with the French education authority of the BC (le Conseil scolaire francophone de la Colombie-Britannique, 2016).

There is no probation period or provisional status, and once ITE has been completed, teachers are fully certified; however, there are a range of different qualification certificates available for those who have not completed the coursework or meet the requirements for a Professional Certificate, for example, the Conditional Certificate, which lasts for 60 months, or a Basic Certificate, which does not expire (BCMOE, 2016d).

TEACHER REGISTRATION

All educators working in the BC kindergarten to grade 12 public education sectors must hold a certificate of qualification from the BC Ministry of Education Teacher Regulation Branch (BCMOE, 2016a). Applicants are evaluated according to:

- Academic record, including teacher education training and subject area studies
- Teaching experience
- Fitness to teach (good character, final practicum report, and criminal check results) (BCMOE, 2016b)

The Teacher Regulation Branch issues certificates under several options (BCMOE, 2016d).

- *Professional Certificate* - does not expire and is only available to those who have met all of the teaching certification requirements. Specifically, successful applicants will have completed a teaching education program, a post-secondary degree and have good moral character (e.g., pass criminal record checks etc.).
- *Basic Certificate* - does not expire, and available to those who have not met all of the requirements, but who have full certificate in another Canadian Provence.
- *Trades Certificate* - does not expire, but restricts the teacher to a particular trade
- *Conditional Certificate -* only valid for a maximum of 60 months, available to those who have partially meet the requirements (and do not have a fully valid certificate from another state), they must participate in additional learning whilst teaching.
- *First Language Teacher Certificate* - issued to proficient First Language speakers who are ineligible for professional certificate

Teachers working in independent schools must have an independent school certificate, issued by the Teacher Regulation Branch (BCMOE, 2016d). Letters of permission are also issued to suitable persons who may be needed by individual schools for a period not exceeding one school year.

COURSE ACCREDITATION

Accreditation of teacher preparation programs is the responsibility of the British Columbia Teaching Council (2012). There are currently nine teacher education programs in British Columbia (BCMOE, 2016c). Programs are reviewed

Teacher Effectiveness Systems, Frameworks and Measures: A Review

regularly for assurance that they continue to meet required standards and produce graduates that meet the standards for educators in BC.

EVALUATION FRAMEWORK

There is no systematic or consistent framework for teacher evaluation in British Columbia. Evaluations are commonly part of a collective agreement, negotiated between the teachers union and school board. While there is variation between the school districts, these agreements tend to specify that school principals evaluate teachers every three years. There is a high degree of variation regarding the ways in which teachers are evaluated. For example, the collective agreement for Abbotsford specified that teacher evaluations should be based on between three and six observations (The board trustees of school district no. 34, 2009), whereas the agreement for Spanish specified a maximum of three observations (British Columbia Public Schools Employers' Association and British Columbia Teachers' Federation, 2011). Furthermore, most agreements give principals a high degree of autonomy in determining the criteria, tools and methods for the evaluation, stating only that these should be selected in consultation with the teacher at the beginning of the appraisal period. All the collective agreements examined were summative in their evaluations, with teachers categorised as satisfactory or unsatisfactory.

METHODS AND METHODOLOGIES

Student Outcomes

Student outcomes are not mentioned in any of the collective agreements examined.

Direct Observation of Teaching

Direct observation of teaching is mentioned in each of the collective agreements examined. While there are differences in the number of observations requested, no school districts give clear guidelines as to how observations should be conducted, with principals given a high degree of freedom in determining if teachers satisfied the standards.

Survey Ratings

Not referenced in reviewed documents.

Teacher Self-Assessment

Self-assessment is mentioned in several of the collective agreements; however, there is little information as to how this should be carried out.

Internal/External Evaluators

Evaluation is internal, with the principal taking responsibility.

Aggregation and Weighting of Results

None of the collective agreements outline how the results would be weighted; the teachers' rating as satisfactory or unsatisfactory is determined by the principal's discretion.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

USES OF EVALUATION

Evaluations are formative in nature. All teachers should receive both written and verbal feedback following the evaluation. However, there are no explicit guidelines as to how this should be carried out and how written reports should be structured. While evaluations form some aspects of teachers' career progressions, they are only one facet, with decisions regarding promotions and terminations made by the school boards on a 'case by case' basis.

CONSIDERATIONS FOR MINORITY GROUPS

Minority groups are mentioned in the standards for teaching, which state that teachers must understand the curriculum in a Canadian, Aboriginal and global context (British Columbia Teachers Council, 2012).

EXECUTIVE SUMMARY

- There is no system for regular and reoccurring teacher evaluation in Alberta. When teacher evaluation does take place, it is typically based on requests from the teacher or school leaders on the basis of employment decisions, assessing teacher growth, or when a teacher is suspected of underperforming according to the standards for teaching.
- The school's principal using general guidelines from the Ministry of Education carries out evaluations.
- Teaching standards are set by the Ministry of Education and differ depending on teacher skill level. This 'Teaching Quality Standard' strongly guides the purpose and focus of teacher evaluations when they occur.
- Recommended tools include student outcomes and direct observation.
- Evaluations appear to be formative in nature, with evaluators providing written and verbal feedback. There are no explicit guidelines regarding the use of evaluation results following feedback to the teacher.

INTRODUCTION

Alberta has a total of 2,388 schools, attended by 69,0844 students (Alberta Education, 2016a). School is compulsory for children aged 6-16, and organised into:

- **Elementary School –** Operates from Grades 1 to 6
- **Junior High** - Operates from Grades 7 to 9
- **Senior High** – Operates from Grades 10 to 12

Most students attend a public school, with Alberta having 1,498 public schools compared to 391 separate and 182 private schools. (Alberta Education, 2016a).

Governance

Education policy and regulation is the responsibility of provincial governments in Canada, and in Alberta is governed by the Ministry of Education. The Ministry is responsible for curriculum development, teaching and learning standards, and evaluation (The Ministry of Education, 2016b).

STANDARDS AND THEORY

The Ministry of Education has regulated 'Teaching quality standards' which outline the skills, attributes, and knowledge expected of teachers (The Ministry of Education, 2016e). The standards differ depending on the teacher's career level, with one set of standards for those with interim certification, and more comprehensive standards for permanently certified teachers. There are 19 standards in total, and some examples of the responsibilities and expectations of teachers are (The Ministry of Education, 2016e):

- A belief that all students can learn, albeit differently
- Contributing, independently and collegially, to the quality of the school community
- The importance of guiding your own actions with a personal, overall vision of the purpose of teaching
- Creating and maintaining environments that are conducive to student learning

The importance of school and classroom context is emphasised throughout the standards and other ministry documents and legislation, and teachers are expected to understand students' varying needs and respond appropriately.

TEACHER PREPARATION

ITE in Canada is based in university faculties of education governed by the Ministry of Education. A Bachelor's Degree in Education (BEd) and a provincial teacher's certificate are required to teach in Canada.

 In Alberta, to be eligible for certification a prospective teacher must gain either:

- Bachelor of Education degree from a faculty of education from a university in Alberta
- Another approved degree, from a university in Alberta or from an approved institution, plus an approved professional teacher education program in a faculty of education of a university in Alberta or at any other approved institution (The Alberta Teachers' Association, 2016)

COURSE ACCREDITATION

The Ministry of Education must accredit teacher-training courses.

TEACHER REGISTRATION

Teachers must be certified by the Ministry of Education through the Office of the Registrar to work in Alberta Schools, with two levels of certification:

1. **Interim Professional Certificate (IPC),** teachers are eligible for IPC once they have completed a recognised teaching degree and gained appropriate work experience. IPC is valid for three years, and all teachers must spend at least two years on this stage before they can progress
2. **The Permanent Professional Certificate (PPC)** is the second level of certification, and teachers must have passed two evaluations based on the teaching qualities standards to reach this level of certification (Alberta Education, 2016b)

EVALUATION FRAMEWORK

There is no system for regular and reoccurring teacher evaluation in Alberta, and evaluations typically occur in the following circumstances (The Ministry of Education, 2016d):

- Upon the written request of a teacher
- For the purpose of gathering information related to employment decisions, such as when a teacher is progressing from IPC to PPC
- For assessing the growth of the teacher in specific areas
- When the principal has reason to believe that the teacher may not meet the standard

These formal evaluations are carried out internally, by the school principal, with general guidelines for evaluation and a high degree of freedom in determining if a teacher is meeting the standards (The Ministry of Education, 2016d). Furthermore, while the Ministry of Education recommends that informal evaluation should take place regularly, there are no mechanisms in place to ensure that teachers are receiving regular reviews and feedback.

There is also a process for whole school evaluations, which encompass some aspects of teacher evaluation. The Accountability Pillar is a mandatory school appraisal system, which aims to improve student outcomes by providing schools with a range of feedback regarding how their students are performing (The Ministry of Education, 2016c). This process involves some methods and tools commonly used for teacher evaluation, such as student's improvement results and parental surveys. Furthermore, these evaluations are used to guide the school's growth plan, which includes their strategies for teacher professional development (The Ministry of Education, 2016c).

Teacher Effectiveness Systems, Frameworks and Measures: A Review

METHODS AND METHODOLOGIES

There is little information as to the specific methods and tools used by principals to determine if a teacher is meeting the quality standards.

Student Outcomes

The outcomes of Student Learning Assessments are made available to teachers and schools, who are encouraged to use this information to reflect on their teaching and classroom management, but this is not enforced. Student outcomes are also a part of the Accountability Pillar and principals may use this information to guide teacher professional development, although this is also not obligatory (The Ministry of Education, 2016a).

Direct Observation of Teaching

Direct observation of teaching is not mandatory, but is a commonly used strategy in teacher evaluation. Furthermore, a number of school districts have listed direct observation as part of their districts system for teacher evaluation.

Survey Ratings

Surveys are not used in teacher evaluation, but are used in the Accountability Pillar as part of the school evaluation, where teachers, students, and their parents are surveyed on their perceptions of the school's effectiveness (The Ministry of Education, 2016a).

Teacher Self-Assessment

Not referenced in reviewed documents.

Internal/External Evaluators

Evaluation is internal, with the principal taking responsibility.

Aggregation and Weighting of Results

Not referenced in reviewed documents.

USES OF EVALUATION

The school-based teacher evaluation process is formative in nature. Following a formal evaluation, teachers should be given an evaluation report in writing, which outlines how their performance relates to the standards. In instances where the principal felt that the teacher did not meet the standards, they must provide the teacher with a 'notice of remediation'. This explains why the teacher was considered unsatisfactory in relation to one or more of the standards, and outlines the ways in which they should improve (e.g., professional development).

CONSIDERATIONS FOR MINORITY GROUPS

Not referenced in reviewed documents.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

METHODS FOR TEACHER EVALUATION

Systems for teacher evaluation vary from sophisticated national systems to localised, informal approaches. Effective systems are differentiated, featuring multiple methods and tools that are implemented by well-trained evaluators (OECD, 2013c). Systemic approaches are typically framed by national or statutory professional standards for teachers that articulate the professional knowledge, practice and engagement expected of teachers relative to career stage. A wide range of methods and tools are employed for teacher evaluation. Instruments in most frequent use are classroom observation, self-appraisal, teaching performance portfolios, and performance interviews (OECD, 2013).

Whilst teacher evaluation remains a highly contested area there is general agreement that regular appraisal of teaching performance promotes quality teaching and continuous professional learning, leading to improved student achievement (OECD, 2013). There is consensus that systems for teacher evaluation are in need of an overhaul (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012). Reports highlighting the inability of current systems to accurately measure teacher quality or to support development of a highly skilled teacher workforce (Bill & Melinda Gates Foundation, 2012; Darling-Hammond et al., 2012; Toch & Rothman, 2008; U.S. Department of Education, 2011; Weisberg et al., 2009) have sparked recent efforts to review methods for measuring and improving teacher expertise.

Reporting from the Stanford Centre for Opportunity Policy in Education (SCOPE), Darling-Hammond et al. (2012) highlights the important distinction between *teacher* quality (personal skills, attributes and dispositions) and *teaching* quality (strong instruction that supports student learning). Darling-Hammond notes influences such as the instructional context, curriculum and assessment systems, class size, facilities, and materials, and calls upon policymakers to consider the teaching and learning environment alongside the capacity of individual teachers. Popham (2013) is in agreement and further posits that teacher evaluators should determine the evaluative weighting for each piece of evidence or demonstration of practice, then adjust according to the instructional setting.

There have been several recent reviews of the literature on teacher evaluation. *Evaluating Teacher Effectiveness* (Darling-Hammond, 2010) explored measures of teacher effectiveness as predictors of value-added student achievement. The report detailed progress towards the development of a continuum of reliable and valid teacher performance measures, from initial teacher education through to recognition of advanced expertise. Standards-based approaches to assessment, such as the Performance Assessment of California Teachers (PACT, now edTPA) for pre-service teachers and National Board Certification for experienced teachers, were outlined in the report. Multiple measures of effective teaching and demonstrated evidence (such as video recording of practice, classroom observation, performance portfolios with commentary, and evidence of student achievement) were detailed along with recent research findings.

Similarly, the *Measures of Effective Teaching* (MET) project (2009) investigated a selection of key measures of teacher effectiveness. The study provided considerable empirical evidence related to teacher effectiveness. It produced and trialled various evaluative measures (e.g. classroom observation, video recording of practice, student surveys, student rating, etc.), as well as value-added modelling procedures for identifying effective teachers and predicting the contribution an individual teacher has on student gains (Kane, McCaffrey, Miller, & Staiger, 2013).

TEACHER EFFECTIVENESS MEASURES

The evaluation team has conducted an extensive review of currently available measures of teacher effectiveness. In this report, a description and overview of each method is provided, along with an analysis of reliability and validity data. Methods examined in the review include classroom observation, teaching performance portfolios, teacher interviews, performance and development interviews, peer ratings and student ratings.

CLASSROOM OBSERVATIONS

OVERVIEW

Classroom observations are one of the most commonly used methods of teacher evaluation. Observations typically involve an evaluator (usually a senior colleague or school leader) observing a lesson taken by the teacher in their classroom. Classroom observations can vary significantly in terms of purpose, scheduling, structure, timeframe, protocols, and measures of effective teaching. The formality of observations can range from a brief and impromptu visit where the evaluator drops by a class, to a formally scheduled visit from an external evaluator who conducts a criterion-based assessment of the teacher's practice. Observation of classroom practice may involve the teacher being filmed, with the video footage viewed and assessed by one or more evaluators. Documentation and tools associated with classroom observation are many and varied, ranging from generic or locally modified versions of standards-based constructs to research-based, commercially produced frameworks. Classroom observations exist both as a singular method of teacher evaluation and a component of multi-faceted evaluation systems.

Classroom observations are commonly employed to assess pre-service teachers and new graduates as part of the teacher training and registration process. For pre-service teachers, observations occur throughout each practicum to provide teacher candidates with constructive and actionable feedback for improvement of their practice. In aggregate, these observations inform the summative assessment, usually finalised at the end of a placement block. It is also noted that many of these focus on the behaviours of teachers (do they use feedback, open questions, use evidence to change teaching) and rarely have items relating to their impact on students (Haigh & Ell, 2014).

Classroom observations are also a typical requirement for graduate teachers seeking full licensure both locally and internationally; in Australia, graduate teachers must successfully pass a series of collegiate observations to be eligible for full registration with the statutory body. Once fully registered, the observation of teachers may be less regulated. While it is suggested that Australian teachers should undergo an annual formal evaluation as part of their performance and development review process, this recommendation is not formally enforced. In some educational contexts regular and systematic evaluations are mandated such as South Korea, where teachers undergo yearly evaluations that must include a classroom observation component. In other systems, such as in Germany, classroom observations only occur in response to a formal complaint against a teacher.

VALIDITY & RELIABILITY

Research investigating the validity and reliability of classroom observations for evaluating teacher quality has produced varying results. Major concerns regarding the validity and reliability of classroom observations include: the variable quality of observational instruments, a paucity of evidence regarding the interpretation of observations, focus on variables other than the impact of the teaching, an apparent lack of systematic collection methods which may affect the representativeness of observations, the number of observations necessary to lead to dependable ratings, and the importance of using observation in tandem with other measures of effectiveness. Further, while there are claims about the formative potential of these instruments, there is little evidence of their impact on improved teaching quality or enhanced student achievement. A summary of validity and reliability studies is presented below.

A particular challenge for classroom observations as a tool for evaluation is the degree to which observations correlate with other measures of teacher effectiveness. Strong, Gargani, and Hacifazlioğlu (2011) conducted three experiments where trained evaluators viewed a series of video footage of teachers' classroom practice and assessed their performance. A major finding was that evaluators repeatedly failed to identify those teachers who produced the greatest learning gains, as indicated by student performance on standardised tests at the beginning and end of the year. Evaluators' assessment scores were consistent with the Value Added Modelling (VAM) scores only 47% of the time. Similarly, Jacob and Lefgren (2008) found that while principals are able to identify teachers at extreme ends of effectiveness scales (i.e., those that fall in the top or bottom 10%) they are not adept at assessing those who

fall somewhere in the middle. Collectively, this research suggests that classroom observation may not provide an accurate assessment of teaching performance when used as a standalone measure.

Conversely, Clark, Martorell, and Rockoff (2009) determined that there was a positive correlation between principals' beliefs about a teacher and their students' VAM data. They also found that principals took student achievement data into consideration when making evaluations, and did not rely solely on their already established impressions of the teacher. A number of authors have argued that observation is a necessary aspect of teacher appraisal, as it is able to capture aspects of teaching that other methods of evaluation cannot (Bell et al., 2012; Darling-Hammond et al., 2012; Hill & Grossman, 2013). In this sense, teacher quality is considered to be a multifaceted construct, with classroom observation the only way to assess particular aspects of teacher effectiveness. The Bill and Melinda Gates Foundation conducted a large study of the reliability of classroom observations in 2011/ 2012. An early report on the measures of effective teaching compared five different instruments for scoring classroom instruction using trained observers. The study demonstrated the need for multiple measures in order to more reliably predict teacher quality, as shown in Figure 2.

Multiple Measures Mean More Accurate Evaluations

Correlation w/ teacher's long-term success (math)

0.24 — Classroom observations only (CLASS)
0.37 — Student surveys only
0.69 — Value-added only
0.72 — All three measures combined

Source: "Gathering Feedback for Teaching," page 51. Other observation rubrics had similar correlations.
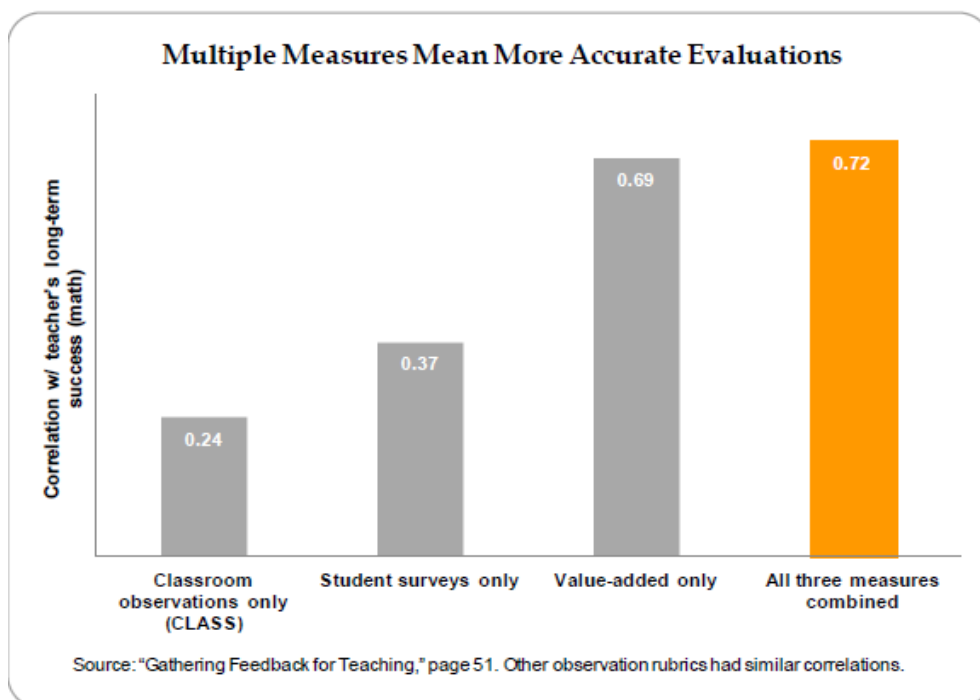
Figure 2 The validity of multiple measures in teacher evaluation.

The report found that scores on each of the five instruments were highly correlated with one another and were positively associated with student gains in maths or English. However, reliability was not consistent; teacher scores varied considerably from lesson to lesson, and for any given lesson, scores varied from observer to observer. A further study evaluated the accuracy and reliability of school personnel in performing classroom observations that led to ten key findings:

1. Observers rarely used the top or bottom categories of unsatisfactory and advanced on a point observation instrument
2. Compared to ratings from peers, ratings from administrators, differentiated teachers more. The standard deviation in underlying teacher scores was 50% higher when scored by administrators than when scored by peers

Teacher Effectiveness Systems, Frameworks and Measures: A Review

3. Administrators rated their own teachers 0.1 points higher than administrators from other schools and 0.2 points higher than peers.
4. Although administrators scored their own teachers higher, their rankings were similar to the rankings produced by others outside their schools.
5. Allowing teachers to choose their own videos generated higher average scores. However, the relative ranking of teachers was preserved whether the videos were chosen or not.
6. When an observer formed a positive or negative impression of the teacher in the first several videos, that impression tended to linger.
7. The reliability of a 15-minute observation was 60% as large as that of a full lesson of observation. There are a number of different ways to ensure reliability of .65 or above. Having more than one observer really does matter.
8. Increasing the number of observations per observer does not reduce the source of error - the use of multiple observers is the best way to reduce error.
9. The element of surprise observation heightened anxiety and reinforced the impression that teacher evaluation is about accountability and not about improvement.
10. Classroom observation instruments are not discerning large *absolute* differences in practice.

Research examining the reliability of classroom observation as an assessment method for teacher effectiveness has also produced mixed results. A number of studies have called into question the reliability of classroom observation, indicating that appraisals may be influenced by extraneous variables. Harris and Sass (2011) found that ratings appear to fluctuate depending on the lesson and course content, with observations in some subject areas appearing to be more reliable than in others. Specifically, classroom observations of mathematics teachers are more closely linked to VAM scores than those of English teachers. Jacob and Lefgren (2008) determined a correlation between class composition and evaluation results, with evaluators tending to rate teachers with more capable students higher than their counterparts. Furthermore, research demonstrated that observational scores of video recorded lessons varied depending on whether the evaluator watched the lesson in 10, 20 or 40-minute segments. The researchers argued that evaluators are vulnerable to primacy and recency effects, whereby information is recalled primarily from the beginning or end of the observation period. It was further suggested that two 20-minute observations produced the most reliable scores. There is strong evidence that having an additional evaluator significantly increases the reliability of the evaluation, and is more effective than multiple observations by the same evaluator (Ho & Kane, 2013; Whitehurst, Chingos, & Lindquist, 2014a).

Regarding the test-retest reliability of classroom observations, Whitehurst, Chingos and Lundquist (2014a) argued that observation scores appear to be more stable over time than Value-added (VAM) scores. They identified a moderately strong positive correlation ($r = 0.65$) between the observation scores given to teachers from one year to the next. This level of consistency was significantly greater than the VAM correlation of 0.38, leading the researchers to conclude that classroom observation is a reliable assessment method. However, Jacob and Lefgren (2008) are critical, maintaining the view that this may be a consequence of biases maintained over time as opposed to the accuracy of evaluations.

The potential for bias is one of the biggest criticisms of classroom observation as a method for teacher evaluation, particularly in instances where the principal acts as the evaluator. The MET study (Kane & Staiger, 2012) provides some evidence of observer bias, by showing that principals rated teachers from their own schools higher than teachers from other schools. However, this bias was considered insignificant, as there was a correlation of .87 between the internal and external evaluation scores. The researchers argued that this type of bias was mitigated by the provision of rigorous training for observers and a standardised assessment rubric. Observer biases were also investigated in a major US report entitled '*The Widget Effect*' (Weisberg et al., 2009). The study examined data from over 15,000 teacher evaluations, finding that teachers were rated as at least satisfactory in 98% of observations conducted by their principals (Weisberg et al., 2009). The report highlights the improbability that this proportion of teachers evaluated was satisfactory, and concludes that classroom observations, as a standalone measure, are not a reliable method for teacher evaluation. Further to this, external observers may also be biased by previous ratings, and are more likely to rate a teacher highly if it is consistent with their previous scores, irrespective of their

105

performance (Hill, Charalambous, & Kraft, 2012). More accurate and reliable ratings may be produced using multiple observers, reducing bias (Whitehorse et al., 2014). In high stakes evaluations multiple observers should always be used (Murphy, 2013).

There is great variation between the expertise and qualification of those observing teachers in their classroom practice for assessment purposes. Senior supervisors within the school, generally a person from the principal class or leadership level, most commonly conduct classroom observations. While assessor training is recommended to encourage reliable and objective appraisals (Mathers & Oliva, 2008), an examination of teacher evaluation in seven U.S states found that only 8% of evaluation frameworks required any form of assessor training (Brandt, Mathers, Oliva, Brown-Sims, & Hess, 2007; Brandt et al., 2007). In some cases, an external evaluator may conduct the observation, but this is generally a second phase observation reserved for specific circumstances such as an appeal or complaint. The New Zealand Education Review Office provides a valuable example of an inspectorate system that has its foundation in evaluation, utilising a search and assist model.

There is inconsistency in the findings of research focused on inter-rater reliability. While several studies have found a strong correlation amongst the scores of different observers (Marzano & Toth, 2013; Pianta & Hamre, 2009; Strong et al., 2011), this is generally reserved for those tools with systematic scoring guidelines, and frameworks that give evaluators more freedom tend to be less reliable (Whitehurst et al., 2014a). The MET Project, involved the video recording and evaluation of over 20,000 lessons with findings that "the same teacher was often rated differently depending on who did the observation" (Ho & Kane, 2013). The study also showed that scores varied significantly depending on the lesson being observed. The researchers recommend evaluation systems that combine classroom observation with other methods of appraisal to increase the validity and reliability of the evaluation.

Importantly, classroom observations are the most commonly used method for teacher evaluation (Goe, Bell, & Little, 2008b), indicating that it is sustainable over time and across contexts. Moreover, there is evidence to suggest that teachers and teacher unions are more supportive of observational methods than those that rely solely on student outcomes, and this may make teachers less resistant to observations and more responsive to evaluation outcomes. For example, 84.4% of teachers in one study reported they believed classroom observation to be important (Lasagabaster & Sierra, 2011). Finally, as observations are generally conducted internally, they are a more accessible form of teacher appraisal in comparison to value-added models that require rigorous data collection and analysis that is more time consuming and expensive.

While many teacher evaluation schemes, and teacher evaluation policies, highlight the importance of classroom observation there is often inconsistency in the observation methodology, including frequency and duration of observations, the use of observational rubrics, and other critical features. An examination of teacher evaluation in seven U.S. states found that policies rarely gave specific instructions as to how evaluations should occur (Brandt et al., 2007). The inconsistency in how observations are used decreases both its validity and reliability as an observation method. Furthermore, observers may have different levels of expertise and ideas of teacher effectiveness, which may impact the way they evaluate and score the teachers. In addition to classroom observations being biased by the evaluator's beliefs, they may be impacted by the evaluator's relationship to the teacher or teaching candidate. While these obstacles may be minimised by using a standardised and reliable classroom observation tool (Murphy, 2014; Whitehurst, Chingos, & Lindquist, 2014b), the tools themselves vary in their reliability, validity and relevance to the educational contexts. Similarly, ensuring that observers are well trained may increase reliability.

THE VISIBLE CLASSROOM

OVERVIEW

The Visible Classroom aims to improve the quality of learning and teaching by providing teachers with opportunities to reflect on their teaching and develop their classroom practice using real-time, credible evidence. The project represents a partnership between Melbourne Graduate School of Education at the University of Melbourne and

Access Innovation Media (Ai-Media). This partnership developed out of a previous investigation of the real-time captioning in classrooms with deaf/hard of hearing students. The Visible Classroom is designed to foster a mindset in participating teachers whereby they a) engage regularly in critical reflection on their teaching goals and values; b) examine how these connect with their teaching practices; and c) feel empowered to adapt their practice in a formative manner based on sound evidence.

Visible Classroom draws on the principles of contemporary pedagogy in particular Visible Learning (Hattie, 2009). In Visible Learning, Hattie notes that teaching and learning is too often hidden, characterised by high levels of teacher talk, but little reflection on the impact of teaching on students. The focus of the Visible Classroom program is the notion of embedded evaluation within the technology, in order to encourage teachers to critically assess what they have done and what their students have learned.

Visible Classroom allows the teacher to audio-record the classroom lessons directly from a smart phone or tablet and then upload the lesson for analysis. The teacher receives a full transcript of their lesson and important metrics about their teaching coded against an evidence-based rubric of 16 effective teaching practices, that has an overlay of Biggs and Collis (1982) SOLO taxonomy. For example, it analyses the time teachers allow for students to participate in their own learning, talk time, and number of questions asked. Visible Classroom allows the teacher to track progress over time and monitor their impact. Periodically (usually after 5 lessons), more detailed analyses are presented to teachers, including whether they were teaching at a surface (i.e., instruction in concepts or ideas) or deep (i.e., exploring connections and relationships between ideas and extending these concepts) level, along with practical suggestions for how they might develop their practice.

RELIABILITY AND VALIDITY

Clinton et al. (2014) demonstrated that The Visible Classroom intervention had an impact on teachers' instructional practice, both in terms of measurable change as evident in lesson transcripts over the intervention period, and based on teacher self-assessments. Challenges were encountered, however, in maintaining this impact on teachers over an extended period of time. A shorter, more intensive version of the intervention appeared to be more effective in engaging teachers in self-reflection, and driving change in instructional practice. This study demonstrates that instructional practice can be enhanced through the provision of feedback. Across this and several studies, teachers in general rated themselves as having improved in the 16 teaching practices as a result of utilising the Visible Classroom over a period of 5 lessons. Further the frequency ratings on the 16 items illustrated moderate to high change with the effect sizes ranging *d*=0.00 - 0.94. This tool holds promise as an alternative to walk-through and video-based classroom observation systems, and may promote improvements in teachers' critical reflection and instructional practice.

EXAMPLE FRAMEWORKS FOR CLASSROOM OBSERVATIONS

Numerous frameworks and tools have been developed locally and internationally that support the measurement of dimensions of teaching evident during classroom practice. Most notably, the Marzano Teacher Evaluation Model (Marzano & Toth, 2013), the Framework for Teaching evaluation instrument (Danielson, 2011), the Classroom Assessment Scoring System (CLASS) (Pianta, La Paro, & Hamre, 2008), and the NSW Quality Teaching (QT) Model (Bowe & Gore, 2011) are acknowledged as research-based, reliable and valid measures of effective teaching. These tools will be examined and discussed in following sections.

MARZANO TEACHER EVALUATION MODEL

The Marzano Teacher Evaluation Model identifies sixty elements of teacher effectiveness, organised across four domains. The tool is primarily designed to examine, adjust and improve the instructional practices of teachers. Domain 1 focuses specifically on *classroom strategies and behaviours* that impact on student achievement. The forty-one components comprising Domain 1 are categorised as: (1) *routine events*, (2) *content*, and (3) *on the spot strategies*. Domain 2 categorises the elements required for *planning and preparing* which are considered to link

107

Teacher Effectiveness Systems, Frameworks and Measures: A Review

directly to classroom strategies and behaviours. The eight elements of Domain 2 are categorised as: (1) *planning and preparing for lessons and units*, (2) *planning and preparing for use of materials and technology*, and (3) *planning and preparing for special needs of students*. Domain 3 targets *reflecting on teaching* and the metacognitive steps involved in teacher development. The five elements in Domain 3 are categorised as: (1) *evaluating personal performance* and (2) *developing and implementing a professional growth plan*. Finally, Domain 4 focuses on *collegiality and professionalism*. The six elements in Domain 4 are categorised as: (1) *promoting a positive environment*, (2) *promoting exchange of ideas and strategies*, and (3) *promoting district and school development*. While comprehensive, these various elements and domains represent highly desired attributes/activities however, like many observation schedules, there is a supposition that the presence of these desired attributes/activities leads to higher impacts on students – a case of confusing correlation as causation.

OVERVIEW

The Marzano Teacher Evaluation Model is a popular and commonly used teacher appraisal tool across the United States of America, Canada, and Australia, as well as countries within Europe, Asia, and South America. This model was chiefly designed for formative uses, and as a part of supervision-based strategy to improve teachers' instructional skill. It is, however, often used for the collection of evaluative evidence regarding summative judgments about a teacher's quality. The evaluation model requires frequent observations across different lessons and groups of students, in combination with the provision of ample focused feedback to build teacher expertise over time. The frequency of these observations is solely dependent on school or district policies, however multiple observations are recommended to increase the accuracy of the rating.

During a classroom observation session, the Marzano-trained examiner identifies the presence of the 60 elements in the class. Specifically, the examiner rates the use of these elements using a five-point scale, where teaching practice is classified as: innovating, applying, developing, beginning, or not using at all. After the observation, teachers are required to answer nine design questions and adjust their practice accordingly. These questions are:

- Design Question 1: What will I do to establish and communicate learning goals, track student progress, and celebrate success?
- Design Question 2: What will I do to help students effectively interact with new knowledge?
- Design Question 3: What will I do to help students practice and deepen their understanding of new knowledge?
- Design Question 4: What will I do to help students generate and test hypotheses about new knowledge?
- Design Question 5: What will I do to engage students?
- Design Question 6: What will I do to establish or maintain classroom rules and procedures?
- Design Question 7: What will I do to recognise and acknowledge adherence and lack of adherence to classroom rules and procedures?
- Design Question 8: What will I do to establish and maintain effective relationships with students?
- Design Question 9: What will I do to communicate high expectations for all students?

Lesson observations are facilitated by the provision of the iObservation tool by Learning Sciences International. The iObservation tool is an interactive web-based data management system for classroom walkthroughs and teacher and leadership evaluation designed for the Marzano Teacher Evaluation Model. Specifically, the iObservation platform allows principals and observers to make observation feedback available to teachers immediately, aligning growth, development, and evaluation in one tool.

There are a number of different stages of training that the evaluator needs to undergo in order to meet the requirements of the Marzano research team. This includes attending professional development and training sessions for the planning, communication and implementation components of the program. Ongoing application of this approach is also reliant on teachers and examiners being allocated sufficient time in their schedule to conduct the operations.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

The Marzano Teacher Evaluation Model is based on a weighting system grounded in research. Observational data is combined with teacher self-evaluation and student academic achievement in order to determine the observational score. Areas of weakness are identified, allowing principals and observers to provide detailed and targeted feedback to help teachers improve. As a component of the evaluation program there is provision of mentoring, demonstrations of the 41 Domain 1 strategies, a library of video and print resources, and associated technology tools to help teachers track student progress and collaborate with their peers to improve specific strategies.

VALIDITY AND RELIABILITY

The Marzano Teacher Evaluation Model has a large evidence base exploring the concept validity of the tool. In particular, the Marzano Teacher Evaluation Model was designed and created using an aggregation of the research on elements traditionally shown to correlate with student academic achievement. A meta-analysis of 300 studies showed that the strategies incorporated were associated with an average effect size of 0.42, and a maximum of 2.0 (Haystead, 2010).

There have also been a number of correlational, quasi-experimental, control studies examining the positive correlations and then causal linkages of the specific strategies with increased student achievement. It should be noted that researchers or organisations associated with the corporate arm of the Marzano Research Laboratory have conducted the vast majority of these studies, which should therefore be interpreted with a degree of caution. These studies include but are not limited to *What Works in Oklahoma Schools* (Marzano Research Laboratory, 2010), *The Adams 50 Instructional Model Study Report on Professional Development* (Marzano Research Laboratory, 2011a) and *Evaluation Study of the Effects of Promethean ActivClassroom on Student Achievement* (Marzano & Haystead, 2009). Each of these four studies examined the Marzano Teacher Evaluation Model in working classrooms and demonstrated positive correlations between the model and student learning.

The Marzano Research Laboratory has also conducted a number of studies examining the reliability of the model (Marzano Research Laboratory, 2011b). Two school districts in the United States, Cherry Creek School District in Denver, Colorado, and Rockwall Independent School District in Rockwall, Texas, were used as the basis for a reliability study in 2011. In particular, the study examined the level of agreement between raters when identifying which of the elements were exhibited and then the level of agreement between raters when assigning scores to teachers on specific elements. Across 41 different elements that were identified in use by the teacher the observers agreed on between 50% and 70% of the coding or scoring observed. It should be noted that this accuracy increased after group consultation. The authors also examined the reliability rating for a particular lesson and calculated the reliability at 0.75.

This model demonstrates strong links between elements of the tool and student outcomes; however, there is limited independent research to validate the tool, and it is linked directly to commercial application; consequently, there is limited information.

THE FRAMEWORK FOR TEACHING (FfT) EVALUATION INSTRUMENT (DANIELSON)

OVERVIEW

The Danielson Framework for Teaching (FfT) incorporates comprehensive and research-informed components of good teaching. The framework is a generic instrument, and is able to be applied to all disciplines. The framework features four domains and 22 teaching skills components. Danielson teamed with Teachscape in 2009 during the MET Project to develop a rater training and certification program that is available through the Teachscape site, Teachscape *Focus*, and a customised observation and evaluation management system, Teachscape *Reflect*. The FfT(Danielson, 2013a) expands on *Enhancing Professional Practice: A Framework for Teaching* (Danielson, 1996) devised for the evaluation of first year teachers for licensure purposes.

109

In 2008, Chicago Public School (CPS) launched the Excellence in Teaching Pilot, employing the Danielson FfT to revamp their model of teacher evaluation. There were two main drivers for the CPS endeavour: (a) the current evaluation systems were failing to provide teachers with meaningful feedback on their performance and (b) traditional teacher evaluation systems were not differentiating the best, good and poor teachers. A research report by Sartain et al. (2011) summarises the findings from a two-year study of Chicago's Excellence in Teaching Pilot that used the Danielson FfT.

Specific findings from the program were: (a) Classroom observation ratings using the Danielson Framework were valid measures of teaching practice and (b) Classroom observation ratings were reliable measure of teaching practice.

Observer training is available through a one-day face-to-face workshop for instructional leaders and district administrators. The workshop aims to coach and help strengthen inter-rater agreement and support observers in accurately rating teaching practices (Teachscape, 2016b). In-classroom as well as video-based observation training for teachers, observers, and teacher trainers is typically conducted over a 3 hour facilitated session to help new users using Teachscape *Focus* and Teachscape *Reflect* (Teachscape, 2016b). Observers need to have a number of skills to conduct fair and reliable observations for teaching, including an ability to collect and interpret evidence against level of performance, and to conduct professional conversations with teachers (Danielson, 2012). The Teachscape training sessions are tailored to the participants (teachers, observers, trainers and district leaders) to introduce them to the modules, scoring practice and reporting guide. Trainers are also exposed to the administrative side of using the tool (navigation features, video app, how to train teachers in their respective schools). Additionally, observers learn test-taking strategies and discuss calibration and feedback (Teachscape, 2016a). Full implementation of the FfT tool may be costly as it would require system/software purchases, activation fee, yearly access to resources and training sessions for teachers, observers and trainers.

Moss (2015) conducted a study in a large rural high school in New Jersey to determine teacher perceptions of the Danielson FfT. Most teachers believed that their supervisors were well trained as evaluators to identify and assess teaching skills against the components and levels of the FfT in order to conduct accurate and meaningful evaluations. However, some teachers who believed the FfT to be subjective were of the opinion that evaluators place their own interpretation on the rubric and established relationships between teachers and evaluators influences the subjectivity of the model, increasing bias. This is consistent with the findings of the Excellent Teaching Pilot (CPS, 2008) that principals gave higher ratings to their own teachers at the top end of the scale than external evaluators due to their established knowledge of the teacher or the influence of previous assessments.

VALIDITY AND RELIABILITY

To test validity, a statistical model was used to compare the FfT ratings (of each component by an observer) against the value-added measures[1] of teachers for reading and math, giving a total of 20 models. Results for math and reading were found to be comparable. On average, teachers who received lower classroom observation ratings also have lower value-added measures. Consistent correlations between observation ratings and value-added measures suggest that the FfT ratings are a valid measure of teacher practice, given that the value-added measures of teachers are accurate[2] (Sartain et al., 2011).

Further work in this area was conducted by Milanowski (2011), who obtained data from three sites in the US: Cincinnati, Coventry and Washoe. Teacher evaluation scores (represented by classroom-level value-added estimates

---

[1] See pg 10-11 and pg 55 of *Rethinking Teacher Evaluation in Chicago Lessons Learned from Classroom Observations, Principal-Teacher Conferences, and District Implementation (Sartain, Stoelinga, and Brown, 2011)* for details on CPS value-added measure, validity and data limitations.

[2] See pg 55-56 of *Rethinking Teacher Evaluation in Chicago Lessons Learned from Classroom Observations, Principal-Teacher Conferences, and District Implementation (Sartain, Stoelinga, and Brown, 2011)* for results from significant testing.

of teachers) and student achievement (test results) were measured. Teacher evaluation scores (using the FfT) were then compared with teachers' value-added estimates to test for validity/reliability. It was concluded that ratings from a system based on the FfT are reliable and are consistent with value-added estimates of teacher effectiveness[3]. Well-established structural and implementation protocols, including a series of observations by multiple (trained) observers, must be incorporated for these findings to be extended to the school district or state level.

To measure reliability, principal and observer ratings of the same lesson were compared. The rating scale was isolated into 3 parts (low end, middle, and high end) to determine if principals and observers rated consistently across the four categories of performance in the FfT (Unsatisfactory, Basic, Proficient, and Distinguished). Two types of reliability analysis were conducted: one to test for inter-rater reliability (using hierarchical modelling), and another to understand differences in principal severity using a Many-Facet Rasch Measurement (MFRM) analysis (Sartain et al., 2011).

For inter-rater reliability, principals and observers generally agreed when practice was Unsatisfactory-Basic (lower end) and Basic-Proficient (middle of the scale). Given this finding, the FfT is concluded to be a reliable tool at the low end of the rating scale. However, at the high end of the scale, principals and observers rated teachers differently[4]. Using a Many-Facet Rasch Measurement (MFRM) analysis, rater severity was measured (this takes into account that there are many different raters/judges). The individual teacher measures of teaching ability generated in this analysis combine all FfT component ratings from the observer and the principal across 2 years for an individual teacher. The individual teacher measure of teaching ability using the FfT was highly reliable (reliability = .93 and separation=3.60). Another important finding is that each of the FfT components was treated as distinct aspects of teaching (reliability= .98, separation = 7.11). In this case, reliability is a measure of how well each of the components can be separated from each other, and the reliability of the FfT components is high (i.e., close to 1).

CLASSROOM ASSESSMENT SCORING SYSTEM (CLASS)

OVERVIEW

The Classroom Assessment Scoring System (CLASS®; (Pianta et al., 2008) is an observational instrument developed at the Curry School Center for Advanced Study of Teaching and Learning to assess classroom quality in Pre Kinder to Year 12 classrooms. The CLASS instrument measures teacher-student interactions in a classroom setting and offers resources for strengthening those interactions across any subject area or age group. CLASS tools examine the emotional, organisational and instructional supports provided by teachers, which are linked to students' social, developmental, and academic achievement. Several forms of CLASS have been constructed for various populations, including: infants, toddlers, pre-kindergarten, elementary, upper-elementary and secondary school settings. The CLASS tools used for assessment of classroom practice in schools are structured across the three domains of Emotional Support, Classroom Organisation and Instructional Support with a range of 10 or 12 related dimensions depending on the stage of schooling. Teachers are observed and rated using the stage-specific standardised manual.

The CLASS tool includes four cycles of 15-minute observations of teachers and students by a certified CLASS observer. Those observations are then rated using a manual of behaviours and responses. Research suggests that from Pre-K programs into the third grade, children in classrooms with higher CLASS ratings realise greater gains in achievement and social skill development. CLASS can be used for both formative and summative evaluation purposes and across contexts. The emphasis on strengthening teachers' relationships with their students differs

---

[3] Many combinations of results/correlations are available for different schools and domains on the FTT on pg 11, 12, 15 and 19 of *Validity Research on Teacher Evaluation Systems Based on the Framework for Teaching (Milanowski ,2011).*

[4] It is suggested that these differences are due to principal's historical knowledge of teacher's previous evaluation ratings. See pg 15-16 for more details.

from most other evaluation systems, despite evidence that student/teacher relationships are a central aspect of students' learning outcomes (Pianta, 1999).

CLASS is a commercial venture, and there are substantial costs associated with purchasing the tool and associated training. CLASS has a demanding training regime relative to other evaluation tools. Specifically, evaluators must participate in 2 full days of training delivered by an accredited CLASS representative, followed by an online reliability test. Access is then provided to additional online reliability tests and professional development materials. To maintain their CLASS registration, assessors must participate in ongoing training and development and regular tests for reliability. The extensive financial and training commitments can act as a barrier for the employment of CLASS as an observational instrument.

VALIDITY AND RELIABILITY

There is conflicting evidence regarding the validity of CLASS, and it is acknowledged that further research is needed to clarify its connection to student academic achievement. Of note is the observation that impartial investigators have conducted few studies investigating CLASS' reliability and validity, with most of the research done by authors involved in the development of CLASS.

Support for the validity of CLASS is provided by Hafen, Hamre, Allen, Bell, Gitomer and Pianta (2014). Their factor analysis, using data from over 1,400 classrooms, showed student-teacher interactions loaded onto three factors that represent the three domains: emotional, organisational and instructional. The paper also provided evidence that these types of interactions are important to student learning. Bell et al. (2012) found that scores of 82 mathematics teachers on CLASS-S, the secondary school component, had a small positive correlation with their Value Added Modelling (VAM) score. However, this effect was not maintained when student factors were controlled. Both CLASS-S and VAM are intended to assess teacher effectiveness; the absence of a statistical relationship between the measures suggests a lack of concurrent validity. The presence of a weak correlation when student factors were not controlled for indicates that these assessments may be measuring student factors instead of teacher effectiveness. More recent research (Allen et al., 2013), which included 1,276 students across 78 classrooms, found a positive correlation ($r = .78$) between teachers' CLASS ratings and student outcomes. This relationship was maintained when classroom and student variables were controlled. Validation studies have also been conducted across cultural contexts, specifically for the kindergarten version of CLASS (Downer, Booren, Lima, Luckner, & Pianta, 2010; Pakarinen et al., 2010).

There have been few investigations into the reliability of CLASS. Although the pre-K version of the tool has been found to have a high inter-rater reliability, with scores consistent 87% of the time (Pianta et al., 2008), there is no evidence for the later years versions. However, several in-built features of the tool are designed to increase its reliability. Evaluators are required to participate in extensive ongoing training, which may promote more consistent and reliable evaluations. Relatedly, CLASS is a structured observation measure, whereby observers follow the same standardised behavioural indicators, which again encourages inter-rater reliability. Finally, these evaluation rubrics differ depending on the student population; appropriate teacher interactions change as students age, and this consideration of developmental differences may also increase reliability (Pianta et al., 2008).

QUALITY TEACHING (QT) MODEL AND QUALITY TEACHING ROUNDS

The Quality Teaching (QT) model is based on the major elements that constitute good classroom and assessment practice. The QT pedagogical framework corresponds with the three dimensions of the New South Wales (NSW) Model of Pedagogy: intellectual quality, quality learning environment, and significance, each of which comprises 6 select research defined elements linked to improved student achievement (New South Wales Department of Education and Training, 2003). The 18-element QT model can be adapted to different contexts and is suitable for all curriculum areas and stages of schooling. The QT model provides a common language for critical reflection and discussion by teachers and schools and is coded to generate powerful professional dialogue.

112

Quality Teaching Rounds (QTR) utilises the QT Model to inform and improve decisions about student learning. Based on educator research at the University of Newcastle, School of Education, QTR (Bowe & Gore, 2011) involves objective observations of school and classroom practice by a Professional Learning Community (PLC).

OVERVIEW

Commissioned and launched by the (then) NSW Department of Education and Teaching, the QT Model was hailed as a 'long-term strategic priority' (Gore, Mockler, Smith, & Bowe, 2012). The QT Model is a framework that builds on Authentic Pedagogy (Newman, Marks, & Garmoran, 1995) and Productive Pedagogies (The University of Queensland, 2001) and identifies the key elements of quality teaching. It focuses on what teachers already know and value, aiming to develop the analytical voice of teachers as they critically reflect upon, describe and diagnose their pedagogical practice both individually and collectively within PLCs.

QTR involves professional reading and discussion as well as classroom observations. Four key questions are employed for consideration by teachers when planning learning experiences and reflecting on their practice: 1) What do I want the students to learn? 2) Why does the learning matter? 3) What do I want the students to do or produce? 4) How well do I expect them to do it? (NSW DET,2008). Members of the PLC observe participants as they teach a lesson. Observers code the lesson using a 1-5 coding system outlined in the Quality Teaching Classroom Practice Guide to inform and frame the feedback and discussion with a focus on teaching and learning and how the lesson exemplifies teachers' philosophical and pedagogical beliefs within the particular school and classroom context (Bowe, Gore, & Miller, 2010).

VALIDITY AND RELIABILITY

A rigorous internal program of research has been undertaken by the University of Newcastle to determine the effectiveness of the model on the tripartite goals of teacher learning, teaching quality and student learning. This research has largely been in the form of ARC Linkage Grants and pilot studies. The authors cite effect sizes in excess of 1.0 for teaching quality and teaching satisfaction in intervention schools participating in QTR (Gore et al., 2012).

The internal nature of this research, conducted by researchers involved in the design of the QT Model, should be considered when interpreting the collective findings:

- A high correlation between the QT model and improved outcomes and greater equity for students from low SES and Indigenous backgrounds; (Systemic Implications of Pedagogy and Achievement in NSW Public Schools [SIPA], ARC Linkage grant, 2003-2007) (Ladwig, Smith, Gore, Amosa, & Griffiths, 2007);
- A correlation between QT Rounds and improved teaching quality, NAPLAN results, and increased teacher satisfaction (Effective Implementation of Pedagogical Reform [EIPR], ARC Linkage grant, 2009-2012);
- Positive impacts can be achieved in a short timeframe using modified versions of the QT Rounds, reducing expense and increasing sustainability of the model (Investigating Quality Teaching Rounds to Support Teacher Professional Learning, ACT ETD pilot study, 2012).

With the aim of connecting teacher professional learning needs to a comprehensive change process focussed on improved student achievement, one project from the Australian Government Quality Teaching Program (AGQTP) conducted by New South Wales Quality Teaching Action Learning Project (QTAL) (2003-2009) studied the QT Framework (QT) (NSW DET,2003) as one of 60 projects under review.

The Meta-Analysis of Quality Teaching Action Project was a quantitative and qualitative analysis of the written reports of 182 QTAL participating schools. It found limited evidence of improved student learning beyond minor inferences about active student learning, quality learning, and improved engagement in the reports of 8 out of 57 schools in a 2007-9 phase (Ewing et al., 2009). QTR was found to play a significant role in focusing teacher professional learning leading to improved pedagogical practices and teaching practice. The report comes with a caveat on the findings citing unequal timing of phases and the late release of QTAL documentation to schools and supporting documents for QT strategies for professional learning. This impacted on the opportunity for teachers in

113

the first phase to learn about and engage with QT compared to those in later phases. There was also no direct question about the use of QT by school teams in their action learning or about impact on teaching practices or student learning.

STANDARDISED ASSESSMENTS, INTERVIEWS AND PORTFOLIOS

The section outlines a series of measures that while in variable forms have been standardised for use at scale.

EXAMPLES OF STANDARDISED ASSESSMENTS AND PORTFOLIOS

TEACHING PERFORMANCE PORTFOLIOS

Teaching Performance Portfolios (TPPs) require teachers to compile documented evidence of current practice. Evidence of teaching performance is typically assembled by the teacher, representing work across a short teaching period to gain understanding from a specific subject area, or over the course of a year. Examples of evidence typically included in a TPP include professional growth plans, lesson plans and resources, instructional videos, student achievement results, and evidence that displays student performance change (Darling-Hammond, 2013). The process of compiling a TPP provides the opportunity for teachers to reflect upon and assess their own practice. TPPs can be used as formative evidence to inform teacher professional development (van der Schaaf, Stokking, & Verloop, 2008), or as a summative tool for evaluating teaching performance against teacher quality standards.

Although TPPs can be a useful teacher evaluation tool, there is mixed evidence regarding the purpose, validity and reliability of the tool. Beck, Livne, and Bear (2005) emphasised that an important dilemma in designing portfolios is whether the portfolio is primarily a vehicle for teacher assessment or for teacher development. TPPs are a complex source of evidence, and standardisation across evidence presented is often considered a design weakness. Fundamental issues for evaluation include the high demand in dedicated time and concentration required to assess portfolios, multiple methods for weighting of the various portfolio components, and the quality of judgement. The quality of portfolio assessment is defined by the clarity and defensibleness of the component parts, the quality and reliability of the scoring rubrics, and the rigor in the standard setting process on which judgments are based.

As teaching portfolios need to be bodies of work that accurately reflect quality of subject knowledge and teaching practice the working context of the situation needs to be taken into consideration (van der Schaaf et al., 2008). The major issue, however, is reliable and valid scoring of the portfolio, ensuring that there is defensible weighting of the various parts of the portfolio, and that the resulting scoring or decisions are replicable over time, contexts, and teachers. Further development needs to be made in determining accurate portfolio criteria and the required content in order to display the level of teacher performance. Currently, there is an absence of consistent and defined measures for teacher portfolio ratings in teacher evaluation.

Teacher formulated portfolios require a significant amount of time and effort in order to prepare. Therefore, teachers need be incentivised in order to see the tangible benefits and then subsequently consistently and accurately format effective portfolios. A conflict of interest may arise as a result of teachers compiling their own body of work as a representation of their teaching capacity. Firstly, the expectations of the portfolio need to be clear, as teachers are required to adhere to external assessment standards when they are compiling the portfolio. A second issue may be associated with the intended purpose and use for the portfolio. This is especially the case when the portfolio is seen as multi-purposed e.g. as both summative evidence of teaching practice and as an evidence base for professional development (Tillema & Smith, 2007*).

Compared to other evaluative tools, portfolio creation and analysis is more time consuming for both teachers and evaluators. Despite the tendency of portfolios to be a combination of a term's or a year's work they are often only formed into a portfolio in a short space of time. The time spent on portfolio formation can vary depending on the evaluation model and the content and the form that the portfolio takes. In order to get valid evidence of teaching

114

practice through the use of a portfolio, time needs to be invested in the formation and collation of evidence in order to give an accurate representation of teaching quality (Martínez, Borko, Stecher, Luskin, & Kloser, 2012).

In order for teachers to present a broad evidence-base on the quality of their teaching practice, they are required to have knowledge of the expected content for a portfolio. It is also necessary that portfolio reviewers receive training in portfolio analysis, especially if content is rated against a rubric.

Portfolios can provide teachers with the opportunity to present their voice and opinion as it relates to their teaching practice. This is evidenced by the inclusion of the teaching reflection piece in the portfolio. As a result portfolios can prove beneficial for teaching practice and improvement (Attinello, Lare, & Waters, 2006). A disadvantage of portfolio use in teacher evaluation is the length of time involved in the process. Teachers are required to allocate time in order to compile and form evidence of their teaching practice. When teachers do not see the benefit of portfolio formation, they may not put the required time or thought into forming their portfolios and not give an accurate impression of their teaching performance (Attinello et al., 2006). Using teacher portfolios in conjunction with other evaluative tools may allow for a broad and rich evidence base for teacher appraisals (Tillema & Smith, 2007).

Explicit evaluation criteria can not only benefit the evaluation process but also provide a base for teacher improvement. Making evaluation criteria clear to teachers as well as assessors can prove beneficial to meet this purpose (Tillema & Smith, 2007).

THE TEACHER PERFORMANCE ASSESSMENT PORTFOLIO (EDTPA)

The Teacher Performance Assessment Portfolio (edTPA) is an evaluation and support framework designed to assess pre-service teachers. EdTPA is used in over 600 registered initial teacher education programs across the United States. The Stanford Centre for Assessment, Learning, and Equity (SCALE) and the American Association of Colleges for Teacher Education created edTPA. Stanford University is the exclusive owner.

The edTPA framework is available in 27 separate tests in order to encompass the required range of disciplinary expertise. Candidates are required to articulate their practice through commentary tasks organised into the three domains: Planning, Instruction, and Assessment. Candidates answer five essential questions that inform the evaluation framework. These are submitted as portfolios to an online platform with accompanying video footage of their clinical teaching practice. Candidates receive feedback on their performance and are given access to online learning tools (such as video tutorials) on how to improve their teaching. There are several fees associated with edTPA. The candidates, who pay approximately $600 per submission, generally incur these costs. The literature points to scepticism amongst teachers and academics about the use of a for-profit program, and the corporatisation of teaching more generally (Greenblatt & O'Hara, 2015; Ledwell & Oyler, 2016).

Candidates' teaching portfolios take at least one month to be evaluated and scored. The timeframe for data collection varies between programs, but generally occurs towards the end of their clinical teaching period. The organisation responsible for the marking of portfolios advertises that expert markers conduct evaluations; however, this claim has recently become a point of contention. Information about marker qualifications is not available, and reports suggest that the qualifications of markers vary widely, which may lead to markers handling portfolios that are inappropriate for their background and level of experience (Parkes & Powell, 2015).

Unlike many teacher evaluation tools, edTPA differentiates between subject with individualised assessment criteria, and thus many have claimed this makes it a more reliable assessment of teaching quality (Harris & Sass, 2011; SCALE, 2014). Furthermore, it uses multiple measures of authentic teaching performance, opposed to relying on pen and paper tests. While edTPA is primarily used in high-stakes summative assessments, there is evidence that it could be used for formative evaluations and to promote professional development (Lin, 2015).

Ledwell and Oyler's (2016) case study of a university as a first time user of edTPA revealed that there were a number of practical barriers to employing the evaluation tool. The study reported ongoing problems accessing and submitting online information. Alarmingly, a number of academics claimed that the introduction of edTPA to their

115

university negatively impacted upon the teaching curriculum in that it was adapted to enable teacher candidates to perform well on the edTPA rather than to promote effective teaching more generally. Greenblatt and O'Hara (2015) are in agreement, noting similar concerns that the introduction of edTPA and similar programs "shifts the focus of the student teacher experience to test preparation" (pp. 62).

VALIDITY AND RELIABILITY

While proponents of edTPA provide an array of evidence for its validity, these results have been challenged. EdTPA is a research-based assessment, and SCALE undertook several rounds of content validation during its creation (SCALE, 2014). One SCALE-conducted study validated edTPA's content, with ratings indicating that expert participants believe the tasks and assessment included in edTPA were valid measures of teacher effectiveness (SCALE, 2014). Unlike most other measures of teacher effectiveness, edTPA has individualised appraisal guidelines for 27 different subject areas. This can be seen to improve the validity of the instrument, as the behavioural indicators for effective teaching have been found to vary across subject areas (Harris & Sass, 2011). Furthermore, Pecheone and Chung (2006) found evidence for the predictive validity of the predecessor of edTPA, PACT (Performance Assessment of California Teachers). A study of 1870 teachers found a positive correlation between participants' PACT scores and their students' VAM scores.

In contrast, a case study of a university using edTPA for the first time revealed that many academics working as initial teacher educators questioned its validity as a measure of teacher performance or classroom readiness (Ledwell & Oyler, 2016). Interviews with 19 teacher educators revealed that many were sceptical of the edTPA and believed that it did not provide an accurate assessment of pre-service teacher's capabilities. Several participants reported instances where top-performing students (who they were confident would make good teachers) had not passed the assessment. The interviewees argued that the edTPA evaluation rubric did not cover the full range of behavioural indicators and therefore refuted its validity as a tool for evaluating teacher effectiveness. Similarly, Greenblatt and O'Hara (2015) argued that edTPA places too much weight on candidates' analytical skills, and not enough emphasis on skills and behaviours integral to teaching, such as adaptability, relating with students and building a cooperative classroom environment. Furthermore, they suggest that edTPA may penalise teachers working in more difficult classroom contexts, as there are extra demands for candidates working in special needs classrooms (e.g., English language learners) in comparison with those working in more homogenous classrooms.

EdTPA has been shown to have high inter-rater reliability and internal consistency. The 2014 administrative data examined how consistently raters scored teacher candidates' portfolios, with 1,808 portfolios being randomly assigned to two raters. Rater scoring was consistent. While the marks were identical in 50.1% of cases, they differed by a score of one with a 93% instance (SCALE, 2014). Furthermore, edTPA found there to be a high level of internal consistency, with an overall Cronbach's alpha score of 0.92. The use of an external evaluator mitigates potential evaluator bias that may be present in other appraisal methods where the observer is familiar with the teacher and their practice. Conversely, internal evaluators have the advantage of observing the candidate regularly and across contexts and time.

THE NATIONAL BOARD FOR PROFESSIONAL TEACHING STANDARDS

The National Board for Professional Teaching Standards (NBPTS), founded in 1987, is a nonpartisan, nonprofit organisation in the United States, dedicated to defining and recognising accomplished teachers (Hattie & Clinton, 2008). National Board Professional Teaching Standards articulate what teachers know and should be able to do based on Five Core Propositions that form the basis for National Board Certification (National Board for Professional Teaching Standards, 2016b). Representing four developmental levels and sixteen subject areas, NBPTS cover 25 certificate areas and apply to the majority of teachers in U.S. public schools. With the majority of states offering financial incentives or assessment bonuses for certified teachers (Exstrom, 2011) over 112,000 educators hold National Board Certification in the United States.

National Board Certification is a voluntary system for advanced certification that employs a teacher performance portfolio assessment. Applicants are required to submit a portfolio containing artefacts to authenticate their instructional practice such as video footage of classroom lessons, exemplars of student work and reflective essays. Applicants also complete online exercises that assess their pedagogical and subject content knowledge. The National Board has engaged Pearson to assess the portfolio and online assessment centre exercises at locations across the United States (Pearson VUE, 2016). Assessors must demonstrate their ability to use the National Board scoring system, the score scale, rubrics, benchmarks and other tools (Pearson, 2016).

VALIDITY AND RELIABILITY

Engelhard, Myford, and Cline (2000) used the FACETS model (many-facet Rasch measurement) (Linacre, 1989) as a framework for analysis of the psychometric quality of NBPTS assessments. They concluded that intensive assessor training and ongoing monitoring of assessor performance, results in extensive disciplining of their judgments and generally consistent use of the rating scale.

Assessment items are evaluated separately and aggregated for the final scale score. Assessor training includes self-examination to identify bias and preferences. Cantrell, Fullerton, Kane, and Staiger (2008) found that assessment scores used by NBPTS are weighted according to professional judgments about the relative importance of various teacher attributes and skills, and may not be the best predictors of teacher quality across various permeations and combinations of the NBPTS sub-scores.

There are numerous observational studies of National Board Certified Teacher (NBCT) effectiveness that report general findings of increased point estimates in the range of 0.01 – 0.03 standard deviations on statewide assessments, or on average about 2-10% of an average year learning gains (Cantrell et al., 2008; Chingos & Peterson, 2011; Clotfelter, Ladd, & Vigdor, 2007; Goldhaber & Anthony, 2007; Harris & Sass, 2009). Hattie & Clinton (2009) detected a difference in the nature of teaching and learning in NBCTs classrooms compared to non-NBCTs classrooms, with the former displaying a greater balance of surface to deep learning and the latter showing a higher proportion of surface-level teaching and learning.

In a comparison of NBCTs and those who hold only advanced degrees, Harvard University's Strategic Data Project (SDP) found that Board-certified teachers, on average, outperform others with the same level of experience in elementary literacy by 0.03 and numeracy by 0.07, equivalent to one month of additional literacy instruction or two months of additional mathematics instruction (Strategic Data Project, 2012). Similarly, in Washington State (with the largest national population of NBCT) Cowen and Goldhaber conducted a longitudinal study based on value-added models in math and reading to find that NBCT are about 0.01 – 0.05 standard deviations more effective than non-NBCTs with similar levels of experience (Cowan & Goldhaber, 2015).

Cavalluzzo, Garrow & Henderson were commissioned by the NBPTS to analyse National Board certification among high school teachers in understudied subject areas and locales (Cavalluzzo, Barrow, & Henderson, 2014). Undertaking classroom observations and examining test scores, they found ratings of the instructional practices of NBC applicants exceeded those of non-applicants on 66% of the teaching quality subscales, and overall. They reported similar effect sizes to those found elsewhere in the literature and endorse National Board Certification as an effective signal of teacher quality.

INTERVIEWS, SURVEYS AND RATINGS

RATING SCALES

Ratings can be considered "a classification or ranking of someone or something based on a comparative assessment of their quality standard or performance" (Oxford Dictionaries, 2016). A rating often implies a continuum from good to bad. Ratings of teachers are typically focused either on one single dimension or multiple dimensions that have been collected separately or subsequently consolidated in the form of a single comprehensive rating. There are a

117

number of concerns that relate to rater bias, as such bias can relate to issues of a predisposition to supply a lower rating or a more generous rating. Further there is the issue of errors of central tendency where a rater is inclined to assign middle quality ratings irrespective of the person being rated. Within the area of teacher evaluation there are often three forms of ratings that provide information to teachers, senior administrators, students, and parents.

Educators frequently disregard student ratings due to reliability concerns.   However, Marsh and Roche (1997) demonstrated high estimates of reliability from student surveys (although most of this research involved university students). Further, the MET study has enhanced the credibility of student ratings and they are now incorporated more frequently in teacher evaluation initiatives.

The quality of evidence collected by ratings is more often than not dependent on the merits of the rating form used, with an evidence-based approach to developing a rating form considered essential. Furthermore, as with observations, preparation of raters in building capacity around evaluative thinking is critical for the success of rating forms. This also applies to the preparation and support of students who will be asked to rate teachers.

TEACHER INTERVIEWS

Teacher interviews are often discussed as an important component of many teacher evaluation programs (Isoré, 2009; OECD, 2013a). This typically involves the teacher engaging in a (mostly structured) conversation regarding their current practice with another individual, most often a senior staff member (Isore, 2009). Teacher interviews also occur prior to formal evaluations for goal setting purposes, and following evaluations for informing ongoing development. Interviews may also be structured or unstructured.

While often cited as an evaluation tool, teacher interviews may more accurately be defined as a delivery system for other evaluation methods, such as teacher self-rating or principal evaluations (Goe, Bell, & Little, 2008a). In the literature, there is often overlap between components of teacher interviews and teacher self-assessment (Ross & Bruce, 2007). For example, teacher interviews typically involve reflection on teaching practice, strengths and weaknesses, and goal setting (Ross & Bruce, 2007). Additionally, the purpose of teacher interviews may be more to consolidate and engage in constructive conversation regarding evaluation outcomes, rather than as an evaluation tool in and of itself.

There is little evidence to support the reliability of interviews. As with other evaluation methods, relationships between the evaluator and teacher have the potential to influence evaluation results. No empirical evidence was found relating to the potential for bias in teacher interviews for evaluation. While there may be some preparation required prior or following the interview, such as pre-interview questions or follow up questions, the time required to implement teacher interviews is entirely based on the duration of the interview. Further, the frequency of interviews can vary considerably, and is entirely based on school preferences.

While there is generally no expertise cited as required, it is generally accepted that a senior staff member, either the school principal or delegated to a senior teacher, conducts interviews. Teacher interviews allow teachers to take a more active role in their evaluation process. As such teachers have the ability to place evaluation outcomes in context, or ask specific questions regarding the evaluation process. One barrier to implementation for teacher interviews may be the time required for implementation. Specifically, the evaluator (often the principal) is often required to meet with the teacher if an interview proves difficult to schedule.

TEACHER SELF-RATING

Teacher self-ratings are a form of evaluation that requires teachers to rate their own teaching methods, skills and classroom practices. Self-ratings assist teachers to reflect on their current performance, strengths and weaknesses and can guide in setting goals that improve their teaching practice. Self-ratings are usually in the form of online survey tools or questionnaires and video-logs.

118

Teacher self-ratings can be an efficient evaluation method as there is no need for an external evaluator and information can be easily gathered for a large population through surveys/questionnaires. However, there are concerns about self-ratings measurements being biased and inaccurate.

Studies on the validity of self-ratings have yielded mixed results. The literature suggests that often teacher self-ratings only occur annually, therefore the accuracy of the self-rating is impacted by the teacher's long-term memory of the evaluation period. The frequency of the evaluation will affect the self-rating with data collected less frequently inevitably being less reliable than data collected frequently. Moreover, different teachers may have different understandings of the same terms, making results less accurate.

Ozogul and Sullivan (2009) investigated the effects of teacher, self, and peer evaluation on 133 pre-service teachers' lesson plans. The results indicated that when student teachers received appropriate training and feedback in conducting formative evaluation, (fractional difference of 16.80-16.61 in comparison to teacher evaluation) peer evaluations (16.80- 15.86) are appropriate alternatives and complementary approaches to teacher-based formative evaluation. This can be helpful when teachers are unwilling to spend the time required for formative evaluation and final evaluation of the student teachers.

Yuan et al. (2014) conducted a study using anchoring vignettes to calibrate 61 mathematics teachers' self-rating of teaching. They compared correlations between teachers' self-rating and value-added scores, student surveys, and observation ratings before and after calibration. The findings report that anchoring vignettes improve the accuracy of teachers' self-reports but do not mention values of the correlation coefficients.

As teacher self-rating requires relatively little external input from other evaluators (e.g. principals, peers, external evaluators, etc.) it may be assumed that it could be a highly sustainable model. It should be noted however that this approach is dependent on teachers being allocated sufficient time in their schedule to conduct self-ratings.

As mentioned earlier, possible biases associated with self-ratings can be a concern as it is common for teachers to favourably self-rate. Self-presentation bias may also be present and it has been acknowledged that it can be difficult to detect problems in one's own work, which might cause someone to reflect inaccurately (Ogozul & Sullivan, 2009). Self-rating results also suffer from potential biases due to differences in respondents' understanding of the latent constructs being measured, as well as interpretation and application of the scale used to quantify their practices (Yuan et al, 2014).

Mathers and Oliva (2008) highlight that time and cultural norms can be a limitation in teacher self-rating. As self-rating requires making time for teachers to engage and reflect, this practice is dependent on how highly administrators prioritise self-rating practices over other duties teachers may need to perform.

Teacher self-reports through surveys and classroom logs are more time efficient and relatively easy to administer over large populations (Yuan et al, 2014). However, as mentioned earlier, self-ratings suffer from biases.

As Yuan et al (2014) noted it is crucial that teachers have an understanding of the latent constructs being measured, as well as interpretation and application of the scale used to quantify their practices.

EXAMPLES OF TEACHER SELF-RATING TOOLS

Specific self-rating tools identified in this review include the Classroom Strategies Scales-Teacher Form (CSS-T) and the Teacher Evaluation Questionnaire (TEQ).

CLASSROOM STRATEGIES SCALES-TEACHER FORM (CSS-T)

Reddy, Dudek, Fabiano, and Peters (2015) tested the reliability and validity of a new teacher self-rating measure called Classroom Strategies Scales-Teacher Form (CSS-T). In order to determine test-retest reliability, teachers across 60 classrooms self-rated their classroom practices using the CSS-T for two occasions within seven days of each other

119

and during different academic subjects. Approximately two weeks later, teachers performed a second administration of the CSS-T during the same two subjects. Test-retest reliability was found for the Instructional Strategies (*r* = 0.79) and Behavioural Management Strategies (*r* = 0.84) total scales.

TEACHER EVALUATION QUESTIONNAIRE (TEQ)

Embiza and Hadush (2015) tested the dimensionality and reliability of the Teacher Evaluation Questionnaire (TEQ) using 459 teacher's rater forms. Completed by principals, the questionnaire had six distinct dimensions of teaching effectiveness. No measures of validity were discussed. Cronbach's alpha was used to determine the internal consistency of the CSS-T tool for Instructional Strategy (IS) and Behavioural Management Strategy (BMS) using 309 classes. The total scales estimates were 0.93 and 0.94 for IS and BMS respectively. The TEQ study yielded alpha coefficients that varied from 0.38 to 0.9, indicating low/moderate to high internal measures of the six factors, confirming the questionnaire reliability.

STUDENT RATING

The most common form for collection of student opinion for the purpose of teacher evaluation is the survey. Student ratings typically include measures of teacher capacity relative to the subject area as well as the classroom environment. There is no standardised student feedback tool that is used to assess teacher performance.

The literature on student feedback is not as extensive for k-12 students compared to students in higher education, where it is common practice for students to do a teacher performance review following the completion of their subject. Though the evidence base for this area is limited, promising results suggest surveys have the potential to serve as a measure of teacher effectiveness (Balch, 2012).

A most critical aspect contributed by the teacher is the quality of their teaching as perceived by the students**.** (Irving, 2004) created a student evaluation of high school mathematics teachers based on the National Board for Professional Standards for this domain (www.nbpts.org). After completing a study on the psychometrics of the instrument in New Zealand, he then located United States teachers who had passed National Board Certification in high school mathematics, and compared their students' evaluation with teachers in the same schools who had not passed. The students were accurate judges of excellence and could discriminate between teachers who were experienced and expert from those who were experienced and non-expert. The dimensions that contributed most to this discrimination had a focus on cognitive engagement with the content of the mathematics curriculum and the development of a mathematical way of thinking and reasoning. It is what teachers get the students to do in the class that emerged as the strongest component of the accomplished teachers' repertoire, rather than what the teacher, specifically, does. Students must be actively involved in their learning, with a focus on multiple paths to problem solving. As mathematical thinkers and problem solvers, the students are also encouraged to go beyond the successful solution of the problem to include the interpretation and analysis of the solution. All the while, students are encouraged to greatly value mathematics and the work that they do in mathematics and always check the quality of their work to strive for the very best standards. As Irving argued, we should not overlook those who are arguably in the best position to evaluate the teachers—the students who share the classroom with the teacher day in and day out. This research does not support myths that students are capricious and that they are likely to award their teachers high grades (Irving, 2004). High ratings were not awarded lightly (Bendig, 1952; Tagomori & Bishop, 1995).

It has been argued that student perceptions of teacher performance are highly important to evaluate effectiveness considering the close proximity that students have to teacher performance. There is a range of different student rating tools which have predominantly been formed in the United States. Depending on the tool, there can be different survey versions for both student age groups and subject areas (Balch, 2012). Combining evaluative tools can improve validity of findings. A study conducted by the Bill & Melinda Gates Foundation (2012) found that including student surveys alongside other teacher evaluation tools improved correlations with student achievement ratings.

120

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Student teacher ratings are a reliable tool for gauging the effectiveness of teacher performance. There is mixed evidence on how much weight can be given to using student ratings in teacher evaluations (Clayson, 2009; Polikoff, 2015). However, when combined with other evaluative tools student ratings can prove to be an effective and reliable method (Balch, 2012). Student ratings are easily administered across contexts, are cost effective and can track change over time (Little, Goe, & Bell, 2009).

There is a potential for bias in student teacher ratings. This may be due to personal relationships and changing student opinions. Surveying a large number of students at one time can help to mitigate any changing and questionable performance ratings (Little et al., 2009; Polikoff, 2015). Relatedly, students may feel pressured to give their teachers a positive rating. This can be minimised by conducting the surveys anonymously, to allow students to present their honest opinions, without fear of punishment (Roe, 2010). Student ratings of teachers are time efficient, depending on the length, and can take at most one class lesson (Little et al., 2009).

No expertise or qualifications are required for students to take teacher review surveys. Students need to have known the teaching practices of the teacher they are evaluating. Furthermore, student surveys are easy to administer, are time efficient, and can be used in both summative and formative methods of evaluation. However, there may be potential for bias as students may favour one teacher over another for personal reasons rather than teaching capability. Aligning tool content with teacher performance standards allows for the gathered data to accurately reflect teacher performance as results align with these standards (The Colorado Education Initiative, 2013).

EXAMPLES OF STUDENT RATING SCALES

TRIPOD STUDENT SURVEY

The Tripod Student Survey is designed to measure teaching, student engagement, school norms, and student demographics. The survey measures student views on teaching, grouping items under seven constructs called the "7 Cs": Care, Control, Challenge, Clarify, Confer, Captivate, and Consolidate. Each survey requires students to respond to a series of statements, indicating their level of agreement on a five-point scale. The Tripod Student Survey is widely used for the purpose of teacher evaluation in the United States as part of a broader project aimed at strengthening content, pedagogy and classroom relationships. The survey was designed by Ferguson (2009) as a measure for understanding and improving teacher performance.

STUDENT PERCEPTION SURVEY

The Student Perception Survey examines student perceptions of specific classroom environments, class content, and teacher practice (NYC Department of Education, 2016). The survey was formed by the Colorado Education Initiative with the primary goal of creating an effective and reliable tool that will give representative results of student perceptions of teacher performance to be used as both a summative and formative tool. The key areas which are focused on in the survey are: teacher utilisation of pedagogical knowledge, how teachers cater to student strengths and weaknesses, classroom environment, and classroom management (The Colorado Education Initiative, 2013). The tool was formed with the Colorado teacher quality standards in mind. Areas examined within the tool align with these standards (The Colorado Education Initiative, 2013). The instrument was tested with 20,000 students, where reliability ratings for student grades 3-5 is 0.94, and for grades 6-12, 0.96. This suggests a high reliability across the board for this instrument (The Colorado Education Initiative, 2013).

Studies have shown that when a student survey is constructed well, key elements of teacher practice can be captured that reliably correlate with levels of teacher performance (White, 2013). However, there are mixed views about whether the student surveys are a valid instrument for measuring teacher effectiveness, as they may be biased depending on the day's lesson (Clayson, 2009). The Colorado Education Initiative found that the 'Student Perception Survey' results from primary students were more strongly correlated with principal or evaluator

evaluative findings. This suggests that student surveys can be used across grade levels (The Colorado Education Initiative, 2013).

PEER RATING

Peer ratings and reviews involve teachers being evaluated by other teachers. A variety of methods and tools may be used in the peer review process, including, but not limited to, observation, surveys and interviews. There are a number of peer review models, which can range from an informal support network within a school, to a comprehensive and mandated framework of teacher appraisal.

Several US States and Districts have adopted Peer Assisted Review (PAR) programs (Goldstein, 2007; Harvard Graduate School of Education, 2016) . These programs typically involve teachers who are failing to meet the education standards being paired with a consulting teacher who provides them with ongoing feedback and support, in addition to conducting more formal observations. A case study of a Californian school district that employed PAR found the program improved teaching practices (Goldstein, 2007). Teachers who did not meet expected standards on an evaluation were matched with a mentor from outside their school, qualified in both their subject area and year level, who observed a lesson each week and provided detailed feedback and support. Goldstein (2007) claimed that provision of a mentor teacher (with a case load not more than ten teachers) who is familiar with the teacher's educational context was instrumental to this particular program's success. However, many peer review programs do not have these conditions, and while there is some research into the efficacy of specific programs or models, there have been few examinations of the use of peer review more generally. This lack of empirical evidence, combined with the range of procedures and protocols, make it difficult to comment on the reliability and validity of peer review in the context of teacher appraisal.

Peer ratings have face-validity, in that it appears to be an appropriate way to assess and support teachers and is supported by a number of educational scholars (Goldstein, 2007; Kılıç, 2016; Wilkins, Shin, & Ainsworth, 2009). However, no formal measures of validity were identified in the present review. The high degree of variability in how peer reviews are conducted makes it difficult to discuss its reliability as a tool for teacher evaluation. However, as with other methods of teacher appraisal, reliability will be increased if evaluators are trained, given clear assessment criteria and follow standardised procedures.

A case study on an Australian school that adopted a mandatory peer review process highlighted several issues which made it difficult to implement and maintain the program (Brix, Grainger, & Hill, 2014). Most saliently, interviews with key stakeholders (teachers, teacher's union representative and the principal) showed that the peer review process increased teacher's workload. This is because they had to participate in reciprocal mentoring and review procedures, in addition to their full time teaching. Some participants believed that the program had a negative impact on teacher effectiveness, by stopping them from completing more important tasks. However, more comprehensive programs that employ mentor teachers on a full time basis are significantly more expensive. Papay, Johnson, Fiarman, Munger, and Qazilbash (2009) reported that programs typically cost between $4,000 to $7,000 per participating teacher, and California ended its PAR program in 2002 due to the high cost of the program, despite believing that it was an effective way to improve teaching (Goldstein, 2007).

There is a high potential for bias in peer reviews that occur internally within the school, as the teachers typically know each other well. While this connectedness may be a benefit in some circumstances (i.e., they will understand the particular barriers well and be able to offer specific strategies) the potential for bias means that internal peer-review should not be used in high stakes summative assessments (Goldstein, 2007). Some peer-review programs have been able to overcome this by using teachers from outside the school, or by using peer review in formative evaluations and having other tools for summative appraisals.

The time needed to implement peers assisted programs varies depending on the framework. Some peer review models will require that teachers engage in regular observations, meetings and evaluations, whereas others may simply be a survey asking teachers to rate their colleagues. Most formal PAR programs used in the US require that the "coach" or "consulting" teacher participate in some training. However, these teachers are already considered to

be 'experts', and have been selected because they have consistently demonstrated excellent teaching. In circumstances where teachers within the school act as informal evaluators for formative purposes training is uncommon. Peer reviewers can take on a range of roles within both a formative and summative evaluation system. Teachers may be more responsive to the opinions of their peers and colleagues. It is also suggested that peer reviews facilitate an environment for teacher collaboration. It has also been noted that there is potential for bias and the cost effectiveness of frameworks may be burdensome.

PARENT RATING

Parent ratings tend to be survey-based and look at teacher performance and quality of practice. Historically, there is little evidence within the literature regarding the use of parent ratings in teacher evaluation. Most of the research accessed as part of this review was based out of the USA or Taiwan. Nonetheless, more recently they have been considered a potential evidence-base for measuring teacher performance. Parent ratings of teacher performance may gain more attention in the future especially considering the evidence that parental involvement in their child's learning is positively associated with student achievement and behavioural gains. Parent ratings may also be used to inform the assessment of teacher quality more broadly through gaining different perspective and may highlight the level of teacher parental and family involvement (Master, 2013).

A criticism of parent surveys as a source of evaluative evidence is that it provides information that is similar to, but less accurate than, the information students themselves can provide about teachers' impacts on student learning. Parent surveys do not present information that is different from student surveys and feedback, which suggests that it is not the most useful source of information, especially if surveying students is a possibility. In the absence of reliable student surveys, when students are from a young age group, parent surveys can prove a reasonable substitute (Master, 2013). There is not enough research on parent rating of teacher performance to determine the reliability of parent teacher ratings. Nonetheless, Masters (2013) did find consistency in parent ratings of the same teacher over multiple years providing some evidence of inter-rater reliability. In a small longitudinal study of 28 teachers they found that parent ratings of specific teachers remained stable across two years with a correlation of 0.41 (Master, 2013). However, more research needs to be conducted on the sustainability of parent ratings as a method of teacher evaluative assessment. One of the main criticisms of parent ratings of teachers is the potential for bias. As most parents are emotionally invested in their child's schooling, their ability to remain objective in the rating process could be limited. Similar to student surveys, parent surveys require little time to administer and information on teacher performance can be gained in one collection (Master, 2013). Aside from the information parents may provide, parents constitute a key stakeholder group for schools and the act of including their perspectives in evaluation systems may by itself help to engender positive parent-school relations. Parent surveys can be beneficial as a tool for providing feedback on student perspective and as a way of understanding the teacher's relationship with the families of the students. There has been more of a focus towards the student experience in teaching in recent years, having parent perspective on teacher efficacy could prove an important adjunct to the student experience. Parents may also be able to share information on student learning at home as it relates to observations of classroom learnt curriculum and rate of homework (Bi, 2012). Parents can draw from what they see and discuss with their child, as well as from interactions with their child's teacher, to inform their perspective on teacher effectiveness. Thus, they may possess unique information about teachers' overall competency (Master, 2013). Bi (2012) argues that evidence gained from teacher surveys are not as valuable compared to student surveys considering data is gained predominantly from a second-hand source as parents tend to gauge teacher effectiveness from their child's perspective. Furthermore, parents provide little insight into teacher efficacy due to a limited knowledge of quality teaching practice.

CLASSROOM ARTEFACTS

Classroom artefacts are documents created during the teaching practice that can be collected by an assessor for evaluation and analysis, or collated by the teacher; these documents are often added to portfolios. These documents, such as lesson plans, teacher created assignments, assessment results or student work, are often evaluated using a rubric to guide the process.

123

Teacher evaluation efforts have traditionally focused on school inputs and assessment evidence of student achievement. However, the practice is changing to include an evaluative focus on the content and delivery of teaching in the classroom. The shift in focus is due to the evidence suggesting the importance of teacher practice and student learning experiences affecting achievement outputs (Martínez et al., 2012).

The use of classroom artefacts in teacher evaluations can be time efficient as the evidence-base for assessment, such as student work, has already been formed. However, there may be more effort required in order to place sufficient time and effort into detailing past work, through annotation and reflection pieces, in order for teachers to provide a sufficient breadth of evidence on teacher performance (Martínez et al., 2012). In cases where there are accepted and consistent standards for measuring teacher effectiveness, results on classroom artefacts have been found to align with other common measures of teacher performance (Goe et al., 2008a).

EXAMPLES OF CLASSROOM ARTEFACTS TOOLS

SCOOP NOTEBOOK AND RATING GUIDE

The Rating Guide was initially developed as an alternative approach to traditional classroom observation based teacher evaluation methodologies. Using specific set criteria, the tool uses teaching practice artefacts to accurately identify the teacher's understanding of teaching practice. Information collected for the notebook, is obtained over the course of a week and includes collecting lesson plans, student work, photographs of classroom set-up and learning materials. Upon collection, this material is categorised into 3 specific subject areas: preparatory materials, in-class materials and outside of class student prepared materials (Goe et al., 2008a). Judgments of the teacher's understanding of teaching practice within the artefacts is based on criteria that is set out in a rubric (Borko, Stecher, & Kuffner, 2007).

INTELLECTUAL DEMAND ASSIGNMENT PROTOCOL (IDAP)

The subject-specific IDAP aims to enable the recognition of quality teaching in the subjects of English, Science and Mathematics. Class material from both students and teachers are measured against a rubric which specifies levels of quality assessment material (Goe et al., 2008a).

INSTRUCTIONAL QUALITY ASSESSMENT (IQA)

The Instructional Quality Assessment (IQA) created by the National Center for Research on Evaluation, Standards and Student Testing (CRESST) at the University of California in Los Angeles, is a subject specific tool that looks to measure teacher performance as it relates to mathematics and English subject areas. Specifically, the IQA tool provides a rating criteria based on extensive research that can be used for identifying key aspects of teaching practice when conducting teacher observation, as well as providing a base for judging teacher prepared assignments (Matsumura et al., 2006).

Coupled with teaching practice observation, teachers were asked to share examples of student work and the direction and feedback that they gave to each student. The evaluation areas of the tool include the quality of in-class activities and assignments and communicating achievement expectations to students. An additional component of this tool is the ability for it to be used in providing feedback for teacher improvement.

The IQA tool provides a valid summative assessment of teacher performance. Matsumura et al. (2006) conducted a study on the relationship between the IQA and student achievement with 34 teachers in American Middle Schools. They found that the IQA in reading comprehension could predict the achievement scores for students in that class. The tool was found to have an inter-rater reliability rating of 0.77 for English Language Arts and 0.76 for mathematics (Matsumura et al., 2006). However, Goe et al. (2008a) suggests that more research needs to be conducted to determine the tools sustainability. Matsumura et al (2006) found that the dependability coefficient of

classroom artefacts form determining teacher performance quality was relatively high with a dependability coefficient of 0.80- 0.82.

QUALITY ASSESSMENT IN SCIENCE (QAS) NOTEBOOK

The QAS Notebook is an artefact-focused instrument developed to collect information about the classroom assessment practices of science teachers. The instrument was specifically designed to gauge teacher effectiveness through artefact analysis.

The instrument was established to look at teaching material as it relates to assessment and student achievement in the science subject area (Martínez et al., 2012). There is potential for bias as the teacher may wish to share specific artefacts that may not be representative of the teacher's general practice.

HIGHLIGHTS OF BEST PRACTICE

The following crosswalk maps the frameworks, measurement processes and uses of evaluation for each of the countries. The crosswalk represents an attempt to highlight best practice through the use of evidence and rigorous search methods.   A judgement is made whether the evidence is present or not; whether or not what is present is comprehensive; and whether the evidence signals that those methods are recommended or mandated. The judgements were made on available evidence and the quality and depth of the evidence presented as opposed to a specific opinion about the nature of the approach and theories presented. This crosswalk should provide a signal about what is available as well as the gaps in the current literature and resources available. Many of the frameworks and theories presented, while providing interesting and in some cases valuable insights, do not highlight what might work in other contexts, hence it may be more appropriate to suggest that there are examples of research, initiatives and approaches that may be adaptable to an Australian context. It is clear that no one country has built a whole system that is an exemplar and that may be adaptable to any other context. There is little evidence that evaluation data is used for improvement across the sample other than where it is mandated. The variability across systems, the lack of use of rigorous methods and the poor use of evaluation data is a concern for education systems globally.

125

Table 14 - Traffic Light Evaluation (Theory and Frameworks)

		AUS	NZ	ENG	SCO	ON	BC	AB	GER	AT	SNG	HK	SK	CAL	WS	VIR	D.C.
STANDARDS AND THEORY	Standards exist	Green	Green	Yellow	Green	Green	Yellow	Green	Yellow	Yellow	Green	Red	Red	Green	Green	Green	Green
	Standards vary for different career levels	Green	Yellow	Red	Green	Grey	Grey	Yellow	Red	Red	Green	Red	Red	Red	Red	Green	Grey
	A recognised theoretical framework	Green	Yellow	Green	Green	Yellow	Grey	Grey	Grey	Grey	Green	Grey	Grey	Grey	Grey	Green	Grey
TEACHER PREPARATION	Accreditation of ITE	Green	Green	Green	Green	Yellow	Yellow	Yellow	Grey	Yellow	Red	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow
	Registration process	Green	Green	Green	Yellow	Yellow	Green	Yellow	Red	Red	Red	Red	Red	Red	Green	Green	Yellow
TEACHER REGISTRATION	Requirements for ongoing registration	Green	Green	Grey	Grey	Grey	Red	Red	Red	Red	Red	Grey	Red	Yellow	Yellow	Green	Grey
EVALUATION FRAMEWORK	Evaluation framework in place	Yellow*	Yellow	Yellow*	Red	Yellow	Red	Red	Red	Red	Green	Yellow*	Yellow	Red	Yellow	Yellow	Green
	Summative Evaluation	Yellow*	Yellow	Yellow	Yellow	Yellow	Yellow	Grey	Grey	Grey	Green	Yellow	Green	Yellow*	Yellow	Yellow*	Green
	Formative Evaluation	Yellow*	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Grey	Grey	Green	Yellow	Green	Yellow*	Yellow	Yellow*	Yellow
	Evaluation timeline	Yellow*	Yellow	Yellow	Red	Yellow	Yellow	Red	Red	Red	Green	Red	Yellow	Red	Yellow	Yellow*	Grey
	Evaluation methods/tools outlined	Yellow*	Yellow*	Yellow*	Yellow*	Yellow	Red	Yellow*	Grey	Grey	Green	Yellow*	Yellow	Yellow*	Green*	Yellow*	Green

Coding Key: Not identified in reviewed material (Grey) — Not present (Red) — Present (Yellow) — Present and Comprehensive (Green)

Note. * Formal recommendations provided rather than prescribed by governing body

126

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Table 15 Traffic Light Evaluation (Methods and Methodology).

		AUS	NZ	ENG	SCO	ON	BC	AB	GER	AT	SNG	HK	SK	CAL	WS	VIR	D.C.
METHODS AND SPECIFIC EVALUATION TOOLS	Inspectorate	Red	Green	Green	Green	Red	Red	Red	Red	Red	Green	Red	Green	Red	Red	Red	Red
	Classroom observations	Yellow	Yellow	Yellow	Yellow	Green	Yellow	Yellow	Grey	Grey	Green	Yellow	Yellow	Yellow	Green	Yellow	Green
	Teacher interviews	Yellow	Yellow	Yellow	Grey	Yellow	Grey	Grey	Grey	Grey	Green	Yellow	Yellow	Yellow	Yellow	Grey	Red
	Teacher portfolios	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Grey	Grey	Yellow	Red
	Teacher self-rating	Yellow	Yellow	Grey	Yellow	Yellow	Yellow	Grey	Grey	Grey	Green	Yellow	Yellow	Yellow	Yellow	Yellow	Red
	Student rating	Yellow	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Yellow	Yellow	Yellow	Grey	Yellow	Red
	Peer rating	Yellow	Grey	Yellow	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Yellow	Yellow	Yellow	Yellow	Grey	Red
	Parent rating	Yellow	Grey	Grey	Grey	Yellow	Grey	Grey	Grey	Grey	Grey	Yellow	Yellow	Grey	Grey	Grey	Red
	Classroom artefacts	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Yellow	Red
	Student achievement	Yellow	Grey	Yellow	Yellow	Yellow	Grey	Yellow	Grey	Grey	Green	Yellow	Yellow	Yellow	Green	Yellow	Green
	Subject knowledge	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey
	Pedagogical knowledge	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey
	Measures of student affect	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Grey	Grey	Grey	Grey	Red
	Value-add model	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Grey	Grey	Grey	Yellow	Green
	Student growth percentile	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Grey	Yellow	Grey	Grey
USE OF EVALUATIONS	Professional development	Yellow	Yellow	Yellow	Yellow	Yellow	Grey	Yellow	Grey	Grey	Green	Yellow	Yellow	Yellow	Yellow	Grey	Green
	Career progression	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Grey	Grey	Grey	Green	Yellow	Yellow	Grey	Grey	Grey	Green

Coding Key

Not present in reviewed material	Grey	Not used	Red	Recommended	Yellow	Mandated	Green

127

Teacher Effectiveness Systems, Frameworks and Measures: A Review

There are multiple models, many theories and perceptions of teacher effectiveness. The review of the nominated country's evaluation frameworks, methods and theories of teacher effectiveness measures, provides a high level view of the specific content dimensions currently in use. Through the synthesis numerous factors associated with teacher effectiveness emerged, and they were often related to a specific evidence based framework or theory (e.g., the Danielson's model of teacher effectiveness). In some cases, the factors related a specific measure that related to research or evaluation tools, and often the factors emerged from professional standards and appraisal frameworks. In an effort to corral these factors into meaningful and manageable areas, the factors have been collapsed into higher order dimensions. It should be noted that these dimensions emerge from the sample of countries and states reviewed. As suggested there are numerous variables that relate or claim to relate to teacher effectiveness or define teacher effectiveness. Consequently, it was important to be systematic about selection of the variables. A non-exhaustive, though extensive, list of factors associated with teaching and teacher effectiveness is provided in Appendix 1. Understanding the evidence base underpinning these variables is beyond the scope of this piece of work. The research team as a conceptual sorting exercise completed the reduction of the variables into dimensions.

The results of this exercise are depicted below in Figure 3. It should also be noted in the majority of the reviewed countries and states many variables or dimensions are treated as simple lists of measures, and causal or relational models are rarely articulated (with the exception of theories developed from evidence e.g., Danielson). Similarly, standards based descriptions of teacher effectiveness tend to illustrate a number of characteristics that are in some way interrelated. These standards based models are often not operationalised into specific behavioural indicators (e.g. Australian Professional Standards for Teaching), and weightings for an aggregated picture are rarely provided. The model posits that teacher effectiveness relates to the impact and influence the teacher has had on the education community. While impact on the learning lives of students is core to this, this review points to the premise that teacher effectiveness, in relation to impact and influence, is multi-dimensional and that there are a number of factors that relate causally to impact and influence. The measurement of teacher and teaching effectiveness rarely relies on one dimension.
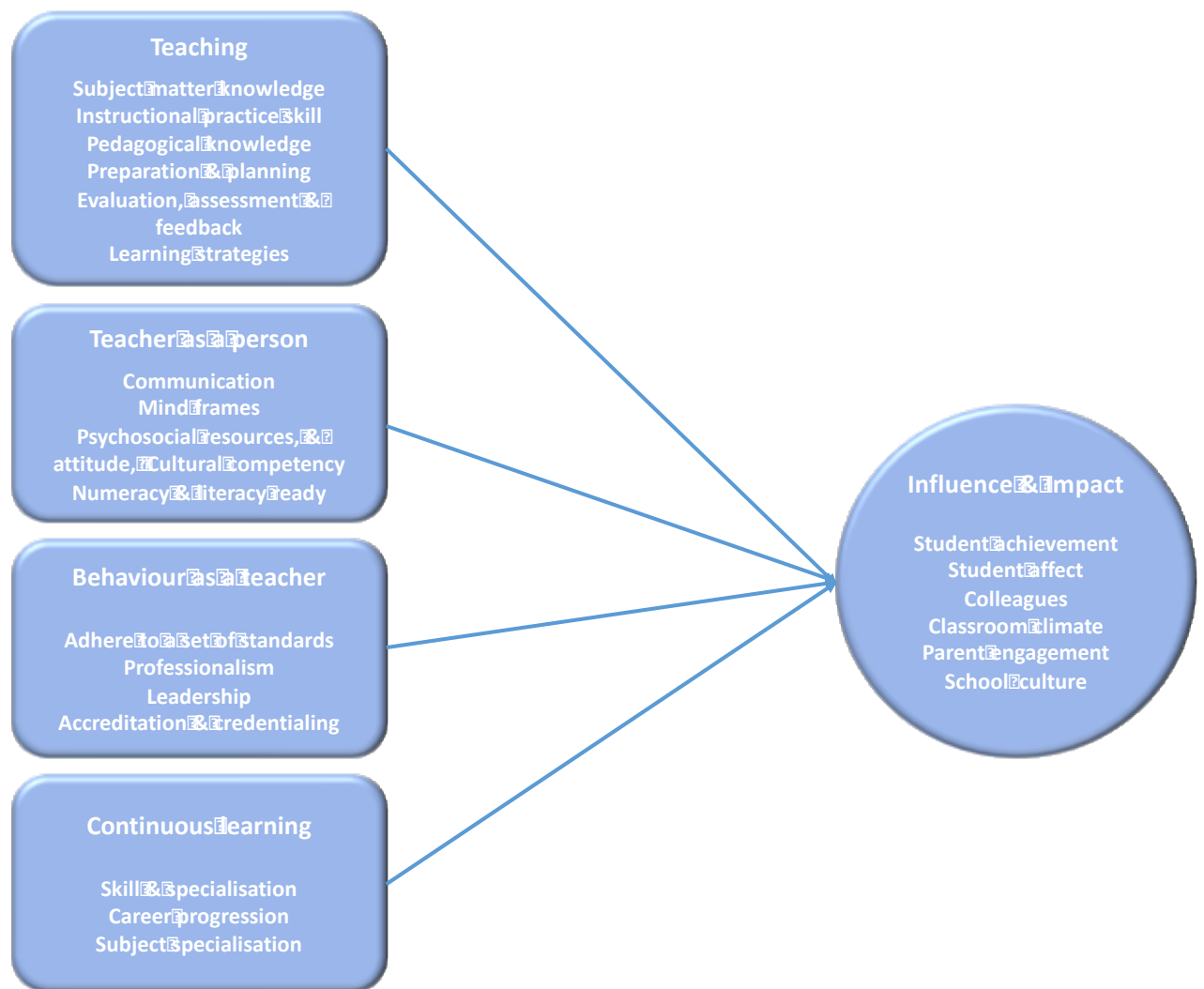
Figure 3 Dimensions of effective teaching.

TEACHERS INFLUENCE & IMPACT

In any model that evaluates the qualities of teachers and teaching there is a necessity to ask about the consequential impact on students. A major claim throughout this paper is that the criterion of success of any evaluation of teachers and teaching needs to be framed in terms of this impact. It is not satisfactory to build a model of desirable teaching methods and beliefs unless it can be demonstrated that they have major impacts on students. This impact, however, is not linear nor necessarily the same when related to the many forms of impacts on students. A major claim is that we need to understand how each teacher answers these questions about what they consider impact, what they consider and how they would know the nature and magnitude of their impact, and which and how many students have this desired magnitude of impact (nature, magnitude, and equity of impact). This nature of impact asks about not only impact on student achievement, but also on their attitude and motivations for learning, on how they can optimally learn (particularly when they do not know), on their feelings of belongingness and willingness to engage in learning, on their dispositions (resilience, respect for self, respect for others), and on their proficiency to become their own teachers. The magnitude must be more than increased in knowledge and understanding, as Hattie (2009) has demonstrated that almost every teacher can increase these attributes – instead we need to ensure the impact of teachers is such that students gain more than a year's growth for a year's input (which begs a robust debate about what a year's growth is). In most classes there are students who make this yearly gain (perhaps with little influence from the teacher), hence the equity issue of how many students are making this gain.

129

Teacher Effectiveness Systems, Frameworks and Measures: A Review

STUDENT FACTORS

Achievement is an omnibus term and can relate to the various dimensions of knowledge (math, literacy, history, arts, etc.) and to the teaching of various strategies for learning this content (i.e., learning strategies). Affect is much broader and can relate to the willingness and resilience to engage in school learning, to the motivations to want to engage in learning, and to the desire to feel that they "belong" in the classroom of learners.

STUDENT ACHIEVEMENT

Hattie (2009) demonstrated that the quality of teachers in terms of their impact on students is the largest source of variance we have any control over. This work is based on synthesising over 800 meta-analyses involving over 100 influences – from the home, family, school, curricula, teacher, and teaching methods. The main idea is that teachers need to know their impact. This begs at least three questions. The first is "What do teachers collectively believe is "impact" in this class, school (or system)". This impact needs to relate to increased achievement, higher test scores, a love of learning, a desire to reinvest in learning, a sense of mastery, the joys of collaboration, social sensitivity to work with others, respect for self, and respect for others – or all of these. The claim was that each teacher should not have a different conception of what impact means – as this creates much of the large variance between teachers in their impact on students. The second question is "What is the expected magnitude of this impact?" Saying "I have evidence that I have increased student's achievement and engagement" is hardly sufficient when over 95 per cent of influences in students can lead to an increase. It is critical, he argued, to have a collective understanding of the desired level of this increase. Hence, it is necessary for schools to collectively have a deeper understanding of what a year's growth looks like relating to the outcomes desired. The third question is the equity one, "How many students in this class are gaining this desired magnitude of growth?"

There are many examples where there is a sole reliance on student academic achievement, such as measured by standardised test results, as an indicator of teacher effectiveness within teacher evaluation models. But this overemphasis can be considered narrow, as there are other critical aspects of learning that teachers need to have a marked positive impact on. A broader notion is to ask about the impact of teachers on the learning lives of students – and this has become a key factor in understanding teacher effectiveness. This focus is becoming more critical than how a teacher teaches, or whether they use the right (evidence based) methods.

The over-use of student achievement or student test scores to assess a teacher is considered one of the controversial approaches in teacher evaluation. In some teacher appraisal systems student achievement is considered a legitimate procedure for evaluating teachers, however this assumes that student learning is directly reflected by tests of academic achievement, which is not always the case. Consider a group of students who start well above the average achievement levels and maintain this above average after a year with the teacher – yes they have high achievement but is this a function of what this teacher has accomplished? Further, gains may or may not be a function of a particular teacher (e.g., increases in literacy in high schools could be a function of the history teacher). There is also concern about the validity and reliability of student test scores or growth measures for evaluating teachers, especially when the results are used to make high stakes decisions (e.g. promotion, termination and pay).

One main concern of the use of student achievement for assessing teachers is whether teachers are the sole determinant of learning outcomes for a particular student. Joshua, Joshua, and Maliki (2007) highlighted that there are various internal and external factors, including a student's family background, peer group, summer learning loss and school attendance that can also contribute to student achievement. As such, student test scores are generally not recommended for use as a measure to evaluate teaching performance.

There are also concerns about how using student results to evaluate teachers can give rise to unintended consequences such as undesirable shifts in curricular focus and poor teaching practices. Lavery, Holloway-Libell, Amrein-Beardsley, Pivovarova, and Hahs-Vaughn (n.d.) explained in their systematic review that the three most common unintended consequences of using student test scores for teacher evaluation were: (a) the likelihood of

unwanted narrowing of the curriculum or syllabus to focus on material directly related to student test scores, (b) development of a disincentive to teach the highest need/weakest students and (c) encouraging cheating or immoral practices amongst teachers. They investigated the use of standardised test scores to evaluate teachers using a framework that includes the interpretation/use argument. Within the 97 peer-reviewed studies included in their systematic review, the overall validity evidence did not fully support the interpretation/use argument of using student scores derived from large scale standardised test to evaluate teachers.

One common way of measuring student outcomes is through the administration of classroom-based tests, often created by teachers themselves. Such tests might have the potential to assist teachers in estimating student's mastery of concepts and content taught in class. This form of measure may also provide a rudimentary indication of the effectiveness of teaching practices. In contrast, poorly designed tests can lead to inaccurate measurement of students' progress and provide false information regarding students' learning outcomes, which can have a great impact on teachers if used for evaluation. It is also important that the test used for pre and post measures are calibrated for difficulty on the same measurement scale – a task beyond most teachers. As such, using teacher made tests may not be an accurate representation of student outcomes.

For example, in Nigerian secondary schools, classroom based tests are extensively used for continuous assessment. However, these tests have been criticised for lack of proper psychometric test properties. Agu, Onyekuba, and Anyichie (2013) developed and validated a Test Construction Skill Inventory (TCSI) to assess teachers' competencies in constructing classroom based tests. The researchers developed an instrument of 30 items that was administered to 543 secondary school teachers. Upon performing factorial analysis, they found that 25 items were valid and reliable with an alpha coefficient of 0.73 for the 25 valid items. Agu et al. (2013) concluded that the TCSI tool might be used as a guide for teachers to construct tests that can give dependable estimates of student performances.

Holler, Gareis, Martin, Clouser, and Miller (2008) found that teacher made assessments were not always adequate measures of student learning. Holler et al. (2008) noted that teacher made tests are often not thorough, with some items being included or constructed out of convenience rather than for the intended purpose of assessing student's knowledge of a particular topic or subject. Additionally, in general, not all learning objectives are measurable on paper and pencil tests, but rather through classroom activities (debates, projects etc.) to assess specific learning outcomes in students. For these reasons, Holler et al. (2008) advise that teacher made tests may not accurately reflect student mastery.

TEACHERS CONTRIBUTION TO STUDENT ACHIEVEMENT

While there are multiple ways of determining student learning gains, there are also a number of methods or tools to determine a teachers' contribution to student learning outcomes. When considering the measure of effective teaching in relation to student achievement it is important to consider these tools.

VALUE-ADDED MODELLING

Value Added Modelling (VAM) is a statistical approach to teacher evaluation that uses student scores on standardised tests in order to isolate and examine the contribution, or value added, that a teacher provides each year. Repeated measurements of student test scores from previous and current school years, as well as the scores of other students in the same grade, are compared in order to formulate models of expected growth for each student. Actual changes in test results are compared against expected results to determine whether the influence of a particular teacher (or school in some cases) has resulted in higher or lower than expected test results.

Papay (2011) examined Value Added estimates from three reading achievement tests: the state test, Scholastic Reading Inventory (SRI), and Stanford Achievement Test (SAT) on teacher performance. He concluded that the correlations between teacher Value-Added estimates range from 0.15 to 0.58 across a wide range of model specifications. Although these correlations are moderately high, these assessments produce substantially different answers about individual teacher performance and do not rank individual teachers consistently. Even using the same

131

test but varying the timing of the baseline and outcome measure introduces a great deal of instability to teacher rankings (Papay, 2011).

Policy makers and researchers also doubt the use of value added measures (that depend on student results) because VAMs can fluctuate over time and be unevenly distributed (Stronge, Ward, & Grant, 2011). Results may also be biased if students are not randomly assigned to teachers. Haertel (2013) discussed the validity and reliability of inferences of a number of studies that tested the VAM estimates. One of the studies included was the teacher VAM reliability as reported from the Measures of Effective Teaching (MET) project. A 2010 MET report gives correlations between VAM scores for 2 successive years, as well as correlations between estimates obtained for two different sections of the same course, taught the same year to different students. On 4 different tests, 2 in mathematics and 2 in English language arts, the correlations between sections taught in the same year range from .18 to .38, and the correlations across years, available for just one of the mathematics tests and one of the reading tests, are .40 and .20, respectively (the MET Project, 2010, p. 18). These numbers are well under .5, which means that over three-quarters of the variation in teachers' single year value-added estimates is random, unexplained, or unstable (Haertel, 2013).

Darling-Hammond et al. (2012) mentioned three criticisms about VAM models as accurate measures of teacher's effectiveness. First, value-added models of teacher effectiveness are inconsistent – researchers have found that teachers' value-added scores differ significantly when different tests are used, even when these are within the same content area. They noted that teacher effectiveness also varied significantly when different statistical methods are used. Second, the students they are assigned affect teachers' value-added performance. Teachers are advantaged or disadvantaged based on the students they teach and this seems unreasonable. Third, value-added ratings cannot disentangle the many influences on student progress.

The American Educational Research Association (AERA) released a statement in 2015 highlighting the scientific and technical limitations of the use of value-added models in evaluating educators. The use of VAM to evaluate teacher continues to be a subject of discussion and debate as research evidence on the accuracy, reliability, stability and the validity and consequences of the use of such indicators in evaluation systems is still accumulating. Some limitations addressed included that student scores on standardised tests vary in the degree to which they fully capture target constructs as well as the level of precision across the range of reported scores. Current state tests, by federal requirement, measure only grade-level standards without including items needed to measure growth for students who perform well below or well above grade level. Therefore, caution about the psychometric quality of the assessment should be exercised if VAM are being considered for purposes of teacher evaluations. Another criticism is that many existing VAM estimates have not been shown to isolate sufficiently the effectiveness of teachers, principals, or other nonteaching professional staff.

Survey based reports indicated that a variety of factors (student interactions with teachers and other students, parental influence, etc.) that influence student's classroom assignments are not directly accounted for in typical VAM estimates. It is argued that the purposeful (non-random) assignment of students into classrooms biases value-added estimates and their valid interpretations (Paufler & Amrein-Beardsley, 2014).

Darling-Hammond (2015) highlighted that in the United States, the No Child Left Behind Act mandated that state tests only measure grade-level standards although these tests do not include items that assess content of skills from earlier or later grade levels. As a result, these state tests do not measure the actual achievement level of the large group of students who are below or above grade-level in their knowledge and skills. Haertel (2013) also noted this feature of test design translates into bias against teachers working with the lowest and highest performing classes.

Following the implementation of the Race to the Top (RTTT) grant program, the US Department of Education reviewed State and district evaluation systems to ensure that it included rating categories which take into account student growth as a significant factor in teacher evaluation. Ballou and Springer (2015) drew attention to some problems in the design and implementation of such evaluation systems that incorporate value added measures. The problems discussed included:

132

(1) Due to measurement error in teacher assessments, teacher value-added estimates can be notoriously imprecise. The magnitude of error in these estimates must be accounted for so that decisions based on VAM scores are fair to teachers.

(2) Revising teachers' scores as more information about students become available. Depending on the teachers' grade level, revision of teachers' estimates can be based on less information on student performance. For example, standardised tests in some states begin in grade 3 and end in grade 8 (and VAM scores are calculated based on a 5-year window). Thus grade 3 teachers would have insufficient data to have an accurate VAM score.

(3) Opportunistic behaviours by teachers during roster verification and supervision of exams. Roster verification raises concern that teachers might fail to claim students that they fear or think will lower their value-added score. For the same reason, teachers may have incentives to cheat by coaching students while monitoring them during an exam.

The authors conclude that evaluation systems that rely on student test scores to measure teacher performance are only as good as the data available and that states/school that wish to implement these systems must address the problems if it is to become a permanent part of the K-12 education. A policy analysis by Newton, Darling-Hammond, Haertel, and Thomas (2010) highlighted that value added models that fail to take student demographics into account can impact the teacher rankings. A teacher who teaches less advantaged students would typically receive lower effectiveness ratings than the same teacher teaching more advantaged students.

Using VAMs for individual teacher evaluation is based on the belief that measured achievement gains for a specific teacher's students reflect that teacher's "effectiveness." However, this attribution assumes that student learning is measured well by a given test, is influenced by the teacher alone, and is independent from the growth of classmates and other aspects of the classroom context. None of these assumptions are well supported by current evidence. Most importantly, research reveals that gains in student achievement are influenced by much more than any individual teacher (Darling-Hammond et al., 2012).

STUDENT GROWTH PERCENTILES

A student growth percentile (SGP) describes a student's growth compared to other students with similar beginning test scores (their academic peers). Although the calculations for SGPs are complex, they are a useful method for understanding student progress against peers. For example, if a student scored 400 on a pre-test, then her score at the post-test would be compared with the scores of all other students who scored 400 on the pre-test. If her Growth Percentile is 50 that means her growth is in the middle of all students who started with 400: half who scored 400 scored higher in the post-test, and half scored lower. A major advantage of growth percentiles is that it allows fairer comparison for students at any level, across different ages, school years, and so on. In relation to teachers, SGPs provide an indication of the magnitude of a teacher's effectiveness in student growth terms by demonstrating the teacher's effect on each student's relative position in the distribution of growth.

STUDENT PROGRESS TESTING

Progress testing is a method that monitors the change in students' knowledge and understanding over the course of the program. It started in the USA and the Netherlands in medicine, and now is widely used in medicine, engineering and sciences. In these domains, progress testing is typically constrained to the knowledge component whereas in teacher education it would be important to also include the experiences and learning when in schools. As an example in medicine, the McMaster program uses a bank of 2500 questions that test knowledge in the basic and clinical sciences from the whole course (overall years). All students sit a stratified (by content) random sample of 180 questions towards the beginning, middle and end of each academic year. Students get a report indicating their performance relative to their cohort, and the information is used to provide remediation or re-teaching (Finucane, Flannery, Keane, & Norman, 2010).

133

Teacher Effectiveness Systems, Frameworks and Measures: A Review

The system provides value added effect scores for an institution (not an individual teacher candidate), and these are calculated for the aggregate of first and second year teachers prepared by the program, then compared across programs and with experienced teachers. They then classify the programs into five bands: Level 1, where the average effect estimate is above the mean for experienced teachers; Level 2, above the mean for new teachers; Level 3, within a standard error of the mean for new teachers; Level 4, below the mean for new teachers; and Level 5 well below the mean for new teachers. They have since closed two programs that scored at level 5 for two years, and asked Level 4 programs to change their programs.

There have been criticisms of progress modelling. For example, too often the measures in the model are too narrow; favour only content knowledge (Jordan, 2012); and they do not include beliefs, knowledge, and practice base (Grossman, Wineburg, & Woolworth, 2001). There is no reason, however, why measures of teacher adaptations to diversity, changes in personal philosophies, and other non-achievement evidence could not be modelled over time to evaluate changes and also to relate back to teacher education experiences. In defence of such criticisms, Fleener and Exner (2011) noted that the Louisiana value added model aimed to link student performance to teachers and *the programs* that prepare them. They wanted to control for the many variables that contribute to, and distract from a student's ability to learn. It is also possible to add many moderators and mediators to a progress model such that critical questions can be asked about their influence. These could include: demographics of the student, nature of the course, the experiences and timing of school placement, the experiences and timing of developing content knowledge and other course attributes, the nature of the schools where students practice, the attributes of the staff (in university and in schools) and the attributes of the school where candidates are employed after graduation.

STUDENT AFFECT

The value of student outcomes aside from academic achievement is becoming increasingly important in evaluating educational, and teacher effectiveness (OECD, 2013). As noted by the OECD: "Student achievement that comes at the expense of well-being is not a full accomplishment" (OECD, 2015). While there are countries that score above average on both academic outcomes and student happiness (e.g. Belgium, Hong Kong China, Japan, Liechtenstein, Shanghai-China, Singapore, Switzerland and Chinese Taipei), countries that score high on current PISA measures of mathematics and reading do not necessarily score well on measures of student happiness. For example Korea, which regularly scores within the top achieving education systems in the world (PISA), had the lowest student happiness ratings of all countries. Across the countries examined by the OECD (2015), students who reported that they had a positive relationship with their teacher were more likely to report that they were happy at school, make friends easily, felt like they belonged at school, and were generally satisfied with their school. As such, positive teacher-student relationships appear to have an impact on the happiness of children at school. Hattie & Clinton (2008) have also noted a strong effect size (.72) between teacher-student relationships and academic achievement.

Further, student self-efficacy (Alivernini & Lucidi, 2011; Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011; Komarraju & Nadler, 2013), subjective well-being (Suldo, Thalji, & Ferron, 2011), and student motivation (Alivernini & Lucidi, 2011) have all shown to be positively related to academic achievement. Student motivation, which has been linked to student perceptions of teacher autonomy support, has also been linked to a student's intention to stay or drop out of school, as well as subsequent academic achievement (Alivernini & Lucidi, 2011). However, despite these associations, there is almost no evidence that demonstrates that these domains are a regular part of the evaluation of teacher quality or effectiveness.

It may be that as education systems are only just beginning to understand the impact of these factors on learning and achievement as well as establishing reliable measures, it is too early to include them in the boundaries of measuring teacher effectiveness. Measures of student affect, such as the Attitudes to School Survey (ATSS) in Australia, have been implemented to understand morale and school connectedness (Victorian Auditor-General, 2010). This survey features questions examining student opinions of statements on a five-point (or seven for some domains) Likert-scale. For example, in examining morale, students are asked to respond to statements such as: "I feel relaxed at school". Scales such as the ATSS, while initially intended as a tool for measuring the impact of well-being programs in schools, may possibly be adapted for teacher evaluation purposes in the future.

134

Furthermore, the introduction of the Australian Curriculum includes student character development across three domains: self-awareness, self-management, and social awareness. With these factors now embedded across all learning areas, in future they may become a measurement domain in the evaluation of teacher effectiveness. While there is a shift towards including student affect as a key indicator of educational success, evaluation models largely do not include these measures in the context of evaluating teachers.

TEACHERS INFLUENCE & IMPACT ON THE SCHOOL AND COMMUNITY

There is little doubt that teachers influence and impact more than students. Evidence suggests that teachers play an integral part in education generally and also influence the whole school culture. As Danielson observes,

> *"Teachers can find a wealth of opportunities to extend their influence beyond their own classrooms to their teaching teams, schools, and districts. Many teachers have a vision that extends beyond their own classrooms—even beyond their own teams or departments. Such teachers recognise that students' school experiences depend not only on interaction with individual teachers, but also on the complex systems in place throughout the school and district. This awareness prompts these teachers to want to influence change" (Danielson, 2007, p. 3).*

This influence can be on their colleagues, on the climate of the school, on their profession, on and with the parents, and on the development of their curricula knowledge. Research relating to teacher's impact beyond students is variable in abundance and quality.

Many attributes of teacher professional standards expect teachers to engage in these out-of-class influences, so it seems reasonable therefore to consider them part of any evaluation of a teacher's effectiveness. Not surprisingly, it is difficult to find claims how specifically these influences directly impact on student learning, how they can be reliably measured, and tend to be considered more in terms of the depth and frequency of involvement. For example, Kraft and Rogers (2015) examined the potential of teacher-to-parent communication, the study's findings indicate that messages from teachers to parents containing actionable information about how a student can improve increases students' academic success. Further, Kraft and Rogers noted: "for participating students, these course credits could be the difference between being on-track or off-track to graduate from high school" (p. 60). There is strong evidence that school culture influences a teacher's satisfaction with their career, and there is strong evidence now to suggest that teachers who are leaders have a potential to reciprocally influence school culture.

TEACHER AND TEACHING FACTORS RELATED TO INFLUENCE & IMPACT

As is noted by Figure 2 there are several factors that contribute to a teachers' impact and influence that are considered in many approaches to understanding and measuring teacher effectiveness. In this case the current review considered the teacher's skill and knowledge in the profession of teaching, teacher dispositions and their behaviour in relation to the profession. The following section attempts to provide a brief summary of the factors that review of teacher effectiveness policies and measures uncovered that relate to teacher's influence and impact. These summaries do not constitute a literature review. Rather, they provide a signal that these are key concepts to explore in future initiatives.

Most pre-service teachers are required to pass a subject knowledge test upon completion of their teaching course in order to graduate or attain accreditation. One country that has this practice is the United States (e.g., EdTPA, Praxis). These tests are usually literacy (and numeracy) tests. In Australia, one such test is the Literacy and Numeracy Test for students in Initial Teacher Education programs. It is designed to test initial education students' personal literacy and numeracy skills to ensure that they are well equipped to meet the demands of teaching (Australian Council for Educational Research, 2016). There has been a long argument about the importance of more subject matter knowledge, pedagogical content knowledge, and other variants of the term to evaluate the breadth and depth of knowing about the curriculum teachers are expected to teach.

Tchoshanov (2011) examined whether teacher's content knowledge is associated with student achievement on a sample of 102 mathematics teachers in the United States. These teachers were tested using the Teacher Content Knowledge Survey (TCKS). TCKS consists of 33 multiple-choice items that measure teachers' content knowledge based on three cognitive types: Knowledge of Facts and Procedures; Knowledge of concepts and connections; and Knowledge of Models and Generalisations. The results of the TCKS Survey were measured against the standardised passing rates for approximately 2,400 middle grade students. It was found that the correlation between the cognitive type of teacher content knowledge and student achievement (student's passing rates) was not significant for knowledge of facts and procedures (Pearson's $r = .06$) and knowledge of models and generalisations ($r = .02$). However, knowledge of concepts and connections was significant with an effect size of .26.

Metzler and Woessmann (2012) estimated the effect of teacher subject knowledge on student achievement based on Peruvian 6th grade data for over 12,000 students in 900 primary schools. They tested both students and teachers in two subjects (mathematics and reading), and found that teacher subject knowledge has a significant impact on student achievement. After correcting for measurement error, their results suggest that a one standard deviation increase in teacher test scores raises student test scores by about 10 percent of a standard deviation. This means that if a student is moved, for example, from a teacher at the 5th percentile of the distribution of teacher subject knowledge to a teacher at the median of this distribution, the student's achievement would increase by 16.6 percent of a standard deviation by the end of the school year. They concluded that teacher subject knowledge is a relevant observable factor that is part of overall teacher quality. However, one of the few studies that showed a link between PCK and student learning is also among the more robust and impressive. Baumert et al. (2010) based multilevel structural modeling on 194 classes. They reported that student content knowledge was not related to the teachers' choice of cognitive levels of tasks and individual learning, and it was only the curricula level of the tasks that increased with increasing levels of CK but it has no direct impact either "on the potential for cognitive activation or on the individual learning support that teachers are able to provide when learning difficulties occur." Both CK and PCK vary independently of effective classroom management. They claimed that the estimate of PCK was higher than CK, but the difference in the fit of the two models was non-significant.

The Australian Council for Educational Research (ACER) developed the Teacher Education Mathematics Test (TEMT). It is being used for selection of students into courses that do not have Year 12 Maths as a prerequisite or as a basic skills test for entrance into such courses. TEMT has both formative and summative capabilities, designed not only to test pre-service teachers' mathematical knowledge, but also to help identify their errors and misconceptions. Ryan and McCrae (2006) reported the outcomes of (TEMT) on one sample of 426 pre-service teachers in Australia and a second sample of 86 pre-service teachers in England. The reliability of the item estimates (item difficulty estimates with mean set at zero) was found to be α = .97 while the reliability of case estimates (student ability estimates) was α = .88.

Metzler and Woessmann (2010) noted that the existing evidence on the association between teacher knowledge and student achievement suffers from bias due to unobserved student characteristics, omitted school and teacher variables, and non-random sorting and selection into classrooms and schools. Exemplars of such bias include motivated teachers inspiring greater student learning but also accruing more subject knowledge; parents with high expectations for educational achievement choosing schools or teachers of higher subject knowledge and also further

their children's learning in other ways; and where principals place students with higher learning gains into classrooms of teachers with higher knowledge.

The evidence from the qualitative literature is similarly not supportive of high relations between PCK and student achievement outcomes. Kennedy (2008) reviewed various studies and concluded that "the relationship between teacher qualification and teaching quality – a relationship that should be self-evident – is not self-evident at all" (p. 345). She used 20 terms relating to teachers' educational background, and 20 terms for evidence of teaching quality to locate research articles. From the 450 studies located, she excluded those not based on real classrooms (e.g., simulations), those that did not establish a link between the qualifications and the quality of practice (e.g., only described both and asserted that they related) - and this led to final corpus of 23 studies. Again, a scarce pool of such a critical topic, pronounced on by so many. Her conclusion is hardly positive news for those who claim that the more nuanced qualitative studies are more likely to demonstrate higher relations. Kennedy concluded "that the qualitative literature agrees with quantitative literature in its inability to distinguish between teachers with different types of certificates or different teacher education backgrounds" (p. 363).

Kennedy noted how this lack of relationship has been most studied in mathematics, and noted that have some have claimed this is because teacher tests are not tapping the content knowledge most relevant to teaching (see below). She reviewed 11 qualitative studies and the message from each is that the teacher(s) tended to have lower levels of deep than surface knowledge, such that they were a) more likely to use questions that (tacitly) discouraged student discussion, b) did not know how to transform his deeper content knowledge into meaningful learning activities c) there was evidence of teachers having missing knowledge, not discrete factors but more the "structural relations among concepts", d) when teachers lack understanding of student misconceptions then their students learned less, e) teachers with greater knowledge "of relations among topics did spontaneously make integrative connections and provided real-life examples during their classroom lessons" (p. 355). It needs to be noted that Kennedy had much difficulty with the quality of these studies, and noted that there was too little control of internal and external validity in the studies. Her overall message was that "overall, this work has been disappointing in its failure to demonstrate the expected relationships" (p. 350). Perhaps, she concluded the very question of how knowledge influences practices is ill conceived, particularly when it assumes knowledge is a relatively fixed entity that can be called on at will and overlooks the role of goals, beliefs, and spontaneous interpretations of events in determining what teachers do at any given moment.

She also reviewed the qualitative studies on the effects of teacher preparation and the conclusion was that there was little discernible effect of these programs on the practices and outcomes of teachers. When students from exemplary programs were contrasted with other graduates, the former had more confidence in their ability to teach, more likely to mention specific features of their programs that they valued, and more likely to base instructional decisions on student needs rather than curricular mandates. But "none of the studies of teacher education found a clear, visible influence on teaching practice" – possibly she suggests is that most programs want to alter the things teachers do (e.g., adopt a particular approach) and most of these require skills often beyond novices (e.g., spontaneous mid-lesson adjustments).

Lederman and Newsome (1992) reviewed studies involving direct observations of classroom practice and found half provided evidence supporting the assumption that a teacher's subject matter knowledge or structures influenced their instructional approaches and half did not. Lederman concluded that "there appears to be little evidence supporting the rather intuitive notion that science teachers' subject matter knowledge or structures directly influence classroom practice" (p. 16). Indeed, if only studies that involve certified science teachers, then the evidence supports "the view that a strong connection between presentation of subject matter and the teacher's subject matter structure does not exist" (p. 16).

KNOWLEDGE OF HOW STUDENTS LEARN:

The current review could find no studies of teacher effectiveness that included teachers' knowledge of research into how students learn. This could be claimed to be an essential form of knowledge that then leads to decisions about

137

selecting teaching strategies and designing learning programs that respond to the diverse physical, social and intellectual learning needs of students. There has been many studies relating to the effectiveness of various strategies for learning, and most indicate that there is much skill needed by teachers to ensure students have multiple learning strategies, and have strategic considerations about when and where to optimise the choice of these strategies (Hattie & Donoghue, 2016).

INSTRUCTIONAL PRACTICE

Instructional practices are a set of teacher created events embedded in purposeful activities that facilitate learning in order to achieve specific objectives (Gagné, Wagner, Golas, & Keller, 2005). Specifically, they can be dialogic and organisational approaches used by a teacher to create a learning environment and to specify the nature of the activities in which the teacher and learner will be involved during the lesson. During the course of the lesson all teachers will use their professional judgement to adapt their instructional practices to the specific context they work in and to the particular cohort of students they teach. It should be noted that the emphasis of instructional practice is on what the teachers are doing, not what students are doing in the classroom or what outcomes the students achieve as a consequence of that instruction. When observed or recorded, instructional practice is the demonstration of how teachers use what they know in the classroom to achieve learning objectives. However, each teacher has a zone of discretion in which there are common practices to draw upon to facilitate learning. A key focus of professional development has been the increase in the knowledge base of the classroom instructional practices in order to improve teachers' impact on learning outcomes. At best, this choice is a correlate on the student learning, and thus lists of effects of various teaching methods speak to the probabilities (not certainties) of impact on students.

PREPAREDNESS AND PLANNING

Initial teacher education providers are required to provide robust evidence that graduates from their programs are well prepared with the subject knowledge and teaching practices required for them to be considered *classroom ready* (TEMAG, 2014b). Preparedness and planning extends beyond curriculum and assessment decisions or resourcing classroom lessons to encompass classroom organisation, transitions and routines and the management of classroom behaviour.

EVALUATION, ASSESSMENT & FEEDBACK

The typical claim is that teachers need to know how to develop, use and interpret appropriate assessment schemes (while teaching and via tests and activities) and know how to evaluate and give and receive feedback to enhance student learning. In most cases there is much research to defend these notions, although the details and priorities are hotly contested. For example, as indicated above, the evidence about the level of content knowledge is not clear (and indeed many meta-analyses show it is a very low effect; Hattie, 2009), there is much debate about *which* teaching methodologies are optimal (and for whom and when in the learning cycle); and the value of assessments is widely debated. What does not appear to be in doubt is that high levels of the above factors, combined, are important. Perhaps, flexibility in choosing between multiple strategies of teaching and assessment is more critical; perhaps more attention to how students receive feedback than how much is given is more critical; perhaps it is how teachers see information from tests as feedback to then modify their instruction is more critical.

EXPERTISE/QUALIFICATIONS REQUIRED

It has been demonstrated that teachers impact on student achievement, and contemporary research suggests that the differences in achievement gains for students who had the most qualified teachers versus those who had the least qualified were greater than the influences of race and parent education combined (Darling-Hammond, 2010). Training, credentialing and registration for newly qualified teachers and importantly teachers progressing through the profession is described by every country and state some degree. There is tremendous variation in the regulatory

138

nature within and across countries. This is similar for quality and the enforcement of the regulations. There is very little evidence to suggest that regulation adds value. The NBPTS has some moderately positive evidence suggesting credentialing has an impact on teacher effectiveness and student outcomes (Ingvarson & Hattie, 2008). Darling-Hammond recently established that those teachers who continue to establish their experience and credentials as lead teachers have greater impact on student learning.

THE TEACHER AS A PERSON

There are many non-cognitive attributes of teachers that have been related to impacts on students. These include their social sensitivity to collaborate with others, their openness to experience, their conscientiousness, their passion, and their motivations to teach. Like the choice of teaching methods, these are more correlates of impact.

MIND FRAMES

There has been a long search for teacher attributes, behaviours, and practices that make the greater impact on student learning. The process-product research has aimed to relate what teachers do to outcomes leading to claims about the power of various teaching methods, teaching strategies, and teaching environments that have the highest correlations with student outcomes. This has been vigorously critiqued as confusing correlates with causation, which has led to continual claims for support for almost everything (given a positive correlation between most influences and outcomes), and there is a plethora of models all claiming an evidence base). An alternative view is that it is more how teachers think, evaluate, and make moment-by-moment decisions that matter, particularly when these decisions are based on using high probability teaching methods (derived from much of the process-product research). Hattie (2009, 2012) for example, claims that there are ten major mind frames or ways of thinking that distinguish high from low impact teachers. These are based on an interpretation of over 1200 meta-analyses relating various influences to student outcomes.

The specific mind frames that teachers have about their role are important—and most critically a mind frame where they ask themselves about the effect they are having on student learning. Among the more critical mind frames is for teachers to see themselves as evaluators of their impact on students. Teachers need to use evidence-based methods to inform, change, and sustain these evaluation beliefs about their effect. The focus on impact leads to three questions: 1) What do teachers consider is appropriate impact? 2) What is the magnitude of this impact? 3) The equity question – How many students benefit from this high level of impact? It is teachers' belief that they are evaluators, change agents, adaptive learning experts, seekers of feedback about their impact, engaged in dialogue and challenge, developers of trust with all, see opportunity in error, and are keen to spread the message about the power, fun, and impact we have on learning. Hattie (2013) outlined and defined 10 critical mind frames:

1. I am an evaluator and Know thy Impact (I engage in Diagnosis, Interventions and Evaluation)
2. I talk about learning
3. I enjoy challenge
4. I see learning as hard work
5. I engage in positive relationships and develop high levels of trust
6. I engage in dialogue and not monologue (I value student voice about my impact)
7. I use the language of learning and see errors as opportunities for learning
8. I am a change agent and value growth (I cause learning)
9. I see assessment as feedback to me
10. I collaborate

TEACHER AS A COMMUNICATOR

In the classroom, clear and often declarative communication is vital for efficient transfer of information. Teaching is a human activity that incorporates the need for efficient information transfer, as well as respect for the humanity and capabilities of those being taught – making the task of communicating even more complex (Bowles, 2010;

139

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Munro, 2008; Webster, 2010). Classroom communication may involve the provision of instruction and feedback to students, the use of questioning and dialogic teaching methods, building rapport and relationships, and conflict resolution. Non-verbal factors include eye contact, proximity and gesturing. A teacher's style of communication plays a key role in the teaching-learning process (Sallinen-Kuparinen, 1992). The teacher as a communicator extends far beyond the classroom, as interaction with parents and colleagues forms a major part of the role. At the subject level, professional learning communities (PLCs) collaborate in environments of trust and respect to analyse and improve classroom practice and student outcomes. At the school level, teachers function in environments that demand leadership, conflict resolution, and collaboration (Bowles, Hattie, Dinham, Scull, & Clinton, 2014).

Early research into teachers as communicators operated on the assumption that learning would not occur without effective communication (Olson, 1969). More recently, Hattie (2009) ranked teacher clarity in the top ten influences on student achievement. Hattie (2009) cites a meta-analysis by Fendick (1990) suggesting that clarity of speech is a major component of overall teacher effectiveness. Research into non-verbal communication reveals that higher levels of teacher non-verbal immediacy (eye contact, proximity, touch, expression, and tone of voice) are moderately correlated with student attitudes and perceptions of learning, with a weaker relationship to cognitive measures of learning (Witt, Wheeless, & Allen, 2004). Students develop strong impressions of their teachers based on their communication style, with teachers who adopt an open, relaxed, and friendly style being perceived as more effective by their students (Andersen, Norton, & Nussbaum, 1981). Others have argued that teachers who engage in dialogic discussions in their class are more effective than those who are more monologic; although the evidence is less clear about under what conditions dialogical teachers are more effective.

Teacher communication with parents is related to improved academic outcomes. Students whose parents received weekly feedback on mathematics lessons outperformed a control group receiving no parental feedback, with the effects particularly strong for lower-achieving students (Sirvani, 2007). Despite this, over half of parents in an American sample reported "never or rarely" communicating with their child's teacher (Patrikakou & Weissberg, 2000).

There is evidence to suggest that teachers who collaborate and communicate about classroom practice and student outcomes through formally organised professional learning communities (PLCs) can improve teaching effectiveness. (Graham, 2007) found that activities that form part of a high-quality PLC meeting can lead to significant improvements in teacher effectiveness. A major mediator of this relationship was the nature of the conversations and communication in the PLC activities, and the development of a sense of community across PLCs.

PSYCHO-SOCIAL RESOURCES

There is growing recognition that non-cognitive teacher factors are an important piece of the teacher effectiveness puzzle. These factors include a teacher's attitudes, sense of self-efficacy, and levels of motivation (Borghans, Duckworth, Heckman, & Ter Weel, 2008; Lindqvist, 2011; Lindqvist & Vestman, 2011). Teacher attitudes may relate to the teacher's feelings towards teaching, or expectations of their students. Teacher self-efficacy relates to a teacher's confidence in their ability to promote students' learning (Hoy, 2000). Teacher motivation involves a personal level of desire to teach and focuses on psychological satisfaction derived during the act of teaching (Reeve, 2014). Motivation has effects on a teacher's enthusiasm and their job satisfaction. Teacher self-efficacy illustrates a strong association with learning satisfaction. A model developed by Tai, Hu, Wang, and Chen (2012) illustrated a strong relationship between teacher self-efficacy with learning satisfaction and learning outcomes.

A teacher's expectations of their students can be one of the most powerful attitudes when considering impact on student achievement. Hattie (2009) suggests that teachers adopt the mantra of "be prepared to be surprised", citing research that teachers seek evidence to confirm their prior expectations of a student's ability rather than focussing on potential progress. In terms of self-efficacy, Hoy (2000) suggests that there is a critical period in the first years of teaching where teacher self-efficacy is developed, related to experiences of mastery during teacher training and graduate years. Individual teacher self-efficacy can be bolstered by a sense of collective self-efficacy at the school level, where teachers are supported to accept challenging goals (Protheroe, 2008). Regarding teacher motivation, it

140

has been suggested that teachers who are intrinsically motivated by learning will pass on this attitude to their students, while those who are not will give students a sense that learning in their class is not rewarding (Csikszentmihalyi, 2014), ultimately affecting student achievement outcomes.

Nearly one-third of teachers in Australia consider leaving the profession during their first five years of employment (Watt & Richardson, 2008) due to stress, burnout, and not being supported by their mentors. Teachers' sense of stress and job satisfaction are linked to their levels of motivation (Barnabé & Burns, 1994) and engagement with the role (Schaufeli & Bakker, 2004), and can ultimately affect student outcomes. It follows that an effective teacher must be well equipped to deal with the high levels of stress, and wavering levels of self-efficacy and job satisfaction associated with the profession. Strong coping skills, such as maintaining a work-life balance, planning and prioritising work tasks, and cognitive restructuring techniques (Kyriacou, 2001) are an important psychological resource for teachers.

Teacher-student relationships have a strong effect on student academic outcomes. (Cornelius-White, 2007) found that teacher-student relationship variables such as empathy, warmth, and encouragement of learning were associated with improved student achievement and attitudinal outcomes. The research also found that students taught by person-centred teachers (where the teacher has high levels of non-directivity, empathy, genuineness, etc.) displayed fewer resistant behaviours. In terms of teacher personality, the two factors consistently shown to relate to higher performance, particularly in jobs requiring social interaction such as teaching, are extraversion and conscientiousness. A number of studies involving teachers support this relationship for teaching (Ayres, Dinham, & Sawyer, 1999; Rockoff, Jacob, Kane, & Staiger, 2011).

CULTURAL COMPETENCY

The increasing diversity apparent in today's classrooms is a reflection of the rapidly changing demographics of Western countries (Moltó, Florian, Rouse, & Stough, 2010). Classroom teachers are relied on to provide and maintain safe and supportive environments that promote student achievement and wellbeing (Hachfeld, Hahn, Schroeder, Anders, & Stanat, 2011). Evidence suggests that beginning teachers who exhibit low cultural sensitivity or who have not received adequate pre-service training (Castro, 2010) run the risk of developing false expectations or hidden biases towards their students (Reiter & David, 2011) as their careers progress.

Cultural competence involves more than an understanding of differences between cultures. It requires an awareness of one's own world view, a development of positive attitudes towards cultures other than one's own, the accumulation of knowledge about the cultural practices and world views of others, and the development of skills in cross-cultural communication and interaction (Department of Education Employment and Workplace Relations, 2010). Culturally competent teachers are able to respect and celebrate the benefits of diversity, and this has a positive impact on the classroom climate they create. Beginning teachers require training to ensure they possess culturally responsible beliefs, attitudes and accordingly, conduct themselves in a pro-social manner both within diverse classrooms and in the wider school context (Harrington & Hathaway, 1995). Teachers who possess cultural sensitivity are more likely to be able to recognise and prevent conflict in the classroom and in turn promote cultural sensitivity in their students. Likewise, a teacher's ability to foster positive classroom environments is likely to build self-acceptance and freedom to focus on learning unencumbered by prejudice from others.

Teachers with sound cultural awareness are able to enrich the learning process for all students by embracing and using cultural differences in their teaching practices (Hachfeld et al., 2011). Similarly, teachers with strong self-awareness of their own beliefs are better equipped to recognise and manage lack of acceptance in their students. These teachers are also more likely to conduct themselves in a pro-social manner, which is stemmed in their belief that people are equal and deserve to be treated equally (Hachfeld et al., 2011). Recent research shows that students of teachers who are assessed as being culturally competent display more inclusiveness in their friendship patterns than teachers who have not been assessed as culturally competent (Thompson, 2010). This finding illustrates the real-world impact of cultural competence on student social outcomes.

141

NUMERACY & LITERACY READY

In examining the literacy skills of over 200 pre-service teachers in Australia, Moon (2014) found that many of the students tested displayed literacy levels below a standard that would allow them to perform the typical tasks required of a teacher. This finding reinforces earlier work by Scriven (1987), who found that up to 40% of pre-service teachers in a Western Australian university failed high school level literacy testing. In contrast, pre-service teachers typically believe that they possess adequate literacy and numeracy skills; just 5% of pre-service teachers admit that their skills are not adequate to teach (Milton, Rohl, & House, 2007). However, at present there is little evidence on the relationship between teacher literacy and numeracy levels, and subsequent impact on student achievement.

The recently introduced Literacy and Numeracy Test for Initial Teacher Education Students signals a push by the Australian Government to ensure that all pre-service teachers are numeracy and literacy ready. The test was implemented following a report published by the Teacher Education Ministerial Advisory Group (2014a), where a key recommendation was that *'higher education providers use the national literacy and numeracy test to demonstrate that all pre-service teachers are within the top 30 per cent of the population in personal literacy and numeracy'*. Hattie (2009) concludes that teachers are among the most powerful influences on student learning, and many have taken this to mean that competency factors such as a teacher's literacy and numeracy levels play an important part in enhancing achievement.

Both NSW and Victorian Ministers for Education recently mandated that all entrants to pre-service teacher education programs must pass the literacy and numeracy test in order to gain registration as a teacher.

TEACHER BEHAVIOUR

While being engaged in the education process and possessing knowledge and skill is deemed critical, teachers across the globe are required to act in a particular way. All countries have a set of regulations that teachers need to adhere to.

ADHERE TO A SET OF STANDARDS

Systems of performance management, increased demand for public accountability and regulation by central government have been the catalyst for the rise of education policies focused on professional teaching standards (Sachs, 2003). This model has dominated government discourse globally for the last thirty years, driven by economic views of education, performativity and accountability (Menter, Hulme, Elliot, & Lewin, 2010).

Accountability frameworks in the form of professional standards that explicitly articulate the elements of high quality teaching have been built into initial teacher preparation and licensure requirements (Darling-Hammond, 2012) in many OECD countries. Standards for teachers and leaders are ideally developed consultatively and provide clear advice at progressive career stages about the indicative knowledge, skills and engagement expected at each level. Teachers must adhere to these standards and provide periodic and documented evidence against each domain and dimension for performance review and promotion purposes. Valid teaching standards and methods for evaluating teachers' performance are essential elements for career structures and incentives that reward quality teaching and retain quality teachers (Kleinhenz & Ingvarson, 2007).

PROFESSIONALISM

Aspects of teacher professionalism permeate the domains and dimensions of professional standards frameworks and measures of teacher effectiveness. Teachers are required to demonstrate increasingly advanced or sophisticated levels of professionalism across these components as they progress during initial preparation, induction, and licensure and across career stages.

142

Based on the 2013 Teaching and Learning International Survey, the OECD report *Supporting Teacher Professionalism* defines the components of teacher professionalism as a knowledge base necessary for teaching; autonomy in decision making over aspects related to teachers' work, and information exchange and support within peer networks (OECD, 2016b). Countries differ in the aspects of teacher professionalism they place most emphasis upon and this is reflected in the incentives offered such as increased professional development, curriculum autonomy or cultivating professional learning networks.

TALIS found a positive correlation between student learning and the teacher professionalism index. Direct links were evidenced between teacher professionalism and teachers' self-efficacy, confidence, work satisfaction and perceptions of the status of the profession (Jensen, 2010; OECD, 2016b). The implications are clear for education systems and policy makers. Professionalisation practices that support early transition, induction and mentoring, teacher professional development and teacher collaboration promote high quality teaching and retention of the teaching workforce.

PROFESSIONAL ETHICS AND CONDUCT

Teachers are required to demonstrate high standards of professional conduct and ethical behaviour, undertaking this responsibility within the framework of the law (NSW DET, 2016). Teachers hold positions of special trust in safeguarding children and adolescents and are accountable for their actions. National and statutory regulators develop codes of conduct and ethics as public statements that inspire the highest professional behaviour and standards. These codes are generally framed by guiding principles of integrity, respect and responsibility (Victorian Institute of Teaching, 2016b) and used in conjunction with local or school based guidelines and policies. They codify the professional and personal conduct, professional competence and ethical behaviour expected of teachers as a guide for interacting with colleagues, students, parents or caregivers, liaising with the public and being responsible and committed members of the profession. If these codes are considered to have been breached protocols are in place to consider each case in accordance with principals of natural justice and procedural fairness. Disciplinary action may be taken, ranging from written admonishment to reduction in career stage or possible termination and deregistration.

TEACHERS AS LEADERS

In the most successful schools, teachers supported by administrators take initiative to improve school-wide policies and programs, teaching, learning, and communication. Teachers as leaders contribute beyond their own school when they participate in school and teacher evaluation or curriculum team, present at a state or national conference, serve on community boards and parent groups, or speak at a school board meeting as the voice of teachers in the community. As noted by Danielson (2007): "these teachers are doing more than teaching their own students (as brilliant as they may be in that work); they are influencing the larger education environment in their communities and perhaps their states" (p. 18).

There are three main domains of school life where involvement of teachers as leaders adds value. 1) Within teacher leaders' own departments or teams. 2) Across the school. 3) Beyond the school into the community and the profession. Many countries and states use the idea of teachers as leaders as often promoted by Danielson and in particular her evaluation model. Muijs and Harris (2003) suggest that teacher leadership could have beneficial effects on school improvement, school and teacher effectiveness and teacher motivation and retention, but that the right conditions need to be in place in order for teacher leadership to flourish.

A recent review of literature demonstrated that teacher leadership as a construct has value, however the quality of research and the difficulty of definitions in particular inclusion of task, renders the construct over encompassing. Wenner and Campbell (2016) suggest that: "teacher leadership, although rarely defined, focused on roles beyond the classroom, supporting the professional learning of peers, influencing policy/decision making, and ultimately targeting student learning" (p. 469). The literature notes a movement towards collective leadership, as the role of

143

the school hence a principal becomes more extensive and complex it is noted that principals alone cannot complete the tasks at hand. Engagement of the whole school in the vision and distribution of the leadership is critical.

TEACHERS CONTINUOUS LEARNING & PROFESSIONAL DEVELOPMENT

Professional learning in most countries is seen as essential to effective teaching and teachers. It is viewed as the desire to engage in further learning but also a regulatory requirement to engage in learning.

CONTINUOUS LEARNING

Effective teachers maintain and renew their knowledge and skills to meet the changing and expanding demands of the profession. National, state and local goals and priorities focus professional learning as a vehicle for school reform and educational excellence. At a time of technological advancement, pedagogical change and great diversity in our classrooms, teachers must be able to reconceptualise their practice in order to prepare students with the knowledge, understanding, skills and values they need for the 21st Century.

Teachers are generally required to demonstrate their commitment to ongoing professional learning by meeting or exceeding prescribed targets for full registration and periodic renewal with regulatory bodies. With recognition that powerful professional learning drives advances in teaching practice and impact on student outcomes, there is increased focus and investment in teacher development. Global trends in professional learning include approaches that are: Integrated, Immersive, Design-led, Market-Led and Open (AITSL, 2016b). Teachers are required to develop their skill and specialisation. In particular, they are often asked to have advanced skills and knowledge in the areas they teach in order to contribute to improved student learning outcomes, confidence and engagement.

Teachers have access to and regularly engage in professional learning opportunities that enable them to acquire and renew professional knowledge and pedagogical skills both generally and in specialised areas. Distinctions in sophistication, with expectations of a growing repertoire of specialised skills relative to the teaching area extending across broader educational contexts, are evident in the career stages of professional standards frameworks. Specialist subject knowledge is now seen as critical to teacher effectiveness despite variable evidence. Effective teachers, both generalist and specialised, have mastery of subject content in the areas they teach. Teacher candidates are required in many contexts to pass subject content tests, typically in the areas of literacy and numeracy, in order to progress in their program or to gain initial teacher registration. Although teacher subject knowledge is often claimed to have a significant impact on student achievement, there are no mandated requirements for renewal of subject knowledge beyond registration or promotion purposes.

Career progression is essential for retention, and demonstrating movement through the level of profession is central to performance appraisal. Career stages are located across the continuum, from graduate teacher to leadership or principal class positions. Career stages are typically represented as benchmarks in professional standards frameworks that represent increasingly sophisticated levels of professional knowledge, practice and engagement across broader and more complex educational contexts. National or state requirements are typically mandated in order to progress from the graduate level. Progression through higher stages is usually via localised demonstrations during performance and development reviews.

Evaluating the performance and impact of teachers is an integral part of ensuring the quality of teaching. This can include evaluating the impact of ongoing professional development, informing staffing decisions, and ultimately determining the quality of teaching that students receive. The evaluation of teachers is typically framed by statutory professional standards that articulate the knowledge, practice and engagement required across teachers' careers, and can vary across countries and contexts. Thus, methods for teacher evaluation should exist within coherent, integrated evaluation systems and accountability constructs. Importantly, comprehensive documentation of the implementation system, valid and reliable assessment tools must support evaluation methods. Purposeful, ongoing training must be provided for evaluators. Evidence that they are worthwhile and enhance the profession must be included. While a suite of evaluation tools can be powerful for transforming praxis (Darling-Hammond, 2013), the quality of currently available tools is varied.

There is agreement that systems for measuring teacher effectiveness are in need of an overhaul (Darling-Hammond, 2012). Reports highlighting the inability of current systems to accurately measure teacher quality or to support development of a highly skilled teacher workforce (Bill & Melinda Gates Foundation, 2010; Toch & Rothman, 2008; U.S. Department of Education, 2011; Weisberg et al., 2009) have sparked recent efforts to develop methods for measuring and improving teacher expertise. The call from the field is for comprehensive evaluation systems that clearly distinguish the features of good teaching and purposefully outline the measures of effective teaching across career-stages and contexts (Darling-Hammond, 2012).

Darling-Hammond (2012) made a distinction between *teacher* quality (personal skills, attributes and dispositions) and *teaching* quality (strong instruction that supports student learning). Darling-Hammond noted influences such as the instructional context, curriculum and assessment systems, class size, facilities and materials, and called upon policymakers to consider the teaching and learning environment alongside the capacity of individual teachers. The researchers add a third dimension, teacher *impact*, whereby the focus is on the consequences of the teacher and teaching on the cognitive and non-cognitive attributes of their students.

The teacher evaluation literature often highlights the trend to identify the attributes of *quality teachers* or *quality teaching,* and then evaluates the degree to which teachers possess or demonstrate these attributes. Considerable variation exists in the use of singular or combined methods and associated tools and the degree to which teacher evaluation methods encourage or allow for local variation in appraising teachers. Popham (2013) posited that a defensible determination about a teacher's quality and performance must be based on quality evidence about the teacher's subject knowledge and instructional skills, and that no single source of evidence is sufficient. Popham acknowledged the diverse settings that teachers work in, finding "no unbending, one size fits all evaluative process will ever work to accurately appraise different teachers in different contexts" (2013, p. 10). Like Darling-Hammond (2013), Popham suggested that multiple methods and sources of evidence are, in aggregate, more reliable and valid. He further stated that teacher evaluators should determine the appropriate evaluative weighting for each piece of evidence or demonstration of practice, and flexibly adjust weightings in consideration of the instructional setting.

There have been several recent attempts to review the teacher evaluation literature. In *Evaluating Teacher Effectiveness,* Darling-Hammond (2010) explored measures of teacher effectiveness as predictors of value-added student achievement. Her report detailed progress towards the development of a continuum of reliable and valid teacher performance measures, from initial teacher education through to recognition of advanced expertise. Standards-based approaches to assessment, such as the Performance Assessment of California Teachers (PACT, now edTPA) for pre-service teachers and National Board Certification for experienced teachers, were outlined in the report. Multiple evaluative methods and sources of evidence of effective teaching practice (such as video recording of practice, classroom observation, performance portfolios with commentary and evidence of student achievement) were detailed along with recent research findings. The report concluded with a synthesis of the main points and recommendations for national policy and an effective national agenda for improvement of teacher quality. Darling-Hammond (2010) also concluded that there was not "a set of widely available methods to support the evaluation and ongoing development of teacher effectiveness throughout the career, along with decisions about entry and

145

continuation in the profession. Meeting the expectation that all students will learn to high standards will require a transformation in the ways in which our education system attracts, prepares, supports, and develops expert teachers who can teach in more powerful ways—a transformation that depends in part on the ways in which these abilities are understood and assessed" (p. 3).

Similarly, the *Measures of Effective Teaching* (MET) project (2009) investigated some of the major measures of teacher effectiveness. This involved a research partnership between academics, teachers and educational organisations dedicated to identifying and studying methods for evaluating teachers and promoting teacher development (Kane et al., 2013). The study provided considerable empirical evidence related to teacher effectiveness. It produced and trialled various evaluative measures (e.g., classroom observation, student rating, etc.), as well as value-added modelling procedures for identifying effective teachers and predicting the contribution an individual teacher has on student gains (Kane et al., 2013). It recommended that:

- Teacher evaluations should be based on multiple measures of performance, including data on student academic progress
- Classroom observations could be improved by making them more frequent and robust
- Evaluators should use or modify an existing observation rubric instead of trying to reinvent the wheel
- Evaluators should be provided with the training and ongoing support they need to be successful
- The use of student surveys as a component of teacher evaluation should be strongly considered

RECOMMENDATIONS

The current review of teacher effectiveness provides insight into previously assessed evaluation tools and their respective regulatory standards, frameworks, reliability and validity. It has been designed to be descriptive and provide a high-level picture of the state of play of teacher evaluation and teacher evaluation measures. It is not meant to be all encompassing. However, the researchers believe it provides a high level guide to future directions to understand teacher evaluation and appropriate measures of teacher effectiveness.

There are a number of factors that need to be considered in implementing an evaluative system to measure teacher effectiveness. Evidence suggests that teacher quality is a multidimensional construct that requires a range of tools and measures to assess the effectiveness. It is apparent from the evidence available that a one-size fits all approach is not applicable. It is especially important to note that there is evidence that policies and practice must relate to context and the specific purpose of the evaluation processes to be useful and feasible. The following provides some high level recommendations in relation to measuring teacher effectiveness.

SYSTEM-LEVEL POLICY SETTINGS AND PREREQUISITES

- Standards must be evidence-based and support career progression
- The evaluation system needs to separate the formative and summative evaluation of quality teaching
- There needs to be a regulatory body in place to articulate and support evaluation, and house the results
- Evaluative judgments must be made against standards that are progressive across career
- Further definitional work around teacher effectiveness and impact is needed, as well as more agreement of agreed definition of teacher effectiveness/ impact

FEATURES OF HIGH QUALITY MEASURES OF TEACHER EFFECTIVENESS

RESOURCES

- Measures need to reflect progressive development
- Measures must be valid, reliable, consistent over time, and unbiased
- Multiple methods of measurement need to be available and utilised simultaneously

146

CONTEXT

- Measures must be adaptable to multiple contexts
- Remote and various communities must be able to utilise measures
- Multiple voices need to be taken into account, for example, student, teachers, and parents

DIMENSION EVALUATION FOCUS ON

- Instructional practice
- Skill and knowledge
- Professionalisation consideration
- Teacher impact on educational community
- Student learning outcomes (cognitive and non-cognitive)

APPLICATION/IMPLEMENTATION OF MEASURES OF TEACHER EFFECTIVENESS

MEASUREMENT ASSUMPTIONS AND PROCEDURES

- Defensible weighting of the various components must take place in relation to individual teachers
- Set of guidelines need to exist for the conglomeration or weighting of measures to make effective judgments
- Measures must operate across different contexts
- Defensible cut-scores for making decisions about high quality

EVALUATOR EXPERTISE

- Expert evaluators are available
- Training for evaluators is accessible
- Capacity building in the form of induction and ongoing support needs to be available for teachers and evaluators

USE OF EVALUATION RESULTS

- Evaluative results must be utilised for educational improvement of school community
- Results need to inform policy and practice within the context being evaluated
- Support decisions are required about professional growth, e.g. professional learning, certification at higher levels of the standards, etc.

EVALUATION PROCEDURES: REMOTE AND INDIGENOUS COMMUNITIES

- Measures must be adaptable to multiple contexts
- Remote and Indigenous communities need to able to utilise measures
- Communities engaged in process and criteria for success

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Quality teachers and teaching lead to successful student outcomes, hence consideration of the evaluation of teachers and teaching is essential. The current Australian system for teacher evaluation is noted by its accreditation systems. While the system at a national level provides strong guidelines for career pathways for teachers and guidance for aspiration for quality teaching as per the standards, there is little agreement about measures or methods for evaluating the levels of quality of teaching across the country. The expansion of APST has gone some way to alleviating this situation and in particular providing support for highly accomplished and lead teachers. While several states have adapted this model, there is still little information and research available relating to the assessment of teachers deemed successful at this level.

There is little doubt that in Australia leadership teams are actively engaging in performance development and utilising the performance development frameworks to ensure teachers have a career pathway. However, this information is all held at a local level and in some cases systematic methods of appraisal is not utilised in any dependable manner (TALIS 2013 results suggest 43 per cent of teachers do not have or use formal feedback about their impact; OECD, 2014c). While this continues to be the case teacher evaluation will remain inconsistent, high cost, and add little value. Consequently, our understanding of quality teaching and our evidence as a nation will remain costly and anecdotal.

The review highlights pockets of excellence; however, no one country or state provides an exemplar of teacher evaluation systems or measures that are applicable to the Australian context. Current Australian evaluation systems appear to add little value to the enhancement of teaching practice despite some recent support to develop career pathways for teachers. There also appears little capacity to evaluate teacher effectiveness at a local, state or national level; hence capacity building in the workforce is needed. There are a number of key features that may be valuable to consider in relation to an Australian teacher evaluation system and measures to understand teacher effectiveness.

Framework Considerations

A successful evaluation approach is a framework that:

- o Ensures an evaluation workforce and encourages evaluation capacity building
- o Ensures standards-based evaluators are identified both internally and externally also at national state and territory and school level
- o Includes and articulates the Australian Professional Standards for Teaching in a transparent way
- o Has solid usability, feasibility, validity, accuracy and accountability-hence adhering to the evaluation standards
- o Involves the whole evaluation community
- o Has transparent procedural systems e.g., defined roles
- o Promotes evaluative thinking
- o Considers teacher effectiveness as multidimensional
- o Takes into account value for money
- o Adopts an evaluation theory approach

System Considerations

An effective evaluation system:

- o Relates to an aspirational process that empowers teachers and celebrates the profession
- o Is evidence based
- o Provides summative evaluation, for example accreditation and credentialing
- o Provides for continuous monitoring and improvement
- o Ensures accountability

148

     o   Encourages life long career development

     o   Is adaptive to context

Measurement and Resource Considerations

Measurements and resources should:

- o Utilise multiple methods
- o Aggregate and weight multiple dimensions
- o Employ multiple iterations of appraisals
- o Ensure multiple uses of data that informs policy and practice and theory
- o Be adapted to teaching standards/experience expectations targets and indicators of success
- o Consider timing of measures and development of teacher skill as well as utilising progressive assessment

Considerations for Remote and Indigenous Communities

Remote and Indigenous regions should be considered such that:

- o Prescribed systems are adaptable
- o Indigenous communities and leaders are engaged
- o Appropriate guidelines are developed for remote and Indigenous areas
- o Resourcing and capacity building is made available

FUTURE RESEARCH AND LIMITATIONS

This summary of international frameworks and measures of teacher effectiveness was achieved through a rapid synthesis of controlled research, grey literature, and policy documents relating to teacher effectiveness and evaluation procedures. Policies, practices, and procedures were summarised and represented in an evaluation crosswalk for reference and clarity. Policies and procedures were then further evaluated using a traffic light system, with categories chosen to enhance understanding with regards to the extent and government level of their implementation.

From this review a list is provided of common dimensions that are utilised, in the measures and frameworks of effective teachers and teaching, by the sample countries. This list is by no means presented to represent a model of effective teaching, it is simply a categorised view of those dimensions utilised.

It is suggested that this review provides grounds for a further systematic review of the dimensions of teacher effectiveness and aligned measures that may be appropriate for Australia. In particular, understanding the adaptation of an evidence base related to remote and Indigenous communities will require a particular focus.

149

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Appendix 1. Non-exhaustive List of Variables Related to Effective Teaching and Teachers

Subject matter knowledge	Instructional practice skill	Pedagogical knowledge
Preparation & planning	Evaluation, assessment & feedback	Learning strategies
Communication	Mind frames	Psychosocial resources
Cultural competency	Numeracy & literacy ready	Adheres to a set of standards
Professionalism	Leadership	Accreditation & credentialing
Skill & specialisation	Career progression	Subject specialisation
Student achievement	Student affect	Colleagues
Classroom climate	Parent engagement	School culture
Philosophy of teaching	Innovative and creative teaching	Mentorship of students and faculty
Participation in teacher development	Scholarship of teaching	Recognition of effective teaching
Creating an optimal classroom climate for learning	Identifies essential representations of their subject,	Can guide learning through classroom interactions
Can monitor learning and provide feedback	Relate what is happening to a deeper sets of principles	Recognizes a sequences of classroom events which affect learning & teaching
Can detect and concentrate more on information that has instructional significance	Can make better predictions based on their representations about the classroom	Can identify a greater store of algorithms
Understand students for problem solving	Predict and determine what types of errors students	Much more responsive to students
Anticipate, plan, and improvise as required by the situation	Adopt a problem-solving stance to their work	Better decision-makers an
Can identify what decisions are important and which are less important decisions	Have a multi-dimensionally complex perception of classroom situations	More context-dependent and have high situation cognition
More adept at monitoring student problems	Assessing their level of understanding and progress	They provide much more relevant, useful feedback
Adept at developing and testing hypotheses about learning difficulties or instructional strategies	High respect for students	Provide appropriate challenging tasks and goals for students
Enhance surface and deep learning	Challenging	Classroom climate
Deep representations	Deep understanding & achievement	Improvisation
Monitors learning & provides Feedback	Multidimensional perception	Passion
Problem solving	Respect	Sensitivity to context
Test Hypotheses	Use of knowledge	Persistence
Organisation and Planning	Values student learning	Theory to Practice
Teach all backgrounds and levels	Approach to students	Survive in Bureaucracy
Fallibility	Equity and High Expectations	Demonstrates Professionalism
Safe, Respectful, and Culturally Sensitive	Responsive Learning Communities	Instructional Plans & Implementation
Has Content Knowledge	Monitoring and Assessment of Progress Reflection	Reflect on practice in collaboration with administrators and colleagues, monitor personal and professional growth
Clear Focus for the Lesson	Offer Overt Instruction	Get Students to Engage with Content
Commitment to Professional Growth	Self-assesses and works to improve classroom practice	Develops and implements a professional growth plan
Seeks out professional development and continuous learning opportunities	Works with colleagues to improve practice throughout the building Commitment to the School Community	Maintains open communication with the entire school community
Assumes appropriate leadership roles	builds a positive school culture Commitment to Professionalism	Maintains a high level of professionalism at all times
Adheres to legal responsibilities and current educational policies	Collaboration, and Personal Growth	

150

REFERENCES

Academy of Singapore Teachers. (2016a). About AST.   Retrieved from http://www.academyofsingaporeteachers.moe.gov.sg/

Academy of Singapore Teachers. (2016b). Teacher Growth Model.   Retrieved from http://www.academyofsingaporeteachers.moe.gov.sg/professional-growth/learning-frameworks/teacher-growth-model

ACARA. (2016). NAPLAN National Reports.   Retrieved from http://www.nap.edu.au/results-and-reports/national-reports#NAPLAN

Agu, N., Onyekuba, C., & Anyichie, A. (2013). Measuring teachers' competencies in constructing classroom-based tests in Nigerian secondary schools: Need for a test construction skill. *Academic Journals, 8*(8), 431-439.

Alberta Education. (2016a). Number of Alberta Schools and Authorities.   Retrieved from https://www.education.alberta.ca/school-infrastructure/schools-and-authorities/

Alberta Education. (2016b). Teacher Certification.   Retrieved from https://education.alberta.ca/teaching-in-alberta-what-you-need-to-know/teacher-certification/

Alivernini, F., & Lucidi, F. (2011). Relationship between social context, self-efficacy, motivation, academic achievement, and intention to drop out of high school: A longitudinal study. *The Journal of Educational Research, 104*(4), 241-252.

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system-secondary. *School Psychology Review, 42*(1), 76.

Andersen, J. F., Norton, R. W., & Nussbaum, J. F. (1981). Three investigations exploring relationships between perceived teacher communication behaviors and student learning. *Communication Education, 30*(4), 377-392.

Attinello, J. R., Lare, D., & Waters, F. (2006). The value of teacher portfolios for evaluation and professional growth. *NASSP Bulletin, 90*(2), 132-152. Retrieved from https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ748003&site=eds-live&scope=site

Australian Council for Educational Research. (2016). Literacy and Numeracy Test for Initial Teacher Education Students. Retrieved from https://teacheredtest.acer.edu.au/

Australian Government. (2016). Australian education system.   Retrieved from https://www.studyinaustralia.gov.au/global/australian-education/education-system

Australian Institute for Teaching and School Leadership. (2013). Australian Professional Standards for Teachers.   Retrieved from http://www.aitsl.edu.au/australian-professional-standards-for-teachers/standards/frequently-asked-questions

Australian Institute for Teaching and School Leadership. (2014). Nationally consistent registration of teachers.   Retrieved from http://www.aitsl.edu.au/registration/nationally-consistent-registration-of-teachers

Australian Institute for Teaching and School Leadership. (2015). Accreditation of initial teacher education programs in Australia. Retrieved from http://www.aitsl.edu.au/docs/default-source/initial-teacher-education-resources/accreditation-of-ite-programs-in-australia.pdf

Australian Institute for Teaching and School Leadership. (2016a). Graduate teachers.   Retrieved from http://www.aitsl.edu.au/australian-professional-standards-for-teachers/standards/career-stage/graduate

Australian Institute for Teaching and School Leadership. (2016b). New Research in Professional Learning and Performance and Development.   Retrieved from http://www.aitsl.edu.au/docs/default-source/media-news-room-resources/articles-resources/edreview_pl_pd_aug2014_web.pdf?sfvrsn=2

151

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Australian Institute for Teaching and School Leadership. (2016c). Reflection on Practice Tool.   Retrieved from http://www.aitsl.edu.au/australian-professional-standards-for-teachers/resources/reflection-on-practice-tool/

Australian Institute for Teaching and School Leadership. (2016d). Self Assessment Tool.   Retrieved from http://www.selfassessment.aitsl.edu.au/

Australian Institute for Teaching and School Leadership. (2016e). Strengthened accreditation of initial teacher education programs.   Retrieved from http://www.aitsl.edu.au/initial-teacher-education/ite-reform/accreditation

Australian Institute for Teaching and School Leadership. (2016f). Teacher Registration in Australia.   Retrieved from http://www.aitsl.edu.au/docs/default-source/default-document-library/teacher_registration_in_australia.pdf?sfvrsn=8

Ayres, P., Dinham, S., & Sawyer, W. (1999). What makes a good HSC teacher? . *Education Network, 16*, 8-15.

Balch, R. T. (2012). *The validation of a student survey on teacher practice.* Vanderbilt University.

Ballou, D., & Springer, M. G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher, 44*(2), 77-86.

Barber, M., & Mourshed, M. (2007). *How the world's best-performing schools systems come out on top*: McKinsey & Company.

Barnabé, C., & Burns, M. (1994). Teachers' job characteristics and motivation. *Educational Research, 36*, 171-185.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y. M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal, 47*(1), 133-180. doi:10.3102/0002831209345157

Beck, R., Livne, N., & Bear, S. (2005). Teachers' self-assessment of the effects of formative and summative electronic portfolios on professional development. *European Journal of Teacher Education, 28*(3), 221-244.

Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment, 17*(2-3), 62-87. doi:10.1080/10627197.2012.715014

Bendig, A. (1952). A preliminary study of the effect of academic level, sex, and course variable on student ratings of psychology instructors. *The Journal of Psychology, 34*, 21-26.

Bi, S.-y. (2012). *Evaluation of teaching effectiveness in Taiwan's EFL context*. Paper presented at the meeting of The Language Training and Testing Center, Twenty-first International Symposium on English Training.

Biggs, J., & Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. New York: Academic Press.

Bill & Melinda Gates Foundation. (2010). Learning about Teaching: Initial Findings from the Measures of Effective Teaching Project.   Retrieved from https://docs.gatesfoundation.org/documents/preliminary-finding-policy-brief.pdf

Bill & Melinda Gates Foundation. (2012). *Gathering Feedback for Teaching: Combining High-Qulaity Observations with Student Surveys and Achievement Gains* Retrieved from http://files.eric.ed.gov/fulltext/ED540960.pdf

Borghans, L., Duckworth, A., Heckman, J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources, 43*(4), 972-1059.

Borko, H., Stecher, B., & Kuffner, K. (2007). *Using Artifacts to Characterize Reform-Oriented Instruction: The Scoop Notebook and Rating Guide. CSE Technical Report 707*. Retrieved from https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED495853&site=eds-live&scope=site

Bowe, J., & Gore, J. (2011). *Quality Teaching Rounds: Facilitating Quality Teaching Rounds*: University of Newcastle.

Bowe, J., Gore, J., & Miller, W. (2010). *Rounding out professional development: Professional learning community, instructional rounds and quality teaching.* Paper presented at the AARE 2010 Conference, Melbourne, Victoria.

Bowles, T. (2010). The association of typologies of communication style with characteristics of relationship conflict. *Australian Journal of Communication, 37*(3), 117-129.

Bowles, T., Hattie, J., Dinham, S., Scull, J., & Clinton, J. (2014). Proposing a comprehensive model for identifying teaching candidates. *The Australian Education Researcher, 41*(4), 365-380.

Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest Region*. Retrieved from

British Columbia Ministry of Education. (2016a). Becoming a Teacher in BC.   Retrieved from https://www.bcteacherregulation.ca/Teacher/BecomingATeacherOverview.aspx

British Columbia Ministry of Education. (2016b). Overview of the Eligibility Requirements.   Retrieved from https://www.bcteacherregulation.ca/Teacher/AppFromBC.aspx

British Columbia Ministry of Education. (2016c). Teacher Education Programs.   Retrieved from https://www.bcteacherregulation.ca/TeacherEducation/TeacherEducationOverview.aspx

British Columbia Ministry of Education. (2016d). Types of certificates.   Retrieved from https://www.bcteacherregulation.ca/Teacher/CertificatesIssued.aspx

British Columbia Public Schools Employers' Association and British Columbia Teachers' Federation. (2011). Provincial Collective Agreement.   Retrieved from http://www.bcpsea.bc.ca/documents/sd-collective-agreements/63-LB-2006-2011%20Working%20Document%20FINAL%20July%2015%202010.pdf

British Columbia Teacher Federation. (2012). 2012 BC Education Facts.   Retrieved from http://www.bctf.ca/uploadedfiles/public/publications/2012edfacts.pdf

British Columbia Teachers Council. (2012). Standards for the Education, Competence and Professional Conduct of Educators in BC.   Retrieved from https://www.bcteacherregulation.ca/Standards/StandardsDevelopment.aspx

Brix, J., Grainger, P., & Hill, A. (2014). Investigating mandatory peer review of teaching in schools. *Australian Journal of Teacher Education, 39*(4), 5.

Bundesministerium für Bildung und Frauen. (2014). Basic Information.   Retrieved from https://www.bmb.gv.at/enfr/school/gen_edu/basic.html

Bundesministerium für Bildung und Frauen. (2015). Welcome to the website for the Austrian Federal Ministry of Education! Retrieved from https://www.bmb.gv.at/enfr/index.html

California Department of Education. (2016). Educator Evaluation Systems.   Retrieved from http://www.cde.ca.gov/pd/ps/evaluationsystems.asp

California State Government. (1971). Education Code Section 44660-44665.   Retrieved from http://www.leginfo.ca.gov/cgi-bin/displaycode?section=edc&group=44001-45000&file=44660-44665

California Teachers Association. (2012). Teacher Evaluation Framework.   Retrieved from http://www.cta.org/~/media/Documents/Issues%20%20Action/Teacher%20Quality/Teacher%20Evaluation%20Framework%204-23-15%20Final%20for%20website%20File.pdf?dmc=1&ts=20160519T1813302611

Cantrell, S., Fullerton, J., Kane, T. J., & Staiger, D. O. (2008). *National board certification and teacher effectiveness: Evidence from a random assignment experiment.* Retrieved from

Caprara, G., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: a longitudinal study. *British Journal of Educational Psychology, 81*, 78-96. doi:DOI: 10.1348/2044-8279.002004

Carnoy, M., Brodziak, I., Luschei, T., Beteille, T., & Loyalka, P. (2009). *Teacher Education and Development Study in Mathematics (TEDS-M)* Retrieved from Amsterdam, the Netherlands: http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/TEDS-M_Cost_Study.pdf

153

Castro, A. J. (2010). Themes in the research on preservice teachers' views of cultural diversity: Implications for researching millennial preservice teachers. *Educational Researcher, 39*, 198-210.

Cavalluzzo, L., Barrow, L., & Henderson, S. (2014). *From Large Urban to Small Rural Schools: An Empirical Study of National Board Certification and Teaching Effectiveness*. Retrieved from http://www.nbpts.org/sites/default/files/Policy/cna_from_large_urban_to_small_rural_schools.pdf

Center on International Education Benchmarking. (2016). Singapore: Teacher and Principal Quality. Retrieved from http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/singapore-overview/singapore-teacher-and-principal-quality/

Centre of Study for Policies and Practices in Education. (2013). *Learning Standards, Teaching Standards and Standards for School Principles: A comparative Study* Retrieved from Chile http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/WKP(2013)14&docLanguage=En

Chingos, M., & Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. . *Economics of Education Review, 30*(3), 449-465.

Clark, D., Martorell, P., & Rockoff, J. (2009). *School Principals and School Performance*. Retrieved from

Clayson, D. E. (2009). Student evaluations of teaching: are they related to what students learn?--a meta-analysis and review of the literature. *Journal of Marketing Education, 31*(1), 16-30. Retrieved from https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ831766&site=eds-live&scope=site

Clemens, J., Palacios, M., Loyer, J., & Fathers, F. (2014). Measuring Choice and Competition in Canadian Education. Retrieved from https://www.fraserinstitute.org/sites/default/files/measuring-choice-and-competition-in-canadian-education.pdf

Clinton, J., Hattie, J., Cairns, K., Dabrowski, A., Abrahams, T., & Brcan, S. (2014). *Visionary feedback: The use of real-time captioning and teaching analytics to make teaching and learning visible.* Paper presented at the 7th International Conference of Education, Research and Innovation (ICERI), Seville, Spain.

Clotfelter, C. T., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682.

Commission on Teacher Credentialing. (2007). Accreditation Framework: Educator Preparation in California.

Commission on Teacher Credentialing. (2009). California Standards for the Teaching Profession. Retrieved from http://www.ctc.ca.gov/educator-prep/standards/CSTP-2009.pdf

Commission on Teacher Credentialing. (2016a). California Teaching Performance Expectations. Retrieved from http://www.ctc.ca.gov/educator-prep/TPA-files/TPEs-adopted-2016.pdf

Commission on Teacher Credentialing. (2016b). Multiple Subject of Sungle Subject Teaching Credential Via Peace Corps Teaching Experience. Retrieved from http://www.ctc.ca.gov/credentials/leaflets/cl535.pdf

Commission on Teacher Credentialing. (2016c). Renew Your Document. Retrieved from http://www.ctc.ca.gov/credentials/renew.html

Commission on Teacher Credentialing. (2016d). Renewal and Reissuance of Credentials. Retrieved from http://www.ctc.ca.gov/credentials/leaflets/cl494.pdf

Commission on Teacher Credentialing. (2016e). Requirements for Teachers Prepared in California. Retrieved from http://www.ctc.ca.gov/credentials/leaflets/cl561c.pdf

Commission on Teacher Credentialing. (2016f). Teaching Performance Assessment (TPA). Retrieved from http://www.ctc.ca.gov/educator-prep/TPA.html

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research, 77*(1), 113-143.

154

Corsi-Bunker, A. (2016). Guide to the Education System in the United States.   Retrieved from
    http://www.isss.umn.edu/publications/USEducation/2.pdf

Cowan, J., & Goldhaber, D. (2015). *National Board Certification: Evidence from Washington State*. Retrieved from Seattle, WA:

Csikszentmihalyi, M. (2014). *Applications of Flow in Human Development and Education: The Collected Works of Mihaly Csikszentmihalyi*. Dordrecht: Springer.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. United States: ASCD.

Danielson, C. (2007). The Many Faces of Leadership. *Teachers As Leaders, 65*(1), 14-19.

Danielson, C. (2011). *Enhancing professional practice: A framework for teaching* (2nd ed.). United States: ASCD.

Danielson, C. (2012). Observing classroom practice. *Educational Leadership, 70*(3), 32-37.

Danielson, C. (2013a). The Danielson Framework for Teaching. Retrieved from http://sbeacares.org/files/2014/06/Danielson-Yellow.pdf

Danielson, C. (2013b). Rubrics from the Framework for Teaching: Evaluation Instrument.   Retrieved from
    http://www.ode.state.or.us/wma/teachlearn/commoncore/danielson-2013-rubric-only.pdf

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *education policy analysis archives, 8*(1), 1-44.

Darling-Hammond, L. (2010). Evaluating Teacher Effectiveness.   Retrieved from http://files.eric.ed.gov/fulltext/ED535859.pdf

Darling-Hammond, L. (2013). *Getting Teacher Evaluation Right: What Really Matters for Effectiveness and Improvement*: Teachers College Press.

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher, 44*(2), 132-137. doi:10.3102/0013189x15575346

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan, 93*(6), 8-15. Retrieved from
    https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ985081&site=eds-live&scope=site

Delannoy, F., McKenzie, P., Wolter, S., & van der Ree, B. (2003). Attracting, Developing and Retaining Effective Teachers. Country Note: Austria.   Retrieved from https://www.oecd.org/austria/33919144.pdf

Department for Education. (2011a). First report of the independent review os teachers' standards.

Department for Education. (2011b). *Teachers' standards*. London Retrieved from
    https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/283566/Teachers_standard_information.pdf.

Department for Education. (2012). Teacher appraisal and capability: A model policy for schools.   Retrieved from
    https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/282598/Teacher_appraisal_and_capability.pdf

Department for Education. (2013). Teachers' Standards: Guidance for school leaders, school staff and governing bodies. Retrieved from
    https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/301107/Teachers__Standards.pdf

Department for Education. (2014). Qualified teacher status (QTS): qualify to teach in England.   Retrieved from
    https://www.gov.uk/guidance/qualified-teacher-status-qts

Department for Education. (2015). Schools, pupils and their characteristics: January 2015.   Retrieved from
    https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/433680/SFR16_2015_Main_Text.pdf

155

Department for Education. (2016a). School Direct.   Retrieved from https://getintoteaching.education.gov.uk/explore-my-options/teacher-training-routes/school-led-training/school-direct

Department for Education. (2016b). Teacher training routes.   Retrieved from https://getintoteaching.education.gov.uk/explore-my-options/teacher-training-routes

Department of Education Employment and Workplace Relations. (2010). Educators Being, Belonging and Becoming: Educators' Guide to the Early Years Learning Framework for Australia.   Retrieved from https://docs.education.gov.au/system/files/doc/other/belonging_being_and_becoming_the_early_years_learning_framework_for_australia.pdf

District of Columbia Public Schools. (2014a). IMPACT: The DCPS Effectiveness Assessment System for School-Based Personnel (Group 1) 2014-2015.   Retrieved from http://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/IMPACT-2014-Grp1f.pdf

District of Columbia Public Schools. (2014b). IMPACT: The DCPS Effectiveness Assessment System for School-Based Personnel (Group 2) 2014-2015.   Retrieved from http://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/Group%202_r.pdf

District of Columbia Public Schools. (2015a). IMPACT: The DCPS Effectiveness Assessment System for School-Based Personnel (Group 1) 2015-2016.   Retrieved from http://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/Group%201.pdf

District of Columbia Public Schools. (2015b). IMPACT: The DCPS Effectiveness Assessment System for School-Based Personnel (Group 2) 2015-2016.   Retrieved from http://dcps.dc.gov/sites/default/files/dc/sites/dcps/publication/attachments/Group%202_r.pdf

District of Columbia Public Schools. (2015c). The Teaching and Learning Framework.   Retrieved from http://dcps.dc.gov/page/teaching-and-learning-framework

District of Columbia Public Schools. (2016a). IMPACT: An Overview.   Retrieved from http://dcps.dc.gov/page/impact-overview

District of Columbia Public Schools. (2016b). Individual Value-Added (IVA).   Retrieved from http://dcps.dc.gov/page/individual-value-added-iva

District of Columbia Public Schools. (2016c). Master Educators.   Retrieved from http://dcps.dc.gov/page/master-educators

District of Columbia Public Schools. (2016d). Partnership for Assessment of Readiness for College and Careers (PARCC).   Retrieved from http://dcps.dc.gov/parcc

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., & Pianta, R. C. (2010). The individualized classroom assessment scoring system (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Childhood Research Quarterly, 25*(1), 1-16.

Education and Training. (2013). How to Get Qualified.   Retrieved from http://www.education.vic.gov.au/about/careers/teaching/Pages/qualifying.aspx

Education Bureau. (2003). Teacher Performance Management.   Retrieved from http://www.edb.gov.hk/attachment/en/sch-admin/sbm/sbm-forms-references/staff-appraisal-system/TPMguide.pdf

Education Bureau. (2016a). Hong Kong: The Facts - Education.   Retrieved from http://www.edb.gov.hk/attachment/en/about-edb/publications-stat/hk_thefacts_education/The%20Facts_Education_Mar_2016_EN.pdf

Education Bureau. (2016b). Teacher Registration.   Retrieved from http://www.edb.gov.hk/en/teacher/qualification-training-development/qualification/teacher-registration/index.html

Education Bureau. (2016c). Teacher Training & Qualifications.   Retrieved from http://www.edb.gov.hk/en/teacher/qualification-training-development/qualification/teacher-training-qualifications/index.html

Education Council New Zealand. (2015). Approval, Review and monitoring process and requirements for ITE programs. Retrieved from https://educationcouncil.org.nz/sites/default/files/Approval,%20Review%20and%20Monitoring%20Processes%20for%20ITE%20programmes%20August%20%202015_0.pdf

Education Council New Zealand. (2016a). Appraisal of Teachers Project. Retrieved from https://educationcouncil.org.nz/appraisal-teachers-project

Education Council New Zealand. (2016b). Getting Started - Certification. Retrieved from https://educationcouncil.org.nz/content/getting-started-%E2%80%93-certification

Education Council New Zealand. (2016c). Graduating Teacher Standards. Retrieved from http://educationcouncil.org.nz/content/graduating-teacher-standards

Education Council New Zealand. (2016d). Initial Teacher Education Providers. Retrieved from https://educationcouncil.org.nz/content/initial-teacher-education-providers

Education Council New Zealand. (2016e). Practising Teacher Criteria. Retrieved from https://educationcouncil.org.nz/content/registered-teacher-criteria-1#h4

Education Council New Zealand. (2016f). Renewing a Full Practising Certificate. Retrieved from https://educationcouncil.org.nz/content/renewing-full-practising-certificate

Education Council New Zealand. (2016g). Studying to be a teacher. Retrieved from https://educationcouncil.org.nz/content/studying-be-teacher

Education Counts. (2016). School Rolls: Time Series Data for Student Numbers. Retrieved from https://www.educationcounts.govt.nz/statistics/schooling/student-numbers/6028

Education Data Partnership. (2016). California Public Schools: Demographics. Retrieved from http://www.ed-data.org/state/CA

Education in New Zealand. (2016). Educational System. Retrieved from http://www.education-newzealand.org/educational-system/

Education Review Office. (2014). Supporting school improvement through effective teacher appraisal. Retrieved from http://www.ero.govt.nz/assets/Uploads/TeacherAppraisal-Final-May-2014.pdf

Education Scotland. (2016a). The curriculum in Scotland.

Education Scotland. (2016b). Key features of PRD: Key feature 1. Retrieved from http://www.educationscotland.gov.uk/professionallearning/prd/keyfeatures/keyfeature1.asp

Education Scotland. (2016c). Professional Update. Retrieved from http://www.educationscotland.gov.uk/professionallearning/prd/profupdate.asp

Education Scotland. (2016d). Validated self evaluation reports. Retrieved from http://www.educationscotland.gov.uk/inspectionandreview/reports/othersectors/validatedselfevaluation/index.asp

Education Service District 112. (2016). eVAL Professional Development. Retrieved from https://web3.esd112.org/tpep/eval/eval-prof-dev

Educational Testing Service. (2003). Preparing Teachers Around the World. Retrieved from https://www.ets.org/Media/Education_Topics/pdf/prepteach.pdf

Embiza, S., & Hadush, S. (2015). Assessing the dimensionality and reliability of teachers' performance evaluation in Eastern Zone high schools, Tigrai National Regional State, Ethiopia. *Journal of Education and Practice, 6*(7), 91-99.

Engelhard, G., Myford, C., & Cline, F. (2000). Investigating assessor effects in National Board for Professional Teaching Standards assessments for early childhood/generalist and middle childhood/generalist certification. Retrieved from http://www.ets.org/Media/Research/pdf/RR00-13-Engelhard%20.pdf

157

European Agency for Special needs and Inclusive education. (2015). Germany: Teacher training basic and specialist teacher training. Retrieved from https://www.european-agency.org/country-information/germany/national-overview/teacher-training-basic-and-specialist-teacher-training

European Agency for Special Needs and Inclusive Education. (2016a). Austria - Teacher training - basic and specialist teacher training. Retrieved from https://www.european-agency.org/country-information/austria/national-overview/teacher-training-basic-and-specialist-teacher-training

European Agency for Special Needs and Inclusive Education. (2016b). Structure and content of initial teacher education courses - Austria. Retrieved from https://www.european-agency.org/agency-projects/Teacher-Education-for-Inclusion/country-info/austria/structure-and-content-of-initial-teacher-education-courses

European Union. (2015). Austria Overview. Retrieved from http://europa.eu/about-eu/countries/member-countries/austria/index_en.htm

Ewing, R., Groundwater-Smith, S., Mockler, N., Loughland, T., Simpson, A., Smith, D., . . . Brooks, D. (2009). *Meta Analysis of Quality Teaching Action Learning Project*. University of Sydney.

Exstrom, M. (2011). *National Board for Professional Teaching Standards certification: what legislators need to know*. Retrieved from Denver, CO: http://www.ncsl.org/portals/1/documents/educ/nationalboard.pdf

Fendick, F. (1990). *The correlation between teacher clarity of communication and student achievement gain: A meta-analysis.* (Doctoral dissertation), University of Florida, Florida.

Ferguson, S. (2009). *Teacher-leaders' use of reflective assessment practices to improve student learning.* Northern Illinois University, ProQuest Dissertations and Theses.

Figazzolo, L. (2013). The Use and Misuse of Teacher Appraisal: An overview of cases in the developed world

 Retrieved from http://download.ei-ie.org/Docs/WebDepot/TeacherAppraisal.pdf

Finucane, P., Flannery, D., Keane, D., & Norman, G. (2010). Cross-institutional progress testing: feasibility and value to a new medical school. *Med Educ., 44*(2), 184-186.

Fleener, M., & Exner, P. (2011). Dimensions of teacher education accountability: A Louisiana perspective on value added. *Teacher Education Policy*, 26-43.

Gagné, R., Wagner, W., Golas, K., & Keller, J. (2005). *Principles of Instructional Design* (5th ed.). Belmont, CA: Thomson Wadsworth.

Goe, L., Bell, C., & Little, O. (2008a). *Approaches to Evaluating Teacher Effectiveness: A Research Synthesis*. Retrieved from https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED521228&site=eds-live&scope=site

Goe, L., Bell, C., & Little, O. (2008b). Approaches to Evaluating Teacher Effectiveness: A Research Synthesis. *National Comprehensive Center for Teacher Quality*.

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National Board certification as a signal of effective teaching. *Review of Economics and Statistics, 89*(1), 134-150.

Goldstein, J. (2007). Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review. *American Journal of Education, 113*(3), 479-508.

Gore, J., Mockler, N., Smith, M., & Bowe, J. (2012). *Response to 'Great Teaching, Inspired Learning'*. Retrieved from http://www.schools.nsw.edu.au/media/downloads/news/greatteaching/submissions/university-of-newcastle.pdf

Graham, P. (2007). Improving teacher effectiveness through structured collaboration: A case study of a professional learning community. *RMLE Online, 31*(1), 1-17.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Great Schools. (2016). Schools in Washington D.C.   Retrieved from http://www.greatschools.org/washington-dc/washington/schools/?gradeLevels%5B%5D=p&gradeLevels%5B%5D=e&gradeLevels%5B%5D=m&gradeLevels%5B%5D=h

Greenblatt, D., & O'Hara, K. E. (2015). Buyer beware: Lessons learned from EdTPA implementation in New York State. *Teacher Education Quarterly, 42*(2), 57-67.

Grossman, P., Wineburg, S., & Woolworth, S. (2001). Toward a theory of teacher community. *The Teachers College Record, 103*, 942-1012.

Group, D. R. (2014). Demographics.   Retrieved from http://www.coopercenter.org/demographics

Hachfeld, A., Hahn, A., Schroeder, S., Anders, Y., & Stanat, P. (2011). Assessing teachers' multicultural and egalitarian beliefs: the teacher cultural beliefs scale. *Teaching and Teacher Education, 27*, 986-996.

Haertel, E. (2013). *Reliability and Validity of Inferences about Teachers Based on Student Scores*. Retrieved from https://www.ets.org/Media/Research/pdf/PICANG14.pdf

Hafen, C. A., Hamre, B. K., Allen, J. P., Bell, C. A., Gitomer, D. H., & Pianta, R. C. (2014). Teaching through interactions in secondary school classrooms: Revisiting the factor structure and practical application of the classroom assessment scoring system-secondary. *The Journal of Early Adolescence, 35*(5-6), 651-680. doi:10.1177/0272431614537117

Haigh, M., & Ell, F. (2014). Consensus and dissensus in mentor teachers' judgments of readiness to teach. *Teaching and Teacher Education, 40*, 10-21.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). *The Market for Teacher Quality*. (NBER Working Paper Series No. 11154).

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1998). *Teachers, Schools, and Academic Achievement*. (NBER Working Paper No. 6691).

Hargreaves, D. (2011). *The challenge for the comprehensive school: culture, curriculum and community* (Vol. 43): Routledge.

Harrington, H., & Hathaway, R. (1995). Illuminating beliefs about diversity. *Journal of Teacher Education, 46*(4), 275-284. doi:DOIs: 10.1177/0022487195046004006

Harris, D. N., & Sass, T. R. (2009). *What Makes for a Good Teacher and Who Can Tell?* Retrieved from

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Eonomics, 95*(7), 798-812.

Harvard Graduate School of Education. (2016). What is PAR?   Retrieved from http://www.gse.harvard.edu/~ngt/par/parinfo/

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.

Hattie, J. (2013). *Visible Learning for Teachers*. London, UK: Taylor & Francis Ltd.

Hattie, J., & Clinton, J. (2008). Identifying accomplished teachers: A validation study. In L. C. Ingvarson & J. Hattie (Eds.), *Assessing Teachers for Professional Certification: The First Decade of the National Board for Professional Teaching Standards* (Vol. 11, pp. 313-344). Oxford: Elsevier.

Hattie, J., & Donoghue, G. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning, 1*. doi:doi:10.1038/npjscilearn.2016.13

Haystead, M. (2010). *Report on professional development: A summary of findings for ten teachers on three critical commitments to school reform*. Retrieved from Englewood, CO:

Headden, S. (2011). Inside IMPACT: D.C.'s Model Teacher Evaluation System.   Retrieved from http://educationpolicy.air.org/sites/default/files/publications/IMPACT_Report_RELEASE.pdf

159

Hill, H., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*(2), 56-64.

Hill, H., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard educational review, 83*(2), 371-384.

Ho, A. D., & Kane, T. J. (2013). The reliability of classroom observations by school personnel. *Bill & Melinda Gates Foundation*.

Holler, W., Gareis, C., Martin, J., Clouser, A., & Miller, S. (2008). Teacher-made assessments: Getting them right. *Principal Leadership, 9*(1), 60-64.

Hoy, A. (2000). *Changes in Teacher Efficacy During the Early Years of Teaching*. Paper presented at the American Education Research Association, New Orleans.

Independent Schools Council. (2016). ISC Census and Annual Report Retrieved from http://www.isc.co.uk/media/3179/isc_census_2016_final.pdf

Ingvarson, L. C., & Hattie, J. (2008). *Assessing Teachers for Professional Certification: The first decade of the National Board for Professional Teaching Standards*. Bingley, UK: Emerald Group Publishing Limited.

Irving, S. (2004). *The development and validatino of a student evaluation instrument to identify highly accomplished mathematics teachers.* (Doctoral dissertation), The University of Auckland.

Isoré, M. (2009). *Teacher Evaluation: Current Practices in OECD Countries and a Literature Review* (1993-9019). Retrieved from

Jacob, B. A., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of labor Economics, 26*(1), 101-136.

Jensen, B. (2010). *Making time for great teaching*. Retrieved from http://grattan.edu.au/wp-content/uploads/2014/03/808-making-time-for-great-teaching.pdf

Jordan, K. (2012). *Beginning teacher TPACK knowledge: Influence of school type.* Paper presented at the ACEC2012: ITs Time Conference.

Joshua, M., Joshua, A., & Maliki, A. (2007). Use of student achievement scores as basis for assessing teachers' instrucitonal effectiveness: Issues and research results. *J. Hum. Ecol., 22*(2), 135-140.

Kane, T. J., McCaffrey, D., Miller, T., & Staiger, D. O. (2013). Have we identified effective teachers? Validating measures of effective teaching using random assignment. *MET Project Research Paper, Bill & Melinda Gates Foundation*.

Kane, T. J., & Staiger, D. O. (2012). Gathering feedback for teaching: combining high-quality observations with student surveys and achievement gains. *Bill & Melinda Gates Foundation*.

Kennedy, M. (2008). Contributions of qualitative research to research on teacher qualifications. *Educational Evaluation and Policy Analysis, 30*(4), 344-367.

Kılıç, D. (2016). An examination of using self-, peer-, and teacher-assessment in higher education: A case study in teacher education *Higher Education Studies, 6*(1), 136-144. doi:doi: 10.5539/hes.v6n1p136

Kim, E.-G., & Han, Y.-K. (2002). Attracting, developing and retaining effective teachers: Background Report for Korea. *Seul: Korean Educational Development Institute. Recuperado el.* Retrieved from https://www.oecd.org/edu/school/2713221.pdf

Kleinhenz, E., & Ingvarson, L. C. (2007). Standards for Teaching: Theoretical Underpinnings and Applications. Retrieved from http://research.acer.edu.au/cgi/viewcontent.cgi?article=1000&context=teaching_standards

Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences, 25*(67-72).

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Korean Educational Development Institute. (2010). Country Background Report for Korea - OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.   Retrieved from http://www.oecd.org/edu/school/49363138.pdf

Kraft, M. A., & Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review, 47*, 49-63.

Kyriacou, C. (2001). Teacher stress: Directions for future research. *Educational Review, 53*, 28-35.

Ladwig, J., Smith, M., Gore, J., Amosa, W., & Griffiths, T. (2007). *Quality of pedagogy and student achievement: Multi-level replication of authentic pedagogy*. Paper presented at the Australian Association for Research in Education (AARE), Fremantle.

Lasagabaster, D., & Sierra, J. M. (2011). Classroom observation: desirable conditions established by teachers. *European Journal of Teacher Education, 34*(4), 449-463.

Lavery, M. R., Holloway-Libell, J., Amrein-Beardsley, A., Pivovarova, M., & Hahs-Vaughn, D. L. (n.d.). *Evaluating the Validity Evidence Surrounding the Use of Student Standardised Test Scores to Evaluate Teachers: A Centennial, Systematic Mega-Review*. Retrieved from http://www.mattlavery.com/wp-content/uploads/2016/05/Evidence-Validating-Test-Based-Teacher-Evaluation.pdf

Learning Sciences International: Learning and Performance Management. (2016). Introducing Dr. Marzano's Teacher Evaluation and Leadership Evaluation Models.   Retrieved from http://www.marzanoevaluation.com/

Lederman, G., & Newsome, J. (1992). Do subject matter knowledge, pedagogical knowledge, and pedagogical content knowledge constitute the ideal gas law of science teaching? *Journal of Science Teacher Education, 3*(1), 16-20.

Ledwell, K., & Oyler, C. (2016). Unstandardized responses to a "standardized" test: The edTPA as gatekeeper and curriculum change agent. *Journal of Teacher Education, 67*(2), 120-134.

Lin, S. (2015). *Learning through Action: Teacher Candidates and Performance Assessments.* University of Washington.

Linacre, J. M. (1989). *Many-facet Rasch measurement* Chicago: MESA Press.

Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics, 3*(1), 101-128.

Little, O., Goe, L., & Bell, C. (2009). *A Practical Guide to Evaluating Teacher Effectiveness*. Retrieved from

Marsh, H., & Roche, L. (1997). Making students' evaluations of teaching effectiveness effective: The critical isues of validity, bias, and utility. *American Psychologist, 52*(11), 1187-1197.

Martínez, J. F., Borko, H., Stecher, B., Luskin, R., & Kloser, M. (2012). Measuring classroom assessment practice using instructional artifacts: A validation study of the QAS notebook. *Educational Assessment, 17*(2-3), 107-131.

Marzano Research Laboratory. (2010). Phase I – What Works in Oklahoma Schools?   Retrieved from http://www. marzanoresearch.com/documents/PhaseI_WWIO.pdf

Marzano Research Laboratory. (2011a). 2010-2011 The Adams 50 Instructional Model Study: Individual Teacher Effectiveness Linked to Student Growth.   Retrieved from http://www. sbsadams50.org/wp-content/uploads/2011/10/ACSD50-InstructionalModel-Final-Validity-Report-2011.pdf

Marzano Research Laboratory. (2011b). Reliability Study: Cherry Creek School District Observational Protocol. .   Retrieved from marzanoresearch.com.

Marzano, R. J., & Haystead, M. W. (2009). *Evaluation Study of the Effects of Promethean ActivClassroom on Student Achievement. Final Report*. Retrieved from http://www.marzanoresearch.com/documents/finalreportonactivclassroom.pdf

161

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Marzano, R. J., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*: ASCD.

Master, B. (2013). *What Can Parents Tell Us About Teacher Quality? Examining the Contributions of Parent Perspectives in Comparison to a Portfolio of Alternative Teacher Evaluation Measures.* (Doctoral dissertation), Stanford University.

Mathers, C., & Oliva, M. (2008). *Improving Instruction through Effective Teacher Evaluation: Options for States and Districts. TQ Research & Policy Brief*. Retrieved from http://www.gtlcenter.org/sites/default/files/docs/February2008Brief.pdf

Matsumura, L. C., Slater, S. C., Junker, B., Peterson, M., Boston, M., Steele, M., & Resnick, L. (2006). *Measuring Reading Comprehension and Mathematics Instruction in Urban Middle Schools: A Pilot Study of the Instructional Quality Assessment*. Retrieved from https://www.cse.ucla.edu/products/reports/R681.pdf

Measures of Effective Teaching. (2009). Measures of Effective Teaching Project Series. Retrieved from https://www.icpsr.umich.edu/icpsrweb/METLDB/series/257

Menter, I., Hulme, M., Elliot, D., & Lewin, J. (2010). *Literature Review on Teacher Education in the 21st Century*. Retrieved from http://www.gov.scot/resource/doc/325663/0105011.pdf

Metzler, J., & Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation. *Journal of Development Economics, 99*(2), 486-496.

Milanowski, A. T. (2011). Validity research on teacher evaluation systems based on the framework for teaching. *Online Submission*.

Milton, M., Rohl, M., & House, H. (2007). Secondary beginning teacher's preparedness to teach literacy and numeracy: A survey. *Australian Journal of Teacher Education, 32*(2), 37.

Ministry of Education. (2015). Bringing out the best in every child. Retrieved from https://www.moe.gov.sg/docs/default-source/document/about/files/moe-corporate-brochure.pdf

Ministry of Education. (2016). About Us. Retrieved from https://www.moe.gov.sg/about

Ministry of Education Korea. (2015). Statistics. Retrieved from http://english.moe.go.kr/web/1722/site/contents/en/en_0219.jsp

Ministry of Education New Zealand. (2010). OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes. Retrieved from https://www.educationcounts.govt.nz/__data/assets/pdf_file/0009/90729/966_OECD-report.pdf

Ministry of Education New Zealand. (2015a). Primary Teachers' Collective Agreement 2015-2018. Retrieved from http://www.education.govt.nz/assets/Documents/School/Collective-Employment-Agreements/Primary-Teachers-Collective-Agreement/PTCA-2016-2018.pdf

Ministry of Education New Zealand. (2015b). Secondary Teachers' Collective Agreement 2015-2018. Retrieved from http://www.education.govt.nz/assets/Documents/School/Collective-Employment-Agreements/Secondary-Teachers-Collective-Agreement/SecondaryTeachersCA20152018.pdf

Ministry of Education New Zealand. (2016a). Education in New Zealand. Retrieved from http://www.education.govt.nz/home/education-in-nz/

Ministry of Education New Zealand. (2016b). Performance management. Retrieved from http://www.education.govt.nz/school/running-a-school/employing-and-managing-staff/performance-management/

Ministry of Education New Zealand. (2016c). Performance management. Retrieved from http://www.education.govt.nz/school/running-a-school/employing-and-managing-staff/performance-management/

Ministry of Education New Zealand. (2016d). Primary Teachers' Collective Agreement 2013-2015. Retrieved from http://www.education.govt.nz/school/running-a-school/employing-and-managing-staff/collective-agreements/primary-teachers-collective-agreement/

162

Ministry of Education New Zealand. (2016e). Ruia: Teacher appraisal for Māori learners' success.   Retrieved from http://appraisal.ruia.educationalleaders.govt.nz/

Ministry of Education New Zealand. (2016f). School structures and governance.   Retrieved from http://www.education.govt.nz/school/running-a-school/school-structures-and-governance/

Ministry of Education New Zealand. (2016g). Secondary Teachers' Collective Agreement 2015-2018.   Retrieved from http://www.education.govt.nz/school/running-a-school/employing-and-managing-staff/collective-agreements/secondary-teachers-collective-agreement/

Ministry of Education New Zealand. (2016h). Teacher appraisal for Maori learners' successs.   Retrieved from http://appraisal.ruia.educationalleaders.govt.nz/

Ministry of Education Singapore. (2016a). Career Information.   Retrieved from https://www.moe.gov.sg/careers/teach/career-information

Ministry of Education Singapore. (2016b). Compulsory Education.   Retrieved from https://www.moe.gov.sg/education/education-system/compulsory-education

Ministry of Education Singapore. (2016c). Desired Outcomes of Education.   Retrieved from https://www.moe.gov.sg/education/education-system/desired-outcomes-of-education

Moltó, M. C., Florian, L., Rouse, M., & Stough, L. M. (2010). Attitudes to diversity: a cross-cultural study of education students in Spain, England and the United States. *European Journal of Teacher Education, 33*(3), 245-264.

Moon, B. (2014). The literacy skills of secondary teaching undergraduates: Results of diagnostic testing and a discussion of findings. *Australian Journal of Teacher Education, 39*(12), 12.

Moss, J. T. (2015). *The Danielson Model of Teacher Evaluation: Exploring Teacher Perceptions Concerning Its Value in Shaping and Improving Instructional Practice.* (Doctor of Education), Seton Hall University.   (Paper 2090)

Muijs, D., & Harris, A. (2003). Teacher leadership: Principles and practice.   Retrieved from http://dera.ioe.ac.uk/5132/1/download%3Fid%3D17417%26filename%3Dteacher-leadership-principles-practice-full-report.pdf

Munro, S. (2008). Opportunity lies in teacher–child interaction. *Education Digest, 73*(6), 46-48.

Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment Evaluation and Accountability, 26*(1).

National Board for Professional Teaching Standards. (2016a). About Certification.   Retrieved from http://boardcertifiedteachers.org/about-certification

National Board for Professional Teaching Standards. (2016b). Five Core Propositions.   Retrieved from http://www.nbpts.org/five-core-propositions

National Institute of Education. (2010). A Teacher Education Model for the 21st Century (TE21).   Retrieved from http://www.nie.edu.sg/docs/default-source/te21_docs/te21_executive-summary_14052010---updated.pdf?sfvrsn=2

National Institute of Education. (n.d.-a). Graduand Teacher Competencies.   Retrieved from http://www.nie.edu.sg/docs/default-source/default-document-library/te21---gtc.pdf?sfvrsn=0

National Institute of Education. (n.d.-b). The New V3SK Model.   Retrieved from http://www.nie.edu.sg/docs/default-source/td_practicum/te21---v3sk.pdf?sfvrsn=2

National Institute of Education Singapore. (2016). Teacher Education.   Retrieved from http://www.nie.edu.sg/teacher-education

NCEE. (2016). Teacher and Principal Quality.   Retrieved from http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/south-korea-overview/south-korea-teacher-and-principal-quality/

163

NCTL. (2016). About Us. Retrieved from https://www.gov.uk/government/organisations/national-college-for-teaching-and-leadership/about

New South Wales Department of Education and Training. (2003). Quality teaching in NSW public schools. Retrieved from http://www.darcymoore.net/wp-content/uploads/2012/02/qt_EPSColor.pdf

New South Wales Department of Education and Training. (2008). Quality Teaching to support the NSW Professional Teaching Standards. Retrieved from http://dei.eduone.net.au/ashford/DEPASCSCURRICULUM/HSIE%20UNIT/Quality%20Teaching%20Framework.pdf

New South Wales Department of Education and Training. (2016). Professional responsibilities. Retrieved from http://www.dec.nsw.gov.au/about-us/careers-centre/school-careers/teaching/your-teaching-career/approved-teachers/casual-teacher-induction/professional-responsibilities

Newman, F., Marks, H., & Garmoran, A. (1995). *Authentic pedagogy and student performance.*

Conference presentation to the American Education Research Association.

Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy and Analysis Archives, 18*(23), 1-27.

NYC Department of Education. (2016). Student Perception Survey Retrieved from http://schools.nyc.gov/Accountability/resources/Studentperceptionsurvey/default.htm

OECD. (2009). Creating Effective Teaching and Learning Environments: First results from the OECD Teaching and Learning International Survey: Briefing Note For Austria. Retrieved from https://www.oecd.org/edu/school/43046693.pdf

OECD. (2011). A Framework for Growth and Social Cohesion in Korea. Retrieved from http://www.oecd.org/social/family/48225033.pdf

OECD. (2012). OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes. Retrieved from http://www.oecd.org/edu/school/49578470.pdf

OECD. (2013a). LEARNING STANDARDS, TEACHING STANDARDS AND STANDARDS FOR SCHOOL PRINCIPALS: A COMPARATIVE STUDY. Retrieved from http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=EDU/WKP(2013)14&docLanguage=En

OECD. (2013b). PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know. Retrieved from https://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

OECD. (2013c). Teachers for the 21st Century: Using Evaluation to Improve Teaching: OECD Publishing.

OECD. (2014a). Education Policy Outlook: Germany. Retrieved from http://www.oecd.org/edu/EDUCATION%20POLICY%20OUTLOOK%20GERMANY_EN.pdf

OECD. (2014b). Indicator D3: How much are teachers paid? Retrieved from https://www.oecd.org/edu/EAG2014-Indicator%20D3%20(eng).pdf

OECD. (2014c). TALIS 2013 Results. Retrieved from http://www.keepeek.com/Digital-Asset-Management/oecd/education/talis-2013-results_9789264196261-en#.V9IHBfl9670

OECD. (2015). *PISA In Focus*. Retrieved from https://www.oecd.org/pisa/pisaproducts/pisainfocus/PIF-50-(eng)-FINAL.pdf

OECD. (2016). *Supporting Teacher Professionalism: Insights from TALIS 2013*. Paris: OECD Publishing.

Office for Standards in Education. (2015). Initial teacher education inspection handbook. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/459282/Initial_Teacher_Eduction_handbook_from_September_2015.pdf

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Office for Standards in Education. (2016). About us.   Retrieved from
https://www.gov.uk/government/organisations/ofsted/about

Office of Superintendent of Public Instruction. (2014). Washington State Leadership Student Growth Rubrics.   Retrieved from
http://www.k12.wa.us/TPEP/StudentGrowth/WA_Leadership_Student_Growth_Rubrics.pdf

Office of Superintendent of Public Instruction. (2015). Update: State of the State for Teacher and Principal Evaluation 2015.
Retrieved from http://www.k12.wa.us/legisgov/2015documents/TPEPUpdateJuly2015.pdf

Office of Superintendent of Public Instruction. (2016a). Education Laws in Washington State.   Retrieved from
http://www.k12.wa.us/ProfPractices/adminresources/faq.aspx

Office of Superintendent of Public Instruction. (2016b). Renew a Professional Teacher Certificate.   Retrieved from
http://www.k12.wa.us/certification/Teacher/Professional.aspx

Office of Superintendent of Public Instruction. (2016c). Statewide Framework Selections.   Retrieved from http://tpep-
wa.org/wp-content/uploads/Statewide%20framework%20overview.png

Office of Superintendent of Public Instruction. (2016d). Teacher/Principal Evaluation Program Retrieved from
http://www.k12.wa.us/edleg/tpep/default.aspx

Office of Superintendent of Public Instruction. (2016e). Teacher/Principal Evaluation Program: Rater Agreement.   Retrieved
from http://www.k12.wa.us/TPEP/RaterAgreement.aspx

Office of Superintendent of Public Instruction. (2016f). Teachers: Field Experience/Student Teaching in Washington.   Retrieved
from http://www.k12.wa.us/certification/Teacher/studentteaching.aspx

Office of Superintendent of Public Instruction. (2016g). Teachers: For Those Who Hold A Residency Teacher Certificate.
Retrieved from http://www.k12.wa.us/certification/teacher/Residency3.aspx

Office of Superintendent of Public Instruction. (2016h). Teachers: Professional Certificate.   Retrieved from
http://www.k12.wa.us/certification/Teacher/ProCert.aspx

Office of Superintendent of Public Instruction. (2016i). Teachers: Residency Certificate.   Retrieved from
http://www.k12.wa.us/certification/Teacher/Residency.aspx

Office of Superintendent of Public Instruction. (n.d.). OSPI Study Group Report.   Retrieved from
http://www.k12.wa.us/TPEP/Frameworks/OSPI_Learning_Environments.pdf

Office of the State Superintendent of Education. (2014). Educator Licensure Exams.   Retrieved from
http://osse.dc.gov/sites/default/files/dc/sites/osse/publication/attachments/Educator%20Testing%20Flyer_11%202201
4_0.pdf

Office of the State Superintendent of Education. (2016a). Educator Preparation Program Approval and Accreditation.   Retrieved
from http://osse.dc.gov/service/educator-preparation-program-approval-and-accreditation

Office of the State Superintendent of Education. (2016b). Interstate Licensure Agreement - Reciprocity.   Retrieved from
http://osse.dc.gov/service/interstate-licensure-agreement-reciprocity

Office of the State Superintendent of Education. (2016c). Teacher Licensure.   Retrieved from http://osse.dc.gov/teacher-
licensure

Ofsted. (2015). *School inspection handbook*. Manchester Retrieved from
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/457199/School_inspection_handboo
k.pdf.

Olson, D. R. P., S.; Robinson, S.; Marcus, L.; Gaite, A.J.H.; Traub, R. (1969). The teacher as communicator: An aspect of teacher
effectiveness.   Retrieved from http://files.eric.ed.gov/fulltext/ED040146.pdf

165

Ontario College of Teachers. (2016a). Accreditation Retrieved from https://www.oct.ca/public/accreditation

Ontario College of Teachers. (2016b). Development of Professional Standards Retrieved from http://www.oct.ca/public/professional-standards/development-of-professional-standards

Ontario College of Teachers. (2016c). Ethical Standards Retrieved from http://www.oct.ca/public/professional-standards/ethical-standards

Ontario College of Teachers. (2016d). Registration.   Retrieved from https://www.oct.ca/becoming-a-teacher/requirements

Ontario College of Teachers. (2016e). Types of Teaching Certificates.   Retrieved from https://www.oct.ca/public/professional-standards/how-teachers-are-certified/certificates

Ontario Ministry of Education. (2002). Policy/Program Memorandum No. 131.   Retrieved from http://www.edu.gov.on.ca/extra/eng/ppm/131.html

Ontario Ministry of Education. (2010). *Teacher Performance Appraisal Technical Requirements Manual*.   Retrieved from http://www.edu.gov.on.ca/eng/teacher/pdfs/TPA_Manual_English_september2010l.pdf.

Ontario Ministry of Education. (2016). Education Facts, 2014-2015 Retrieved from http://www.edu.gov.on.ca/eng/educationFacts.html

Oxford Dictionaries. (2016). rating.   Retrieved from http://www.oxforddictionaries.com/definition/english/rating?q=ratings

Ozogul, G., & Sullivan, H. (2009). Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Educational Technology Research and Development, 57*(3), 393-410.

Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., & Nurmi, J.-E. (2010). A validation of the classroom assessment scoring system in Finnish kindergartens. *Early Education & Development, 21*(1), 95-124. doi:10.1080/10409280902858764

Papay, J. P. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *Am. Educ. Res., 48*(1), 163-193.

Papay, J. P., Johnson, S. M., Fiarman, S. F., Munger, M. S., & Qazilbash, E. K. (2009). *Beyond Dollars and Cents: The Costs and Benefits of Teacher Peer Assistance and Review*. Paper presented at the American Educational Research Association, San Diego.

Parkes, K. A., & Powell, S. R. (2015). Is the edTPA the right choice for evaluating teacher readiness? *Arts Education Policy Review, 116*(2), 103-113.

Patrikakou, E. N., & Weissberg, R. P. (2000). Parents' perceptions of teacher outreach and parent involvement in children's education. *Journal of Prevention & Intervention in the Community, 20*(1-2), 103-119.

Paufler, N., & Amrein-Beardsley, A. (2014). The random assignment of students into elementary classrooms: Implications for value-added analyses and interpretations. *American Educational Research Journal, 51*(2), 328-362. doi:doi: 10.3102/0002831213508299

Pearson. (2016). Become a National Board Assessor.   Retrieved from http://scorenbpts.pearson.com/become-a-national-board-assessor.html

Pearson VUE. (2016). National Board for Professional Teacher Standards Certification Testing.   Retrieved from http://www.pearsonvue.com/nbpts/

Pecheone, R. L., & Chung, R. R. (2006). Evidence in Teacher Education The Performance Assessment for California Teachers (PACT). *Journal of Teacher Education, 57*(1), 22-36.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*: American Psychological Association.

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109-119. doi:10.3102/0013189x09332374

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system (CLASS) manual, pre-K*: Paul H. Brookes Publishing Company Baltimore, MD.

Polikoff, M. S. (2015). The stability of observational and student survey measures of teaching effectiveness. *American Journal of Education, 121*(2), 183-212. Retrieved from https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1051313&site=eds-live&scope=site

Popham, W. J. (2013). *Classroom Assessment: What Teachers Need to Know*: Pearson.

Protheroe, N. (2008). Teacher efficacy: what is it and does it matter? *Principal, 87*(5), 42-45.

Reddy, L. A., Dudek, C. M., Fabiano, G. A., & Peters, S. (2015). Measuring teacher self-report on classroom practices: Construct validity and reliability of the Classroom Strategies Scale–Teacher Form. *School Psychology Quarterly, 30*(4), 513.

Reiter, A. B., & David, S. N. (2011). Factors influencing pre-service teachers' beliefs about student achievement evaluation of a pre-service teacher diversity awareness program. *Multicultural Education, 19*(3), 41-46.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247-252.

Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy, 6*(1), 43-74.

Roe, J. W. (2010). *Student self-assessment and student ratings of teacher rapport in secondary student course ratings.* Utah State University.

Ross, J., & Bruce, C. (2007). Self-assessment and professional growth: The case of a grade 8 mathematics teacher. *Teaching and Teacher Education, 23*(2), 146-159.

Ryan, J., & McCrae, B. (2006). Assessing pre-service teachers' mathematics subject knowledge. *Mathematics teacher education and development, 7*, 72-89.

Sachs, J. (2003). Teacher professional standards: Controlling or developing teaching? *Teachers and Teaching, 9*(2), 175-186. doi:DOI: 10.1080/1354060032000089487

Sallinen-Kuparinen, A. (1992). Teacher communicator style. *Communication Education, 41*(2), 153-166.

Sandoval, A. (2015). Student Progress Ignored: An examination of California school districts' compliance with the Stull Act. Retrieved from http://edvoice.org/sites/default/files/STUDENT%20PROGRESS%20IGNORED.pdf

Sartain, L., Stoelinga, S. R., Brown, E. R., Matsko, K. K., Miller, F. K., Durwood, C. E., . . . Glazer, D. (2011). Rethinking teacher evaluation in Chicago. *Chicago: Consortium on Chicago School Research*.

SCALE. (2014). Educative Assessment & Meaningful Support. Retrieved from https://secure.aacte.org/apps/rl/res_get.php?fid=2183&ref=edtpa

Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior, 25*, 293-315.

Scottish Council of Independent Schools. (2016). Pupil Numbers. Retrieved from http://www.scis.org.uk/facts-and-figures/pupil-numbers/

Scottish Credit and Qualifications Framework. (2016). Scottish Credit and Qualifications Framework (SCQF). Retrieved from http://scqf.org.uk/

Scriven, M. (1987). Critical thinking and the concept of literacy. *Informal Logic, 9*, 93-110.

167

Singapore Government. (2013). School Boards (Incorporation) Act (Chapter 284A). Retrieved from http://statutes.agc.gov.sg/aol/search/display/view.w3p;page=0;query=DocId%3A447805a4-57b6-4de0-a697-7518a65172f2%20%20Status%3Ainforce%20Depth%3A0;rec=0

Sirvani, H. (2007). The effect of teacher communication with parents on students' mathematics achievement. *American Secondary Education*, 31-46.

Slater, H., Davies, N., & Burgess, S. (2009). Do teachers matter? Measuring the variation in teacher effectiveness in England. Retrieved from http://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/wp212.pdf

Steiner, L. (2010). Lessons for Singapore. Retrieved from http://opportunityculture.org/images/stories/singapore_lessons_2010.pdf

Strategic Data Project. (2012). SDP Human Capital Diagnostic for Los Angeles Unified School District. Retrieved from http://cepr.harvard.edu/publications/sdp-human-capital-diagnostic-los-angeles-unified-school-district

Strong, M., Gargani, J., & Hacifazlioğlu, Ö. (2011). Do we know a successful teacher when we see one? Experiments in the identification of effective teachers. *Journal of Teacher Education*, 0022487110390221.

Stronge, J., Ward, T., & Grant, L. (2011). What makes good teachers good? A cross-case analysis of the connected between teacher effectiveness and student achievement. *Journal of Teacher Education, 62*(4), 339-355.

Suldo, S., Thalji, A., & Ferron, J. (2011). Longitudinal academic outcomes predicted by early adolescents' subjective well-being, psychopathology, and mental health status yielded from a dual factor model. *The Journal of Positive Psychology, 6*(1), 17-30.

Tagomori, H., & Bishop, L. (1995). Student evaluation of teaching: Flas in the instruments. *Thought and Action, 11*, 63-78.

Tai, D. W., Hu, Y.-C., Wang, R., & Chen, J.-L. (2012). *What is the impact of teacher self-efficacy on the student learning outcome?* Paper presented at the 3rd WIETE Annual Conference on Engineering and Technology Education, Pattaya, Thailand.

Tan, C. Y., & Dimmock, C. (2014). How a 'top-performing' Asian school system formulates and implements policy: The case of Singapore. *Educational Management Administration & Leadership, 42*(5), 743-763. doi:10.1177/1741143213510507

Tchoshanov, M. A. (2011). Relationship between teacher knowledge of concepts and connections, teaching practice, and student achievement in middle grades mathematics. *Educational Studies in Mathematics, 76*(2), 141-164.

Teach in Scotland. (2016). Qualifications required. Retrieved from http://www.teachinscotland.org/doihavewhatittakes/qualifications/scqf.asp

Teach.com. (2015). Become a Teacher in Washington. Retrieved from https://teach.com/states/become-a-teacher-in-washington/

Teacher Education Ministerial Advisory Group. (2014a). Action now: Classroom ready teachers. Retrieved from https://docs.education.gov.au/system/files/doc/other/action_now_classroom_ready_teachers_print.pdf

Teacher Education Ministerial Advisory Group. (2014b). *Action Now: Classroom Ready Teachers*. Retrieved from https://docs.education.gov.au/system/files/doc/other/action_now_classroom_ready_teachers_accessible.pdf

Teacher.org. (2016). How to Become a Teacher In Virginia. Retrieved from http://www.teacher.org/state/virginia/

TeachFirst. (2016). What we do. Retrieved from https://www.teachfirst.org.uk/what-we-do

Teachscape. (2016a). Framework for teaching Retrieved from https://www.teachscape.com/solutions/framework-for-teaching

Teachscape. (2016b). Teachscape Reflect: Manage and conduct evaluations that matter to instructional leaders, teachers, and students. Retrieved from https://www.teachscape.com/products/reflect#training

The Alberta Teachers' Association. (2016). Alberta Certification.   Retrieved from
       https://www.teachers.ab.ca/Teaching%20in%20Alberta/Becoming%20a%20Teacher/Pages/Alberta%20Certification.as
       px

The Australian Bureau of Statistics. (2016a). 4221.0 - Schools, Australia, 2015.   Retrieved from
       http://www.abs.gov.au/ausstats/abs@.nsf/mf/4221.0

The Australian Bureau of Statistics. (2016b). Primary and Secondary Education.   Retrieved from
       http://www.abs.gov.au/AUSSTATS/abs@.nsf/0/93EB4563583425CCCA25773700169C91?opendocument

The Australian Institute for Teaching and School Leadership. (2013). Australian Teacher Performance and Development
       Framework.   Retrieved from http://www.aitsl.edu.au/docs/default-source/default-document-
       library/australian_teacher_performance_and_development_framework

The board trustees of school district no. 34. (2009). Provincial collective agreement between: British Columbia public school
       employers association/ the board of school trustees of school district no. 34 Retrieved from
       http://www.bcpsea.bc.ca/documents/34-pca-0611.pdf

The Colorado Education Initiative. (2013). *STUDENT PERCEPTION SURVEY TECHNICAL REPORT*. Retrieved from Colorado:
       http://www.coloradoedinitiative.org/wp-content/uploads/2014/09/technical-report-CEI.pdf

The Consortium for Policy Research in Education. (2007). A Comparative Study of Teacher Preparation and Qualifications in Six
       Nations.   Retrieved from http://www.cpre.org/sites/default/files/researchreport/823_sixnationsfinal.pdf

The Department of Education and Training. (2016). Managing Teacher Performance.   Retrieved from
       http://www.education.vic.gov.au/school/principals/management/Pages/manageteacherperf.aspx

The Educational Institute of Scotland. (2015). Welcome.   Retrieved from http://www.eis.org.uk/About/Welcome.htm

The General Teaching Council for Scotland. (2006). Policy Statement: Accreditation of Programmes of Initial Teacher Education
       in Scotland.   Retrieved from http://www.gtcs.org.uk/web/FILES/education-in-scotland/policy-statement-accreditation-
       of-programmes-of-ITE-in-scotland.pdf

The General Teaching Council for Scotland. (2012a). The Standard for Career-Long Professional Learning: Supporting the
       development of teacher professional learning.   Retrieved from http://www.gtcs.org.uk/web/FILES/the-
       standards/standard-for-career-long-professional-learning-1212.pdf

The General Teaching Council for Scotland. (2012b). The Standards for Leadership and Management: Supporting leadership and
       management development.   Retrieved from http://www.gtcs.org.uk/web/FILES/the-standards/standards-for-
       leadership-and-management-1212.pdf

The General Teaching Council for Scotland. (2012c). The Standards for Registration: Mandatory requirements for Registration
       with the General Teaching Council for Scotland.   Retrieved from http://www.gtcs.org.uk/web/FILES/the-
       standards/standards-for-registration-1212.pdf

The General Teaching Council for Scotland. (2013). Guidelines for Initial Teacher Education Programmes in Scotland.   Retrieved
       from http://www.gtcs.org.uk/web/FILES/about-gtcs/guidelines-for-ite-programmes-in-scotland.pdf

The General Teaching Council for Scotland. (2016a). Additional Support Needs Registration.   Retrieved from
       http://www.gtcs.org.uk/registration/apply-for-registration/additional-registration/additional-support-needs-
       registration.aspx

The General Teaching Council for Scotland. (2016b). Professional Registration.   Retrieved from
       http://www.gtcs.org.uk/registration/apply-for-registration/additional-registration/professional-registration.aspx

The General Teaching Council for Scotland. (2016c). Professional Review and Development.   Retrieved from
       http://www.educationscotland.gov.uk/professionallearning/prd/index.asp

The General Teaching Council for Scotland. (2016d). Registration.   Retrieved from
       http://www.gtcs.org.uk/registration/registration.aspx
169

The General Teaching Council for Scotland. (2016e). Scotland's Education System. Retrieved from
    http://www.gtcs.org.uk/TeacherJourney/scotlands-education-system.aspx

The General Teaching Council for Scotland. (2016f). Self-evaluation. Retrieved from
    http://www.educationscotland.gov.uk/professionallearning/clpl/selfevaluation.asp

The General Teaching Council for Scotland. (2016g). Standards for Registration. Retrieved from
    http://www.gtcs.org.uk/professional-standards/the-standards/standards-for-registration.aspx

The General Teaching Council for Scotland. (2016h). Teacher Journey. Retrieved from
    http://www.gtcs.org.uk/TeacherJourney/teacher-journey.aspx

The General Teaching Council for Scotland. (2016i). Using Evidence of Impact and Examples of Professional Learning Journeys.
    Retrieved from http://www.gtcs.org.uk/professional-update/professional-learning/professional-learning-journey-using-
    evidence-of-impact.aspx

The Government of the Hong Kong Special Administrative Region. (2007). Accreditation of Academic and Vocational
    Qualifications Ordinance. Retrieved from
    http://www.legislation.gov.hk/blis_pdf.nsf/6799165D2FEE3FA94825755E0033E532/08A299C8E01C2F21482575EF001F
    FE6F/$FILE/CAP_592_e_b5.pdf

The Government of the Hong Kong Special Administrative Region. (2016). Quality Assurance Mechanism Underpinning QF.
    Retrieved from https://www.hkqf.gov.hk/en/qa/index.html

The Hong Kong Council for Accreditation of Academic and Vocational Qualifications. (2016). Accreditation. Retrieved from
    http://www.hkcaavq.edu.hk/en/services/accreditation

The Ministry of Education. (2016a). Accountability Pillar Evaluation Information. Retrieved from
    https://education.alberta.ca/accountability-pillar/how-the-pillar-works/

The Ministry of Education. (2016b). Alberta Education Retrieved from https://education.alberta.ca/alberta-education/

The Ministry of Education. (2016c). An Overview of the Accountability Pillar. Retrieved from
    https://education.alberta.ca/accountability-pillar/about-the-pillar/

The Ministry of Education. (2016d). Teacher Growth, Supervision and Evaluation Policy. Retrieved from
    https://education.alberta.ca/media/1626684/tgse.pdf

The Ministry of Education. (2016e). Teaching in Alberta: What you need to know. Retrieved from
    https://education.alberta.ca/teaching-in-alberta-what-you-need-to-know/

The National Archives. (2012). The Education (School Teachers' Appraisal) (England) Regulations 2012. Retrieved from
    http://www.legislation.gov.uk/uksi/2012/115/contents/made

The Scottish Government. (2013). Schools: Frequently Asked Questions. Retrieved from
    http://www.gov.scot/Topics/Education/Schools/FAQs

The Scottish Government. (2015). Summary statistics for schools in Scotland - No. 6: 2015 Edition. Retrieved from
    http://www.gov.scot/Publications/2015/12/7925/0

The University of Queensland. (2001). *School Reform Longitudinal Study*. Retrieved from Brisbane:
    https://www.researchgate.net/publication/37621426_The_Queensland_School_Reform_Longitudinal_Study

Thompson, J. (2010). *What impact to culturally competent teachers have on the social inclusiveness of their students?*. (Doctoral
    dissertation), Utah State University, Utah.

Thomson, S., De Bortoli, L., & Buckley, S. (2013). *PISA 2012: How Australia Measures Up*. Victoria, Australia: ACER.

Tillema, H., & Smith, K. (2007). Portfolio appraisal: In search of criteria. *Teaching and Teacher Education, 23*, 442-456.
    doi:10.1016/j.tate.2006.12.005

170

Toch, T., & Rothman, R. (2008). *Rush to Judgment: Teacher Evaluation in Public Education.* Retrieved from http://educationpolicy.air.org/sites/default/files/publications/RushToJudgment_ES_Jan08.pdf

U.S. Department of Education. (2011). *Preparing and Credentialing the Nation's Teachers: The Secretary's Eighth Report on Teacher Quality Based on Data Provided for 2008, 2009, 2010*. Retrieved from http://www2.ed.gov/about/reports/annual/teachprep/2011-title2report.pdf

U.S. Department of Education. (2012). The Federal Role in Education.   Retrieved from http://www2.ed.gov/about/overview/fed/role.html

U.S. Department of Education. (2016). Every Student Succeeds Act (ESSA).   Retrieved from http://www.ed.gov/essa?src=rn

van der Schaaf, M. F., Stokking, K. M., & Verloop, N. (2008). Teacher beliefs and teacher behaviour in portfolio assessment. *Teaching and Teacher Education: An International Journal of Research and Studies, 24*(7), 1691-1704. Retrieved from https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ806004&site=eds-live&scope=site

Victorian Auditor-General. (2010). *The Effectiveness of Student Wellbeing Programs and Services*. Retrieved from http://www.audit.vic.gov.au/publications/2009-10/290110-Student-Wellbeing-Full-Report.pdf

Victorian Institute of Teaching. (2016a). Moving to (full) registration.   Retrieved from http://www.vit.vic.edu.au/registered-teacher/moving-to-full-registration

Victorian Institute of Teaching. (2016b). The Victorian Teaching Profession Code of Conduct.   Retrieved from http://www.vit.vic.edu.au/media/documents/imported-files/spl/Code-of-Conduct-2015v2.pdf

Virginia Department of Education. (2011a). Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers. Retrieved from http://www.doe.virginia.gov/teaching/performance_evaluation/guidelines_ups_eval_criteria_teachers.pdf

Virginia Department of Education. (2011b). The Research Base for the Uniform Performance Standards for Teachers Retrieved from http://www.doe.virginia.gov/teaching/performance_evaluation/research_base_ups_teachers.pdf

Virginia Department of Education. (2011c). Virginia Standards for the Professional Practice of Teachers.   Retrieved from http://www.doe.virginia.gov/teaching/regulations/uniform_performance_stds_2011.pdf

Virginia Department of Education. (2012). Teacher Performance & Evaluation Retrieved from http://www.doe.virginia.gov/teaching/performance_evaluation/teacher/training_phase1/index.shtml

Virginia Department of Education. (2016a). Board of Education Teacher Performance Standards & Evaluation Criteria.   Retrieved from http://www.doe.virginia.gov/teaching/performance_evaluation/teacher/index.shtml

Virginia Department of Education. (2016b). College & University Programs.   Retrieved from http://www.doe.virginia.gov/teaching/educator_preparation/college_programs/index.shtml

Virginia Department of Education. (2016c). Enrollment and Demongraphics.   Retrieved from http://doe.virginia.gov/statistics_reports/enrollment/local_regional_schools_centers/2015_16_local_reg_sch_ctrs_a.pdf

Virginia Department of Education. (2016d). Licensure.   Retrieved from http://www.doe.virginia.gov/teaching/licensure/index.shtml

Virginia Department of Education. (2016e). Teacher & Administrator Preparation.   Retrieved from http://www.doe.virginia.gov/teaching/educator_preparation/index.shtml

Virginia Department of Education. (2016f). *Virginia Licensure Renewal Manual*. Richmond, Virginia Retrieved from http://www.doe.virginia.gov/teaching/licensure/licensure_renewal_manual.pdf.

Walla Walla Valley Education Association. (2015). Walla Walla Public Schools Teacher Professional Evaluation Process (TPEP). Retrieved from http://www.wwps.org/images/departments/personnel/tpep/TPEP_Handbook.pdf

171

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Washington State Legislature. (1991a). RCW 28A.405.100 Minimum criteria for the evaluation of certified employees - Revised four-level evaluation systems for classroom teachers and for principals - Procedures - Steering committee - Models - Implementation - Reports.   Retrieved from http://apps.leg.wa.gov/rcw/default.aspx?cite=28A.405.100

Washington State Legislature. (1991b). RCW 28A.405.120 Training for evaluators.   Retrieved from http://apps.leg.wa.gov/rcw/default.aspx?cite=28A.405.120

Washington State Legislature. (1991c). RCW 28A.405.220 Conditions and contracts of employment - Nonrenewal of provisional employees - Notice - Procedure.   Retrieved from http://apps.leg.wa.gov/rcw/default.aspx?cite=28A.405.220

Washington State Legislature. (1991d). WAC 392-191-007 Evaluation requirements.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191-007

Washington State Legislature. (1991e). WAC 392-191-010 Minimum evaluation criteria- Certificated classroom teachers.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191-010

Washington State Legislature. (1991f). WAC 392-191-010 Minimum evaluation criteria - Certificated classroom teachers.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191-010

Washington State Legislature. (1991g). WAC 392-191-025 Minimum procedural standards - Purposes of evaluation.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191-025

Washington State Legislature. (1991h). WAC 392-191-030 Minimum procedural standards - Frequency of evaluation.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191-030

Washington State Legislature. (1991i). WAC 392-191-035 Minimum procedural standards—Conduct of the evaluation.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191-035

Washington State Legislature. (1991j). WAC 392-191-045 Minimum procedural standards - Use of evaluation results.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191-045

Washington State Legislature. (1991k). WAC 392-191A-080 Minimum procedural standards - Conduct of the comprehensive evaluation.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191A-080

Washington State Legislature. (1991l). WAC 392-191A-090 Minimum procedural standards - Outcomes of the student growth rating.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191A-090

Washington State Legislature. (1991m). WAC 392-191A-200 Summative performance ratings - Descriptors.   Retrieved from http://apps.leg.wa.gov/WAC/default.aspx?cite=392-191A-200

Washington State Legislature. (2010). Engrossed Second Substitute Senate Bill 6696.   Retrieved from http://apps.leg.wa.gov/documents/billdocs/2009-10/Pdf/Bills/Senate%20Passed%20Legislature/6696-S2.PL.pdf

Washington State Teacher/Principal Evaluation Project. (2016a). Cel 5D+ Teacher Evaluation Rubric 2.0 At a Glance Version 1.1.   Retrieved from http://tpep-wa.org/wp-content/uploads/CEL-at-a-glance.pdf

Washington State Teacher/Principal Evaluation project. (2016b). eVAL System Features.   Retrieved from http://tpep-wa.org/resources/eval/features/

Washington State Teacher/Principal Evaluation Project. (2016c). Framework and Rubrics.   Retrieved from http://tpep-wa.org/the-model/framework-and-rubrics/

Watt, H. M., & Richardson, P. W. (2008). Motivations, perceptions, and aspirations concerning teaching as a career for different types of beginning teachers. *Learning and Instruction, 18*(5), 408-428.

Webster, C. A. (2010). Relating student recall to expert and novice teachers' instructional communication: an investigation using receiver selectivity theory. *Physical Education and Sport Pedagogy, 15*(4), 419-433.

Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. *New Teacher Project*.

172

Teacher Effectiveness Systems, Frameworks and Measures: A Review

Wenner, J., & Campbell, T. (2016). The theoretical and empirical basis of teacher leadership: A review of the literature. *Review of Educational Research*. doi:doi: 10.3102/0034654316653478

White, S. A. (2013). *A Comparison of Teacher Evaluation, Student Surveys and Growth Scores to Identify Effective Teaching Traits.* ProQuest LLC. Retrieved from https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED561474&site=eds-live&scope=site

Whitehurst, G., Chingos, M., & Lindquist, K. (2014a). Evaluating teachers with classroom observations. *Brown Center on Education Policy: Brookings Institute*.

Whitehurst, G., Chingos, M., & Lindquist, K. (2014b). *Evaluating Teachers with Classroom Observations: Lessons Learned in Four Districts*. Retrieved from https://www.brookings.edu/wp-content/uploads/2016/06/Evaluating-Teachers-with-Classroom-Observations.pdf

Wilkins, E. A., Shin, E.-K., & Ainsworth, J. (2009). The effects of peer feedback practices with elementary education teacher candidates. *Teacher Education Quarterly, 36*(2), 79-93.

Witt, P., Wheeless, L. R., & Allen, M. (2004). A meta-analytical review of the relationship between teacher immediacy and student learning. *Communication Monographs, 71*(2), 184-207.

Yuan, K., Engberg, J., Kaufman, J., Hamilton, L., Hill, H., Umland, K., & McCaffrey, D. (2014). Using Anchoring Vignettes to Calibrate Teachers' Self-Assessment of Teaching. *Society for Research on Educational Effectiveness*.