Development of a Teaching Performance Assessment in Australia: What Did We Learn?

Rebecca Spooner-Lane
Queensland University of Technology
Tania Broadley
University of Canberra
Elizabeth Curtis
University of Southern Queensland
Peter Grainger
University of Sunshine Coast

Abstract: Following increasing criticism of the variability in graduate teachers' readiness to enter the profession, the Australian Institute for Teaching and School Leadership (AITSL) introduced a program accreditation requirement that all initial teacher education (ITE) providers must implement a Teaching Performance Assessment (TPA) in the final year of their teacher education programs. AITSL were not prescriptive in how ITE providers must meet the program standard which has resulted in 12 TPAs being implemented across 42 ITE providers. This paper outlines the development and implementation of one endorsed TPA designed to measure the readiness of graduating teachers, whilst taking into consideration the learnings from well-known TPAs and our own experiences. With this being one of the earlier unfunded TPAs in Australia to have been approved through the accreditation endorsement process, the paper offers some insights into meeting the additional accreditation program requirements and raises some longer-term considerations associated with implementing TPAs.

Introduction

The professional standards of teacher education programs are critical to quality assuring initial teacher education (ITE) programs. The accreditation of teacher education programs in Australia relies on each professional standard being taught, practiced, and assessed multiple times throughout an ITE degree. However, increasing scrutiny and rising concerns about the variability in Australian teaching graduates' readiness for teaching practice led to the introduction of a capstone Teaching Performance Assessment (TPA) which also features in several ITE degrees internationally (Allard et al., 2014; Cochran-Smith et al., 2013; Darling-Hammond et al., 2013; Stacey et al., 2020). The TPA is a capstone, performance assessment, aligned to the professional standards for teachers, and tied to national teacher education program accreditation. The TPA is a benchmark for

> *measuring graduates' ability to demonstrate that they have the requisite knowledge, skills and practices to (1) plan and deliver a series of lessons, (2) monitor and assess student learning and to (3) reflect on the impact of their teaching on the learners in their classroom. (Mascadri et al., 2022, p. 2)*

Since 2019, all pre-service teachers in Australia must pass a TPA to graduate. The introduction of TPAs has brought about a shift in focus from professional standards as inputs used to inform program development, to standards as outputs assessed in actual classroom teaching practice (Adie & Wyatt-Smith, 2021). The decision to implement a TPA resulted from a set of recommendations following the formation of a Teacher Education Ministerial Advisory Group (TEMAG) funded by the federal government to review ITE programs. The report *Action Now: Classroom Ready Teachers* (TEMAG, 2015) highlighted inconsistencies between ITE programs in the assessment of pre-service teachers against the Australian Professional Standards for Teachers (APSTs) at program completion. The report called for a consistent and transparent graduate assessment at an agreed benchmark of Graduate Teacher Standard to be a key feature of profession entry requirements. The foundations of this recommendation were primarily influenced by the well-known Performance Assessment for California Teachers (PACT) and the Education Teacher Performance Assessment (edTPA) in the United States which are linked to teacher licensure and have been well researched (Nelson et al., 2014; Peck & McDonald, 2013).

As was the case in Australia, the advent of TPAs in the United States came about due to increasing calls to professionalise the field. Conceived as a means to develop more positive perceptions about the teaching profession (Greenblatt, 2018), the aim of the TPA was to address the rising "problem of teacher education" (Cochran-Smith et al., 2013, p. 16). The wide variance in approach and a "lack of accountability and standardization of expectations across programs" (Cochran-Smith et al., 2013, p. 16) was considered problematic. Thus, TPAs emerged as a means to authentically measure 'teacher readiness' by providing a consistent assessment of teaching practice (Dover & Shultz, 2016; Greenblatt, 2018). In the next section, the background to the development of multiple TPAs in Australia, only two of which received finding via the Australian Institute for Teaching and School Leadership (AITSL) grant scheme, will be explained.

AITSL Accredited TPAs

AITSL is funded by the Australian Government, to lead the Commonwealth, state and territory governments to promote excellence in teaching and school leadership. Through the release of AITSL's (2015) ITE Program Standards, the inclusion of Standard 1.2 required ITE providers to include an assessment of classroom teaching performance, across a sequence of lessons that reflects the range of teaching practice. With the release of the program standards, higher education providers of ITE began to discuss the impact of the requirement for a capstone assessment in terms of development and implementation. In May 2016, the teacher regulation authority Queensland College of Teachers (QCT) who accredit ITE programs in the state of Queensland, called together ITE providers for a presentation of a concept TPA developed by the Institute for Learning Sciences and Teacher Education at Australian Catholic University (ACU) and the QCT. Providers were then invited to engage in a pilot of TPA implementation with ACU. Given Queensland ITE providers were scheduled to be among the first providers in Australia to undertake accreditation according to the new program standards, which included providing a long-term plan to collect and show evidence of the quality required of graduates to be considered 'ready' for entry into the profession, it was critical that ITE providers had a clear plan for development and implementation of a TPA. Later that same year, in October 2016, AITSL developed a national grant process to offer financial support for groups of ITE providers to "stimulate the development of TPAs in line with Program Standard 1.2 that will be used across multiple ITE providers" (AITSL, 2017a). In 2016, the Queensland University of Technology (QUT) led an expression of

interest for the tender which included five ITE providers across three states and territories. The application was one of six expressions of interest that were received for consideration, with only two out of six expressions of interest taken to further development and awarded funding in 2017. The successful consortia were led by the Institute for Learning Sciences and Teacher Education at ACU resulting in the Graduate Teacher Performance Assessment (GTPA) and the Melbourne Graduate School of Education at Melbourne University resulting in the Assessment for Graduate Teaching (AfGT).

With QUT moving into accreditation of three postgraduate programs in June 2017, the development work for the TPA continued. A small group of teacher educators formed a community of practice (Lave & Wenger, 1991) to develop and implement the TPA assessment tool and assessment rubric. A community of practice recognises that learning is a social process situated in a cultural and historical context (Wenger, 1998). The members comprised a diverse set of teacher educators with teaching experience across the three preservice teacher degree programs (early child, primary and secondary education). The community of practice enabled the members to collaboratively design a TPA fit for purpose. Capitalising on individual strengths, members of the community undertook different tasks but came together to share and refine ideas, to discuss challenges and difficulties, and problem solve. In this article, the authors share their experience of contributing to the design and implementation of the Quality Teaching Performance Assessment (QTPA).

At the time, there was a high level of uncertainty about the TPA in Australia (e.g., what the TPA should look like and how it should be assessed), which was evolving alongside the accreditation of ITE programs that were being assessed at the same time. In Queensland, the optimal intention was to accredit ITE providers through the use of one TPA (the newly funded consortium developing the GTPA), however, AITSL was clear at the national level that some individual providers would have the capability, capacity and resources to develop their own TPAs. A context statement from AITSL indicated,

> Throughout 2017, AITSL is funding two Consortia of ITE providers to develop, trial and/or validate TPA tools under the TPA Grant Program. All ITE providers are required to have a TPA in place for their 2018 graduate cohort. As such there will also be non-Consortia TPAs being developed, trialled and implemented. (AITSL, 2017b, p. 15)

Darling-Hammond and Snyder (2000) argued that,

> If such [teacher performance] assessments are treated largely as add-ons at the end of a course or program rather than as integral components of ongoing curriculum and instruction, the time, labor, and expense of conducting them could be overwhelming within the institutional constraints of teacher education programs. (p. 527)

It is plausible that the significant investment of resources required in the development and implementation of a TPA (including program redesign, assessment expertise, school partnership development, staff professional development and additional administrative resources and on-going assessor training) may have precluded some smaller ITE providers from developing their own TPA. Many providers, looking for certainty in the accreditation process, joined AITSL's funded and already endorsed TPAs.

The QTPA was the first TPA in 2019 to be endorsed by AITSL's Expert Panel as an individual submission from an ITE provider in Australia. However, endorsement was only given on the understanding that further improvements to the TPA were planned and completed. These included: a) ensuring that training and moderation processes ensured reliable and consistent judgements between panels and different pre-service teacher programs over time; b) ensuring appropriate processes were put in place to periodically recalibrate TPA assessors' ongoing consistency; c) a review of inter-rater reliability to determine whether the

training had been effective and d) evidence of cross-institutional moderation for providers within a consortium with the expectation that users of the TPA submit a selection of TPA submissions for purposes of moderation and data analysis. Endorsement of the QTPA was achieved on the fourth submission for review by AITSL's Expert Panel. However, it is important to note that there were some conflicting judgments and suggestions from review to review. This may be because members of the panel were not consistent from review to review and there was a lack of published guidelines at the time as to what was expected of ITE providers designing their own TPA.

Fortunately, endorsement of the QTPA resulted in other ITE providers expressing interest to form a consortium. The QTPA is implemented across two Australian States (Queensland and Western Australia) and four individual ITE providers (QUT, University of Sunshine Coast, Edith Cowan University, Murdoch University). The members are responsible for further refining the TPA tool and assessment rubric and engage in cross-institution moderation to ensure consistency in assessing preservice teachers at a Graduate Teacher Standard.

Using a self-study method (Loughran, 2006), the authors reflect on some of the key decision-making points involved in the development and implementation of an accredited TPA. The decisions underpinning the development of the QTPA are primarily informed by learnings and shortcomings from the two most commonly researched TPAs in the United States. Adopting a reflexive approach enabled the authors to reflect on highlights and criticisms of the edTPA and PACT, in an attempt to further strengthen the QTPA process.

Developing the TPA

A TPA is used to make a judgement as to whether a graduating teacher is ready to enter the teaching profession. AITSL's Program Standard 1.2 stipulated that a TPA must: a) be a reflection of classroom teaching practice including the elements of planning, teaching, assessing and reflecting; b) be a valid assessment that clearly assesses the content of the Graduate Teacher Standards; c) have clear, measurable and justifiable achievement criteria that discriminate between meeting and not meeting the Graduate Teacher Standards; d) be a reliable assessment in which there are appropriate processes in place for ensuring consistent scoring between assessors; and e) include moderation processes that support consistent decision-making against the achievement criteria.

While traditional assessments of pre-service teacher competency have been criticised for lack of authenticity, there is growing evidence that performance assessments better evaluate teaching practices (Darling-Hammond & Snyder, 2000; Weir, 2009). The QTPA is an authentic assessment of teacher readiness as it evaluates tasks regularly performed by a teacher (planning, teaching, assessing, reflecting) and combined with the oral component, demonstrates teachers' ability to communicate their knowledge and teaching skills that teachers use every day. Furthermore, these knowledge and skills align to the APSTs at a Graduate Teacher Standard.

As limited research on the development of TPAs in Australia had been published at the time of development (e.g., Deakin Authentic Teacher Assessment; Dixon et al., 2011; Deakin University, 2012), the development team relied on evidence-based research and practice from primarily the United States to inform its design. In the U.S., both the edTPA and PACT focus on a portfolio of teaching practices designed to assess pre-service teachers' classroom readiness by reviewing planning documentation, videotapes of classroom practice, evidence of student work and learning, commentary on student work and reflections (Darling-Hammond, 2012; Pecheone & Chung, 2006; Stillman et al., 2013). These are

organised into four categories of teaching: planning, instruction, assessment, and reflection (Pecheone & Chung, 2006), all characteristics of authentic teaching practice. QUT also selected a portfolio assessment design comprising four components. The four components are described in the following section.

Specifically, Component A comprises a personal teaching statement that requests pre-service teachers to reflect on their current teaching beliefs and practices across three domains of the APSTs (professional knowledge, professional practice and professional engagement), and articulate how their teaching beliefs and practices have been informed by relevant research and/or theory. Component B requires pre-service teachers to collect and interpret assessment data (formative and/or summative) to determine student learning and achievement levels prior to documenting evidence of their planning and delivering a sequence of four to six lessons within a school environment. For Component C, pre-service teachers provide a written reflection of their impact on student learning upon completion of teaching their lesson sequence. Finally, Component D is a 15 minute oral in which pre-service teachers articulate to a panel of assessors their thinking and decision making in planning, teaching and evaluating their impact on student learning. The oral concludes with the candidate preparing a response to two reflective questions aligned to professional standard 6 (engaging in professional learning) and 7 (engaging with colleagues, parents/carers and the community) posed by the assessment panel.

Darling-Hammond and Snyder (2000) outline five aspects of authentic assessments of teaching: (1) assessments sample the actual knowledge, skills, and dispositions desired of teachers as they are used in teaching and learning contexts, rather than relying on more remote proxies; (2) assessments require the integration of multiple kinds of knowledge and skill as they are used in practice; (3) multiple sources of evidence are collected over time and in diverse contexts; (4) assessment evidence is evaluated by individuals with relevant expertise against criteria that matter for performance in the field; and (5) the assessment includes opportunities for learning and practicing the desired teaching tasks. Weir (2009) defined authentic assessment as consisting of five aspects, the first two of which are core to her model and the TPA. Specifically, she referred to authentic assessment as being *connected* to the curriculum and to everyday lives and real-world contexts. Secondly, she characterised authentic assessment as being *performance-oriented,* and encourages performance of learning in a real context in front of a "real" audience. The QTPA is well positioned to meet the aspects of authentic assessment as detailed in the following section.

The work of Linda Darling-Hammond and her colleagues in the United States focused on "high quality preparation, evaluated authentically through performance assessments that both develop and measure beginning teacher effectiveness" (2012, p. 9). That concept of further developing as well as assessing graduates was important to the community of practice. The assessment tool we developed prioritised building the teaching capacity of our pre-service teachers and benefitting them in terms of readiness for teaching and working professionally with future colleagues, parents and families, students and communities. This meant making the assessment applicable to the real-world context and practice of teaching where there would be opportunities for our pre-service teachers to develop and examine "thinking and actions in situations that are experience based and problem oriented and that include or simulate actual acts of teaching" (Darling-Hammond & Snyder, 2000, p. 524). Research surrounding PACT highlighted the risk that pre-service teachers "might have to choose between being a 'good student' or 'task completer' in order to achieve a high score...versus being successfully and critically engaged in learning to teach" (Reagan et al., 2016, p. 7). Despite the TPA being a summative assessment and not a formative assessment tool for learning, we wanted to ensure the actual process of engaging in the TPA was educative and empowering for pre-service teachers. We aimed to design a TPA where pre-

service teachers felt equipped to hold productive professional conversations about their teaching practices with their supervising teacher, provide students with more usable feedback and enhance their capacity to teach by changing the focus from themselves to their students as a result of fulfilling the TPA requirements (Darling-Hammond, 2012).

In our context, the use of video recordings was not suitable for three key reasons. First, the issue of ethics and gaining consent for pre-service teachers work with children and adolescents aged birth to 18 would be problematic. Second, if a pre-service teacher is placed in a more challenging school context, the behaviour of their learners may be unpredictable. A potential concern of pre-service teachers who submit real-time, real world videoed teaching segments, is that an assessor may not understand why the videoed teaching segment does not feature well-behaved, engaged students (Parkes & Powell, 2015). Third, research has demonstrated that the assessor of the video is the only one who interprets what is seen in the video and then makes a judgement; the performance is either satisfactory or it is not. The only part that is judged is 'what is seen' meaning the pre-service teacher is "subjected to anonymous and invariant evaluative procedures, without reference to the development of teachers/learners" (Caughlan & Jiang, 2014, p. 383).

Evidence of planning documentation is crucial in the inclusion of any TPA. Through the planning of lessons, pre-service teachers demonstrate the ability to organise instruction that promotes effective learning for all students in diverse contexts; and how teaching outcomes are analysed to optimise student success in future teaching and learning tasks (Darling-Hammond, 2012). The QTPA requires pre-service teachers include planning for a sequence of four to six lessons to demonstrate how they considered the learning environment, and the learners within, in designing the sequence. In reflecting on the learners, we expected pre-service teachers to show consideration of school and classroom context and student diversity. Research emerging from PACT found that one of the most valuable elements of the TPA was this aspect where evaluating what students were learning, or not learning, and reflecting and modifying lesson sequences actually continued into the candidates' teaching practices after graduation (Stewart et al., 2015). The inclusion of annotated student work samples to demonstrate the links between what the student needed to do, the teaching enacted and what the student had achieved after the teaching were necessary and vital components of the QTPA.

A pilot of the written components of the QTPA was first conducted in September 2017 with the Bachelor of Education students across our Early Childhood, Primary and Secondary programs (n=400). Important modifications were made to the QTPA as a consequence. For instance, a word limit was imposed for the personal teaching statement (Component A) because there was significant variability in the length of statements. Pre-service teachers were also asked to use the three professional domains of teaching (professional knowledge, professional practice and professional engagement) as sub-headings and to cite the corresponding APSTs throughout their statement. This was designed to assist assessors to determine whether candidates understood the intention and meaning of the professional domains and professional standards. Component B initially consisted of a table where pre-service teachers responded to each APST the corresponding descriptors with one to two examples of evidence that demonstrated the achievement of that standard. However, the table used for Component B for evidencing the APSTs was cumbersome and demonstrated superficial learnings by pre-service teachers and was removed in the next iteration. Component B was revised to reflect the full planning, teaching and assessment cycle with a structured template to support graduate teachers documenting their work. Candidates were also provided with a structured template for Component C to support presenting their impact on student learning and their written reflection.

The oral component of the QTPA was highly significant and remained central in its development (Mascadri et al., 2022). In order to adhere to the principles of authentic assessment, particularly in terms of a real-world audience (Weir, 2009), we wanted to give pre-service teachers the opportunity to clearly explain the planning of their lessons and pedagogical decision making to the assessors. As Darling-Hammond and Snyder (2000) noted when discussing authentic assessment in relation to the TPA that in order to understand a classroom event,

> *one needs relevant information about students and their prior learning and approaches as well as about the event itself and about the teacher's decision-making processes i.e. what goals he/she is trying to achieve and what aspects of content, student needs, and classroom or community context he/she is taking into account. (p. 528)*

The deliberate decision to include an oral component so that pre-service teachers could articulate their teaching and learning practices was important. Teachers regularly engage in professional conversations around their work and that of their students' learning and outcomes. Furthermore, the oral component may reduce susceptibility of assessors to privilege good writers (Clayton, 2018). When the impact of the pre-service teacher teaching on student outcomes is examined, we are mindful that the competency of the pre-service teacher is not solely marked on how successful the student has been in representing his or her teaching competency through annotations on student work samples. Assessors of the QTPA are looking for the preservice teachers' analysis and reflection of their own teaching and learning across the lessons in supporting student learning.

Research on portfolio assessment has demonstrated that it can often be inequitable (Meeus et al., 2009) due to the significant knowledge and skills required for its successful compilation. The research conducted by Campbell et al. (2016) in the United States found that TPA candidates spent an average of 22 hours compiling their TPA and assessors spent three hours scoring each assessment (Parkes & Powell, 2015). Taking this into consideration, we provided a structured template to help reduce the workload for preservice teachers and to make marking more efficient for assessors. Thus, in our first full iteration of the QTPA in September 2018 with 300 pre-service teachers we found that the templates for Components B and C clarified the salient expectations of the assessment, provided equity in terms of visual aesthetics and use of technology, and ease for the assessors in terms of knowing where to locate the appropriate information. We developed a set of pre-service teacher training materials for teacher educators to unpack the core features of each assessment component and guidelines on how to complete the associated templates. This meant two things. First, regardless of which lecturer a pre-service teacher had or which institution they attended there is guaranteed equity in the training information provided. Second, in not providing a specific modelled response for each component, we allowed for some creativity and variation, avoiding the likelihood of cookie cutter responses.

The QTPA was positioned as part of the usual teaching practice that teachers engaged in on a daily basis and not as an add on feature of their final teaching placement. Research in the USA demonstrated that some candidates viewed the TPA as extra to their teaching and noted it detracted from their teaching: "I was so focused on completing the assignment with a passing score that I was unable to fully devote myself to teaching" (Campbell et al., 2016, p. 66). We explained to pre-service teachers that the requirements of the QTPA were what they would already be doing on their final placement on a daily basis as part of the teaching and learning cycle including submitting lesson plans and evaluating student work that was a part of that teaching cycle, not in excess.

The QTPA is designed as a final semester capstone experience that is connected to actual teaching practice in a teaching context and hence is genuinely 'authentic' (Weir, 2009)

and requires discussion and reflection on the teaching cycle of planning, teaching, assessing and impact of teaching on student learning. Whilst the assessment requires some aspects of a traditional portfolio approach where pre-service teachers show and discuss their teaching, evidence of student learning, the feedback they have given to student work, and students' responses to these teaching efforts (Darling-Hammond & Snyder, 2000), it goes beyond a mere collection of artefacts and requires the pre-service teachers to curate the portfolio around one specific teaching and learning sequence of four-six lessons. The pre-service teachers must first demonstrate an understanding of the school context and student cohort through talking to their supervising teacher and collecting assessment data on students' achievement levels to determine the 'what' and 'how' of their teaching and learning sequence. A requirement of the QTPA in this component is for the pre-service teacher to clearly articulate the interpretation of the data, but more importantly how this knowledge and understanding of their learners' needs informs their teaching practice and is usually represented in an annotated table. Rather than disparate, seemingly unrelated items of teaching practice curated in a portfolio over many weeks or months of teaching, the QTPA requires all elements to be related to one teaching and learning sequence of four-six lessons. This allows assessors to examine a chain of events and thinking, to see the quality of deliberation and the pedagogical decisions made by the candidate as well as evaluate the appropriateness of actions taken (Darling-Hammond & Snyder 2000). When portfolios are used hand in hand with an oral component they can be "beneficial to the validity of the assessment [as] the teacher educator then has additional information at his or her disposal, including a better appreciation of the student's overall progress" (Meeus et al., 2009, p. 404). We believe the oral component will ensure that pre-service teachers not only learn by engaging in the assessment, but that they are later able to enact these practices in their daily teaching (Darling-Hammond et al., 2013).

The oral component of the QTPA sets it apart from written only TPAs currently in use. It allows for a more just and authentic experience as well as providing pre-service teachers with powerful preparation for job interviews which typically require that they verbally articulate their teaching beliefs and practices (Huxham et al., 2012), their impact on students' learning, and the reasons behind their teaching content and pedagogical choices. An oral assessment is also a method for understanding pre-service teachers' knowledge and critical reasoning (Gent et al., 1999). Oral examinations may prevent "superficial regurgitation" (Huxham et al., 2012, p.126) found in written assessments and can strengthen academic integrity (Joughin, 1998) as pre-service teachers must explain their own understanding in their own words. We believe that incorporating an oral is appropriate for graduate teachers as the education of students relies primarily on dialogue and learning conversations (Huxham et al., 2012).

The Assessment

The TPA was introduced to make ITE providers accountable for producing 'classroom ready' graduate teachers. Following the development of the assessment tool, attention turned to planning a rigorous assessment criteria and moderation process.

Developing the Assessment Criteria

An important aspect of the quality assurance process of the QTPA was establishing assessment criteria that measured the TPA outcomes using the Graduate Teacher Standards as

benchmarks for making judgements. Research suggests however, that objective portfolio assessment is challenging (Joosten-ten Brinke et al., 2010). Assessors may omit steps in the assessment process or evaluate the portfolio without adhering to the assessment criteria (Joosten-ten Brinke et al., 2010). Another limitation is that preservice teachers' TPAs are not necessarily assessed by someone with expertise in their teaching area. It is therefore imperative that each assessment criterion is clearly articulated. Van der Schaff et al. (2005) argue that teacher educators' judgements are influenced by previous assessment ratings and experiences. Furthermore, assessors may give attention to features that are not explicitly being assessed or give greater weighting to some assessment features over others (Newton & Meadows, 2011). Sadler (1985) also notes that the meaning of words may be interpreted differently by different assessors. To enhance the quality of the assessment criteria, advice was sought from teacher educators with expertise in assessment both internal and external to the institution about the wording of the assessment criteria and the standards descriptors used to describe the behaviours at each of the standards. This process contributed to tightening the alignment between the assessment task requirements and the performance criteria. This meant that the wording of the assessment criteria mirrored the wording used in the pre-service teachers' task requirements to prevent misconceptions and misunderstandings. For example, for the personal teaching statement (Component A) pre-service teachers must explicitly demonstrate the main intent of the professional domain, and convey a coherent alignment of valid theory/relevant research to inform beliefs and an aligned example of how beliefs are enacted in practice. This wording is also used on the assessment criteria. Feedback was also sought from assessors following the first full implementation of the QTPA to further sharpen the language clarity of the assessment criteria. This process was essential in reducing overlap between criteria and unclear wording (Wyatt-Smith & Klenowski, 2013).

TPAs are assessed in different ways in Australia depending on the assessment criteria established. For the QTPA, the decision to pass or fail a student on the QTPA was based on a cut-score (Liu & Liu, 2008). Following the scoring of the TPA, an experienced statistician led a group of assessors through a cut-score process. The assessors were selected due to their expertise in teaching into the early childhood, primary and secondary teacher education programs and familiarity and understanding of the Graduate Teacher Standards. A borderline group method was used to determine the cut-score (Downing et al., 2006). This method required a group of assessors to divide the QTPA assessments into three groups – satisfactory, borderline and unsatisfactory. The borderline group were defined as those performances where the assessors were unsure whether the preservice teacher was at Graduate Teacher Standard or not. When the assessors were satisfied with the three groupings based on the preservice teacher's TPA performance, the designated cut-score was determined by the mean total score of the borderline group. Preservice teachers' TPAs were rated as satisfactory if they met or scored above the cut score. TPAs were rated as unsatisfactory if they scored below the designated cut-scores.

Assessor Training

The primary goal was to establish a community of assessors that come together with a shared understanding of the assessment criterion and Graduate Teacher Standards to determine the graduate teachers' achievement on the TPA. Drawing upon the research of Adie et al. (2013), the authors took the position that the TPA assessors must be teacher educators who are responsible for teaching pre-service teacher education. Ensuring the quality of assessor judgments required deconstructing the QTPA assessment tool, the assessment criteria and stepping assessors through a rigorous step-by-step moderation

process. This is because achieving consistency in assessor judgments improves when assessors adhere to the assessment criteria, determine the credibility of the evidence provided and have shared understanding of the Graduate Teacher Standards (Bloxham et al., 2016).

A full day of assessor training and refresher training is provided twice per year to ensure that all teacher educators who teach into the pre-service teacher education programs understand the QTPA assessment task requirements, the performance criteria and corresponding Graduate Teacher Standards. Assessors are provided with an assessor manual comprising QTPA samples depicting satisfactory, borderline and unsatisfactory TPAs.

During the training, facilitators unpack each QTPA component using samples of pre-service teachers' TPAs as a reference. Assessors carefully examine benchmarked QTPAs and work in small groups to develop a strong understanding of how the standard of Graduate Teacher is applied. Furthermore, performance standard guidelines are used to help support assessors' judgments. Assessors are taught how to assess the variations in the quality of work provided by pre-service teachers. Through the training process, assessors come to realise that pre-service teachers' final professional experience placements are diverse and context-based and therefore no two QTPAs will look exactly the same (Sadler, 2013). However, templates completed by pre-service teachers for Components B and C assist the assessors to locate the evidence to make their judgements. The templates provide some standardisation in QTPA process (Herbert et al., 2014) and ensure assessors appropriately assess the correct evidence for each assessment criterion. For the QTPA, assessors are allocated one hour per student for marking (i.e., 30 minutes for marking the written components, and 30 minutes for the oral component).

Assessors participate in a simulated assessment process during training. Assessors form groups of three and follow the assessment process of marking the written components of independently before coming together to assess the oral component. The assessors then engage in a moderation conversation to establish the final assessment standard of satisfactory or unsatisfactory. At the end of the simulated activity, assessors receive feedback from the training facilitator to assist them calibrate their marking and achieve consistency in judgements. The facilitator helps assessors to identify their personal and professional biases (e.g., how curriculum should be taught, tendency to mark poor spelling and grammar harshly) as assessors explaining their marking decisions out loud.

Assessment Moderation

Teacher educators may bring varying assessment expertise, experience and ideas about graduate expectations (Sadler, 2010; Wyatt-Smith & Klenowski, 2013) and consistency in assessor judgements is dependent on rigorous training and equitable and fair assessment moderation processes. Sadler (2011) maintains that, through a social moderation process, ideas about specific standards can be "clarified, refined and transformed – and in the process shifted from essentially private knowledge to collegially held knowledge" (p. 5). Moderation of QTPAs involves three assessors (two assessors and one Chairperson) meeting to discuss their judgments (Watty et al., 2014). The use of a three-person moderation panel provides pre-service teachers with some confidence that their QTPA has been scored fairly and accurately. In higher education, assessments are primarily scored by one assessor and a sample of scored assessments may be verified by one other assessor. The QTPA design team argued that due to the high-stakes nature of the assessment, having three assessors in agreement could be beneficial in enhancing candidates' trust in the assessment process. The three assessors independently mark the written components of the QTPA prior to coming together to view the pre-service teacher oral presentation. The Chairperson invites the two

assessors to provide their assessment judgement for each component item. When both assessors agree on the scoring, the score is accepted and recorded. However, where there is disagreement on the scoring, the Chairperson will share their score and the three assessors will discuss the item, using evidence from the TPA to justify their judgment. Moderation is completed when agreement is reached. Once the moderation panel have evaluated each item, the final total score is recorded and an overall grade of satisfactory/unsatisfactory is determined based on the cut-score. In the rare instance when agreement cannot be reached, the assessment panel seek advice from a more experienced QTPA supervisor who oversees the panel process.

Cross-institution Moderation

Cross-institution moderation occurs in a similar way to the ITE providers' internal moderation process and is an essential component of the quality assurance process (Bloxham et al., 2016). An important aspect of scoring the QTPA is ensuring partnering universities within the consortium are assessing individual graduates at a comparable passing standard (Coates, 2010). QTPA partnering universities participate in a cross-institution moderation process. De-identified QTPA samples of each performance standard (satisfactory, borderline and unsatisfactory) from each university are submitted and are assessed by another university. Three members from each university blind mark TPAs from another university. Consortium members then meet face-to-face or virtually to moderate the assessments together. This moderation process occurs at the end of each university semester and contributes to a shared understanding of the assessment criteria and consistency in judgements between universities. Consortium members engage in moderation until agreement is reached. Moderation notes are recorded and key learnings are used to improve future assessor training and/or pre-service teacher preparation for the QTPA. These collaborative and transparent moderation discussions are important for ITE providers to learn from each other and to make agreed revisions to the assessment tool and criteria or TPA training (Zahra et al., 2017), further strengthening the quality of teaching and learning provided by the ITE programs.

Implications of TPAs

The design of the TPAs in Australia are multi-faceted and are informed by two recognised and well researched TPAs in the U.S. Similar to the edTPA and PACT, TPAs are linked to preservice teacher graduation and are aligned to a set of professional teacher standards at Graduate level. TPAs incorporate authentic teaching tasks including planning documentation in the form of lesson plans, implementation of a teaching sequence, assessment of student learning, and a reflection of teaching impact on student learning. However, there are also some noteworthy differences of TPAs in Australia when compared to the edTPA and PACT.

In Australia, ITE providers practice and assess the components of the TPA throughout education programs prior to pre-service teachers undertaking their TPA in the final semester of their program. Furthermore, assessors often teach the pre-service teachers about the TPA requirements within their own courses. Over time, pre-service teachers and assessors develop a deep and shared understanding of the TPA requirements.

Pre-service teachers in Australia undertake professional experience placements in diverse school contexts (e.g., urban, rural, regional, and remote schools) with differing

pedagogical approaches. A strength of pre-service teachers fulfilling the requirements of the TPA is the focus on planning and implementing a teaching and learning sequence that is responsive to the strengths and needs of the learners (Aitken et al., 2013). Unlike TPAs in the U.S. (Parkes & Powell, 2015), some teaching regulation authorities across different education jurisdictions in Australia do not permit pre-service teachers to collect video evidence of their teaching in practice. ITE providers can therefore ensure they meet the additional requirements of their teaching regulation authority when designing their TPA. For the QTPA, evidence is provided in the submission of detailed lesson plans and lesson reflections detailing how pre-service teachers adapted their teaching practices to meet the needs of all learners in their context. However, the teaching contexts widely vary and TPAs in Australia must accommodate these differences. For example, some pre-service teachers deliver their lesson to an entire class or a small group.

ITE providers in Australia are responsible for determining whether pre-service teachers satisfactorily pass or fail the TPA. For the QTPA, teacher educators who teach into the pre-service teacher programs are trained to assess the TPA. It is not outsourced to an external education provider such as Pearson where tensions concerning the financial costs to candidates, questions regarding assessors' expertise and backgrounds, ownership of data, and issues of assessor fatigue have been raised (Parkes & Powell, 2015). However, it does mean that ITE providers in Australia assume the additional costs associated with implementing a TPA in the final year of their education programs and resourcing ongoing revisions and testing to further enhance the reliability and validity requirements of TPAs in order continue to meet accreditation program standards.

The administrative costs and logistical load for ITE providers implementing TPAs in Australia vary widely. For example, the administration of the oral component for the QTPA is both labour and resource intensive and the substantial costs are absorbed by consortium members. The QTPA requires a sophisticated learning management system to manage online scoring and moderation processes. Similarly, an online booking system is necessary for pre-service teachers to book their oral presentation. Timetabling and staffing oral panels can be an issue and the frequency of oral panels has meant that assessors are required to sit on three assessment panels throughout the course of a year. The assessment load of assessors has to be monitored by the ITE provider to minimise assessor fatigue. For example, is encouraged that QTPA assessors spread out their required panel work over the assessment periods. The costs associated with implementing TPAs has meant that smaller, regional ITE providers may be disadvantaged when it comes to what is possible when implementing a TPA for accreditation requirements.

In conclusion, the introduction of the TPA as a capstone assessment is still a relatively recent educational reform in Australia. Every ITE provider is required to implement a TPA (AITSL, 2017b) and all ITE providers must commit to resourcing ongoing future research to investigate the reliability and validity of the TPA as a measure graduate teachers' classroom readiness, to maintain accreditation program standard requirements. ITE providers, who took the opportunity to design and implement their own TPA in Australia, did so in a relatively short period of time with limited guidance and understanding of the long-term financial and resource implications. TPAs are a mandatory requirement of meeting accreditation requirements, however, further refinements to the assessment tool, its labour-intensive assessment and moderation processes, and the need for longer term research, raises the issue of the long-term sustainability of ITE providers resourcing TPAs in their current form. It remains to be seen whether in the future, a national TPA will be implemented, outsourced to an external education provider, with graduate teachers wholly or partly responsible for the financial costs of TPA submission.

References

Adie, L., Lloyd, M. & Beutel, D. (2013). Identifying discourses of moderation in higher education. *Assessment and Evaluation in Higher Education, 38(8)*, 968-977. https://doi.org/10.1080/02602938.2013.769200

Adie, L. & Wyatt-Smith, C. (2021). The conceptualisation of a teaching performance assessment: Designing for evidence of graduate competence. In C. Wyatt-Smith, L. Adie, & J. Nuttall (Eds.). *Teaching performances assessments as a cultural disruption in initial teacher education. Standards, evidence and collaboration* (pp. 19-37). Springer. https://doi.org/10.1007/978-981-16-3705-6_2

Aitken, G., Sinnema, C., & Meyer, F. (2013). *Initial teacher education outcomes: Standards for graduating teachers. A paper for discussion*. The University of Auckland.

Allard, A. C., Mayer, D., & Moss, J. (2014). Authentically assessing graduate teaching: outside and beyond neo-liberal constructs. *Australian Educational Researcher, 41*(4), 425-443. https://doi.org/10.1007/s13384-013-0140-x

Australian Institute for Teaching and School Leadership. (2015). *Accreditation of initial teacher education programs in Australia.* https://www.aitsl.edu.au/docs/default-source/ default-document-library/accreditation-of-ite-programs-in-australiace118891b1e86477b58fff0000 6709da.pdf?sfvrsn=86f9ec3c_2

Australian Institute for Teaching and School Leadership. (2017a). *TPA FAQs.* https://www.aitsl.edu.au/deliver-ite-programs/teaching-performance-assessment/teaching-performance-assessment-grant-program/frequently-asked-questions

Australian Institute for Teaching and School Leadership. (2017b). *Request for tender TPA national expert panel.* https://www.aitsl.edu.au/docs/default-source/initial-teacher-education-resources/tpa/aitsl_tpa-national-expert-panel-rft.pdf?sfvrsn=61b2e93c_2

Bloxham, S., Hughes, C., & Adie, L. (2016). What's the point of moderation? A discussion of the purposes achieved through contemporary moderation practices. *Assessment & Evaluation in Higher Education, 41*(4), 638-653. https://doi.org/10.1080/02602938.2015.1039932

Campbell, C., Cuauhtemoc Ayala, C., Railsback, G., Freking, F. W., McKenna, C., & Lausch, D. (2016). Beginning teachers' perceptions of the California Teaching Performance Assessment (TPA). *Teaching Education Quarterly, 43*(2), 51-71.

Caughlan, S. & Jiang, H. (2014). Observation and teacher quality: Critical analysis of observational instruments in preservice teacher performance assessment. *Journal of Teacher Education, 65*(5), 375–388. https://doi.org/10.1177/0022487114541546

Clayton, C. D. (2018). Policy meets practice in New York state: Understanding early edTPA implementation through preservice candidates' eyes. *Teacher Education Quarterly, 45*(3), 97-125.

Coates, H. (2010). Defining and monitoring academic standards in Australian higher education. *Higher Education Management and Policy, 22*(1), 1–17. https://doi.org/10.1787/17269822

Cochran-Smith, M., Piazza, P., & Power, C. (2013). The politics of accountability: Assessing teacher education in the United States. *The Educational Forum, 77*(1), 6-27. https://dx.doi.org/10.1080/00131725.2013.739015

Darling-Hammond, L. (2012). The right start. *Phi Delta Kappan, 94*(3), 8-13. https://doi.org/10.1177/003172171209400303

Darling-Hammond. L., Newton, S. P., & Wei, R. C. (2013). Developing and assessing beginning teacher effectiveness: The potential of performance assessments. *Educational Assessment, Evaluation and Accountability, 25*(3), 179–204. https://doi.org/10.1007/s11092-013-9163-0

Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education*, *16*(5-6), 523-545. https://doi.org/10.1016/S0742-051X(00)00015-9

Darling-Hammond, L., Wise, A. E., & Klein, S. (1998). *A license to teach: Building a profession for 21st century schools*. Jossey Bass.

Deakin University. (2012). *Deakin authentic teacher assessment (ATA) handbook: Master of Teaching EPR 703 reflecting on practice in professional experience*. Faculty of Arts and Education, School of Education.

Dixon, M., Mayer, D., Gallant, A., & Allard, A. (2011). *Deakin authentic teacher assessment: "authentically assessing beginning teaching: professional standards and teacher performance assessment" (Version 1).* https://hdl.handle.net/10536/DRO/DU:30048315

Dover, A. & Shultz, B. (2016). Troubling the edTPA: Illusions of Objectivity and Rigor. *The Educational Forum, 80*(1), 95-106. https://dx.doi.org/10.1080/00131725.2015.1102368

Downing, S.M., Tekian, A., & Yudkowsky, R. (2006). Research methodology: Procedures for establishing defensible absolute passing scores on performance examinations in health professions education. *Teaching and Learning in Medicine, 18*(1), 50-57, https://doi.org/10.1207/s15328015tlm1801_11

Gent, I., Johnston, B., & Prosser, P. (1999). Thinking on your feet in undergraduate computer science: A constructivist approach to developing and assessing critical thinking. *Teaching in Higher Education, 4*(4), 511-522. https://doi.org/10.1080/1356251990040407

Greenblatt, D. (2018). Neoliberalism and teacher certification. *Policy Futures in Education*, *16*(6), 804–827. https://doi.org/10.1177/1478210318771827

Herbert, I. P., Joyce, J., & Hassall, T. (2014). Assessment in Higher Education: The Potential for a Community of Practice to Improve Inter-marker Reliability. *Accounting Education*, *23*(6), 542-561. https://doi.org/10.1080/09639284.2014.974195

Huxham, M., Campbell, F., & Westwood, J. (2012). Oral versus written assessments: a test of student performance and attitudes. *Assessment and Evaluation in Higher Education, 37(1),* 125-136, https://doi.org/10.1080/02602938.2010.515012

Joughin, G. (1998). Dimensions of oral assessment. *Assessment and Evaluation in Higher Education, 23*, 367-378. https://doi.org/10.1080/0260293980230404

Joosten-ten Brinke, D. Sluijsmans, D. M. A & Jochems, W. M. J. (2010). Assessors' approaches to portfolio assessment in assessment of prior learning procedures. *Assessment and Evaluation in Higher Education, 35*(1), 55-70, https://doi.org/10.1080/02602930802563086

Lave, J., & Wenger, E., (1991). *Situated learning: legitimate peripheral participation.* Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Liu, M., & Liu, K-M. (2008). Setting pass scores for clinical skills assessment. *The Kaohsiung Journal of Medical Science, 24*(12), 656-663, https://doi.org/10.1016/S1607-551X(09)70032-4

Loughran, J. J. (2006). *Developing a Pedagogy of Teacher Education: Understanding Teaching and Learning About Teaching*. Routledge.

Mascadri, J., Spina, N., Author., & Briant, E. (2022). Assessors' perspectives of an oral assessment within a teaching performance assessment. *Assessment and Evaluation in Higher Education*, 1-14. https://doi.org/10.1080/02602938.2022.2122930

Meeus, W., Van Petegem, P. & Engels, N. (2009). Validity and reliability of portfolio assessment in pre-service teacher education. *Assessment and Evaluation in Higher Education, 34*(4), 401-413. https://doi.org/10.1080/02602930802062659

Nelson, C. C., Waechter-Versaw, A., Mitchener, C. P., & Chou, V. (2014). *Teacher Performance Assessment in the Age of Accountability: The case of the edTPA* [Policy brief]. UIC Research on Urban Education Policy Initiative. https://shorturl.at/etPT2

Newton, P., & Meadows. M. (2011). Marking quality within test and examination systems. *Assessment in Education: Principles, Policy and Practice, 18*(3), 213-216. https://doi.org/10.1080/0969594X.2011.586842

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. https://doi.org/10.1080/10632913.2014.944964

Pecheone, R. L., & Chung, R. R. (2006). Evidence in teacher education: The performance assessment for California teachers. *Journal of Teacher Education, 57*(1), 22-36. https://doi.org/10.1177/0022487105284045

Peck, C. A., & McDonald, M. (2013). Creating cultures of evidence in teach education: Context policy, and practice in three high-data-use programs. *The New Educator, 9*(1), 12-28. https://doi.org/10.1080/1547688X.2013.751312

Reagan, E. M., Schram, T., McCurdy, K., Chang, T. & Evans, C. M. (2016). Politics of policy: Assessing the implementation, impact and evolution of the Performance Assessment for California Teachers (PACT) and edTPA. *Education Policy Analysis Archives,* 24(9), 1-24. http://dx.doi.org/10.14507/epaa.v24.2176

Sadler, D. R. (1985). The origins and functions of evaluative criteria. *Educational Theory, 35*(3), 285-297. https://doi.org/10.1111/j.1741-5446.1985.00285.x

Sadler, D. R. (2010). Fidelity as a precondition for integrity in grading academic achievement. *Assessment and Evaluation in Higher Education, 35*(6), 727-743. https://doi.org/10.1080/02602930902977756

Sadler, D. R. (2011). Academic freedom, achievement standards, and professional identity. *Quality in Higher Education*, *17*(1), 85-100. https://doi.org/10.1080/13538322.2011.554639

Sadler, D. R. (2013). Assuring academic achievement standards: From moderation to calibration. *Assessment in Education: Principles, Policy and Practice, 20*(1), 5-19. https://doi.org/10.1080/0969594X.2012.714742

Stacey, M., Talbot, D., Buchanan, J., & Mayer, D. (2020). The development of an Australian teacher performance assessment: lessons from the international literature. *Asia-Pacific Journal of Teacher Education*, *48*(5), 508-519. https://doi.org/10.1080/1359866X.2019.1669137

Stewart, A. R., Salzo, J. N., Merino, N. & Nilsen, K. (2015). Beyond the criteria: Evidence of teacher learning in a performance assessment. *Teacher Education Quarterly, 42*(3), 33-58.

Stillman, J., Anderson, L., Arellano, A., Lindquist Wong, P., Berta-Avila, M., Alfaro, C., & Struthers, K. (2013). Putting PACT in context and context in PACT: Teacher educators collaborating around program-specific and shared learning goals. *Teacher Education Quarterly, 40*(4), 135-157.

Teacher Education Ministerial Advisory Group (2015). *Action now: Classroom ready teachers*. https://www.studentsfirst.gov.au/teacher-education-ministerial-advisory-group

Van der Schaaf, M. F., Stokking, K. M., & Verloop, N. (2005). Cognitive representations in raters' assessment of teacher portfolios. *Studies in Educational Evaluation 31*, 27-55. https://doi.org/10.1016/j.stueduc.2005.02.005

Watty, K., Freeman, M., Howieson, B., Hancock, P., O'Connell, B., De Lange, P., & Abraham, A. (2014). Social moderation, assessment and assuring standards for accounting graduates. *Assessment and Evaluation in Higher Education, 39*(4), 461-478. https://doi.org/10.1080/02602938.2013.848336

Weir, C. (2009). Curriculum and Assessment: Aligning what you value with how you teach. In J. Millwater and D. Beutel (Ed.), *Stepping out into the real world of education* (pp. 109-130). Pearson Australia.

Wenger, E., (1998). *Communities of practice: learning, meaning and identity.* Cambridge University Press. https://doi.org/10.1017/CBO9780511803932

Wyatt-Smith, C. & Klenowski, V. (2013) Explicit, latent and meta-criteria: types of criteria at play in professional judgement practice. *Assessment in Education: Principles, Policy and Practice*, *20*(1), 35-52. https://doi.org/10.1080/0969594X.2012.725030

Zahra, D., Robinson, I., Roberts, M., Coombes, L., Cockerill, J., & Burr, S. (2017). Rigour in moderation processes is more important than the choice of method. *Assessment and Evaluation in Higher Education 42*(7), 1159-1167. https://doi.org/10.1080/02602938.2016.1236183