

# Assessing accomplished teaching with reliability and validity: The ACER Portfolio Project



Dr Lawrence Ingvarson  
Australian Council for Educational Research

*Lawrence Ingvarson is a principal research fellow at the Australian Council for Educational Research. His major research interests centre on the professionalisation of teaching. Lawrence is a Fellow of the Australian College of Educators and a recipient of a Distinguished Service Award from the Australian Science Teachers Association. In 2014, he was awarded the Sir James Darling Medal for outstanding and sustained contribution to Victorian Education by the Australian College of Educators. His publications include Assessing teachers for professional certification: The first decade of the National Board for Professional Teaching Standards, which brings together the rigorous research and development work conducted by the National Board for Professional Teaching Standards since its establishment in the United States in 1987. With colleagues at ACER, Lawrence recently prepared the background research report for the Teacher Education Ministerial Advisory Group and the background report on initial teacher preparation in Australia as part of the OECD TALIS 2018 study. He recently co-directed a study for the IEA on the preparation of mathematics teachers in 17 countries (TEDS-M).*

## Abstract

We know that good teachers are worth their weight in gold. But if good teaching is to be truly valued, the teaching profession must be able to demonstrate that it can evaluate itself in ways that are reliable, valid and fair. This capacity is central to any profession. It is also central to lifting the status of teaching, rewarding accomplished teaching and enabling teaching to compete with other professions for our ablest graduates. Recent OECD reports emphasise the necessity of strengthening the teaching profession, which depends upon widespread use of evidence-based teaching practices.

Building the capacity for evaluation is the purpose of the ACER Portfolio Project: to develop valid and feasible methods by which teachers can demonstrate the ways in which they meet the Australian Professional Standards for Teachers at the Highly Accomplished level. This presentation reviews the work of the Portfolio Project in developing an assessment and evaluation framework for Highly Accomplished teaching, piloting the assessment tasks with teachers, training assessors, setting standards, and identifying benchmarks for highly accomplished teaching.

For more information go to: <https://portfolio.acer.org/>

## The ACER Portfolio Project

In 2012, all Australian education ministers endorsed a set of principles and processes for the certification of teachers who met the Australian Professional Standards for Teachers (the Standards) at the Highly Accomplished and Lead Teacher levels (the Guide).<sup>1</sup>

The Guide sets out two central stages in the assessment process:

- *Assessment stage 1* involves the assessment of evidence submitted by the applicant against the Standards.
- *Assessment stage 2* consists of direct observation of the applicant's practice by an external assessor, and discussion with the applicant and the applicant's supervisor (and perhaps other colleagues nominated by the applicant).

The purposes of the ACER Portfolio Project<sup>2</sup> were relevant to Assessment stage 1, which requires applicants to submit evidence about their practice for assessment by certifying authorities. The project team developed methods designed to assist teachers in that process; in particular, a set of portfolio tasks<sup>3</sup>, which together, would help them demonstrate how they met the Standards at the Highly Accomplished level.

Each portfolio task provided teachers with a set of guidelines for preparing an entry to be placed in their professional portfolio, each based on teaching a unit of work with one class. Each provided a structure within which teachers could show how they provided quality opportunities for students to learn. Four portfolio tasks were prepared for generalist primary teachers and four for secondary science teachers, and trialled with teachers to test their feasibility, clarity, validity and reliability, as well as the impact of preparing an entry on a teacher's professional learning.

We then investigated whether it was possible to train assessors to assess portfolio entries to high levels of consistency and whether it was possible, thereby, to identify benchmark entries and to set standards. Two groups of assessors were trained, one to assess entries from primary teachers, the other to assess entries from secondary science teachers. Assessors also evaluated the entries for their fairness, clarity and validity, and reported on the impact of the training on their professional learning.

## Stages in the Portfolio Project

While the Standards describe what Highly Accomplished teachers know and do, they are not 'standards' in the strict meaning of that term. They needed to be operationalised; that is, valid and reliable methods for providing evidence needed to be developed, as well as methods for judging whether that evidence met the Standards (Ingvarson & Hattie, 2008).

Three questions had to be addressed in making the Standards operational and in developing a framework for the assessment of Highly Accomplished teaching.

1. *What is to be assessed?* In this case, the Standards defined what was to be assessed in the certification system: what Highly Accomplished teachers know and do.
2. *How will it be assessed?* The second stage in the Portfolio Project was to identify how the Standards would be assessed. Valid and reliable methods were needed by means of which teachers can demonstrate how their practice meets the Standards in their school context.
3. *How is the evidence to be judged and the standard set?* The purpose of the third stage was to investigate whether it was possible to train assessors to identify portfolio entries illustrating different levels of performance (i.e. benchmarks) in relation to the Standards, with high levels of consistency.

## Methods for assessing teacher performance: The portfolio tasks

Table 1 (p. 61) provides summaries of the four portfolio tasks for primary teachers. Details of the task guidelines can be found at <https://portfolio.acer.org/guidelines>. Each task has clear sections with question prompts and strict page limits for each section. Tasks are accompanied by criteria indicating how each entry will be assessed.

<sup>1</sup> AITSL 2012, Certification of Highly Accomplished and Lead Teachers in Australia, Education Services Australia, Carlton South, viewed April 2015, [http://www.aitsl.edu.au/docs/default-source/default-document-library/certification\\_of\\_highly\\_accomplished\\_and\\_lead\\_teachers\\_-\\_principles\\_and\\_processes\\_-\\_april\\_2012\\_file.pdf](http://www.aitsl.edu.au/docs/default-source/default-document-library/certification_of_highly_accomplished_and_lead_teachers_-_principles_and_processes_-_april_2012_file.pdf).

<sup>2</sup> Members of the Portfolio Project team included Hilary Hollingsworth, Elizabeth Kleinhenz, Marion Meiers, Anne Semple and Lawrence Ingvarson.

<sup>3</sup> A portfolio is simply a container into which samples of professional performance and accomplishments are entered. A portfolio task is a set of guidelines for preparing an entry for a professional portfolio. A portfolio entry is a completed portfolio task ready to be entered into a portfolio.

**Table 1** Summaries of four portfolio tasks for primary teachers

Task number	Primary teaching portfolio entries
1	English: This portfolio entry invites you to show how you have taught students to develop their capacities in writing for a range of audiences and purposes.
2	Mathematics: This portfolio entry invites you to demonstrate how you have built students' understanding of important mathematics content through class discussion.
3	Inquiry skills: This portfolio entry invites you to demonstrate how you have engaged students in collaborative investigations that have strengthened their inquiry skills and deepened their conceptual understanding, as described in the Australian Curriculum.
4	Engaging colleagues in an improvement initiative: This portfolio task invites you to initiate and manage a project in collaboration with colleagues that improves teaching practice and learning opportunities for a targeted group of students in your school.

The portfolio tasks are designed as authentic performance assessment tasks. They are tasks that match the normal duties of teachers, no matter the context in which they are teaching.

Each portfolio task aims to give teachers a structure, and prompt questions that would enable them to provide clear and convincing evidence of their ability to teach at the Highly Accomplished level. Each task provides a scaffold for documenting a coherent case of their teaching. It is hoped that teachers would regard the tasks as consistent with their normal teaching responsibilities: not as additional work, or irrelevant hurdles to jump over for assessment purposes.

Each portfolio task provides teachers with a clear idea of what they are asked to demonstrate and how it will be assessed. The structure of each task reflects the basic architecture of good teaching; from knowing where students are at in their learning, setting worthwhile goals based on this knowledge, implementing learning activities clearly linked to the goals, assessing student learning in light of the goals, providing timely and useful feedback, and reporting student learning and moving on to set new worthwhile learning goals. Accomplished teaching has this basic underlying structure, and demonstrates strong links between its components and coherence in the overall performance.

## The Portfolio Project assessment framework

The purpose of an assessment framework is to ensure that the portfolio entries, as a group, provide a representative sample of evidence about a teacher's practice in relation to the Standards. The number of tasks is determined by the need to ensure that a sufficient number of independent pieces of evidence are gathered to provide a reliable basis for generalising about a teacher's performance in relation to the Standards. It is impossible for one task to do this.

Table 2 (p. 62) shows the Portfolio Project assessment framework for Highly Accomplished primary teaching. It shows that, together, the four entries provide a sample of evidence relevant to *all of the Standards*. However, if the framework is to provide a valid basis for making decisions about a teachers' performance, it is necessary not only to cover the Standards, but also to ensure that, together, the portfolio entries provide:

- evidence covering the main components of the curriculum that a teacher is responsible for teaching (to ensure this, a primary teacher's entries provide evidence of teaching in several subject areas, not just one)
- evidence covering several core teaching skills reflective of accomplished teaching (to ensure this, each entry provides evidence of a different pedagogical skill).

Table 2 illustrates how these requirements were met. The dark purple shows where the main emphases rests for each entry in terms of the Standards. Entry 1 is designed to provide evidence particularly relevant to Standards 3 (Plan for and implement effective teaching and learning) and 5 (Feedback and assessment skills), based as it is on samples of students' writing over time. It also provides some evidence in relation to Standards 1, 2, but not Standards 4, 6, or 7.

**Table 2** Assessment framework for Highly Accomplished primary teaching

		Portfolio entries			
APST Standards	Curriculum focus	Entry 1 English: Writing	Entry 2 Mathematics	Entry 3 Inquiry skills	Entry 4 Teacher leadership
	Teaching skill	Monitoring and feedback	Glass discussion	Collaborative group work	
	Main sources of evidence	Student work samples	Video-based	Video-based	Documented accomplishments
1. Know students and how they learn					
2. Know the content and how to teach it					
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					
6. Engage in Professional Learning					
7. Engage professionally with colleagues, parents/carers and the community					

Entries 2 and 3 are based on video recordings and are designed to provide evidence in relation to Standard 4 (the Classroom Learning Environment) as well as Standard 3. With Entry 2, the focus shifts to mathematics and the quality of discourse that a teacher can create to promote mathematical understanding, also providing evidence in relation to Standard 3.

With Entry 3, the focus shifts to inquiry skills, as emphasised in the Australian Curriculum, and the teacher's capacity to promote student investigations through collaborative group work to develop those skills. Importantly, teachers are asked not only to provide evidence, but to indicate also how that evidence shows they are meeting the Standards.

Entries 1 to 3 thereby ensure that a teacher's portfolio includes evidence of planning and teaching units of work in several subjects, not just one. Similarly, Entries 1 to 3 also ensure that a teacher's portfolio includes evidence of the ability to implement several core teaching skills, not just one. These requirements increase the number of independent sources of

evidence and thereby promote greater reliability in making judgements about a teacher's performance.

Entry 4 directly addresses Standards 6 and 7 in the APST, and the emphasis through all the Standards that Highly Accomplished teachers are active contributors to their schools as professional learning communities. Entry 4 requires a teacher to initiate and document a small project that engaged colleagues in improving the quality of learning opportunities for a designated group of students. It also provides evidence in relation to Standards 1 to 5.

### The field test

The portfolio tasks were subjected to a careful review of their clarity, validity, and fairness by external panels of teachers in each field. After revision, and a second review by the external panels, a general invitation was sent out mid-2015 to teachers in each field to undertake one portfolio task. Twenty-one teachers completed a portfolio entry and submitted it electronically by the end of 2015. Trial teachers were also provided with an

evaluation guide<sup>4</sup> for each portfolio entry, indicating the evidence that assessors would be looking for and how they would make their overall judgment using a four point scale.

Trial teachers rated the clarity, validity and fairness of the portfolio tasks highly. They also reported that preparing their entry was a valuable professional learning experience that improved their teaching.

## Setting standards for portfolio entries

### Training assessors

The second stage of the Portfolio Project investigated whether it was possible to train assessors to score the portfolio entries consistently and identify benchmark entries; that is, entries rated consistently by assessors as illustrating different levels of performance on score scale.

Invitations were distributed widely to teachers interested in being trained as assessors, generating considerable interest. Sixteen primary teachers from different states and school systems participated in the training over three days to assess the primary portfolio entries and 12 science teachers from different states and school systems participated in similar training for the secondary science portfolio entries.

Training of assessors took place at ACER late in 2015. The first aim was to ensure that assessors had a clear understanding of the four portfolio tasks – what each task was measuring and what evidence to look for, as described in the relevant evaluation guide. The second was to minimise bias and to ensure that assessors carefully documented the evidence they saw and made their judgments independently using an Assessment Record Form (ARF). The ARFs ensured that the assessment process was transparent and reproducible, providing records of how assessors arrived at their judgments, and thereby also legally defensible.

Assessors were trained to use a four-level scale for judging portfolio entries, where a score of 3 meant assessors agreed the entry provided clear evidence of meeting the key criterion for that entry. A score of 2 meant there was evidence, but it was insufficient and a score of 1 meant there was little or no evidence. A score of 4 meant the evidence more than met the certification level and was uniformly convincing, coherent and consistent.

As they read each entry, assessors were trained to follow a 'scoring pathway' consisting of two stages: an 'analytic stage' that required them to first record the evidence they saw relevant to the criteria for each

section in their ARF and where they saw it, before making judgments for each section, followed by a 'holistic stage' in which they 'stepped back' and reviewed the entry as a whole and judged the extent to which there was clear, consistent and convincing evidence across the entry that the key criterion for that entry had been met. The key criteria summarised what assessors were to look for in the entry.

The key criterion for Entry 1, for example, asks the assessor to judge whether:

The entry provides clear evidence that the teacher has engaged students in writing for a range of purposes and audiences, catering for the diverse learning needs of students in planning classroom activities, and enabled all students to make progress in their knowledge and understanding of writing.

In making their overall judgement, assessors were trained to focus on the coherence and consistency across the stages documented in an entry. For example, they were asked to look for clear links between:

- evidence about the students and the selected learning goals
- the learning goals and the learning activities, materials and resources
- the learning goals and the methods of monitoring and assessing student learning
- the teacher's analysis of and reflection on their teaching and the evidence of their students' learning.

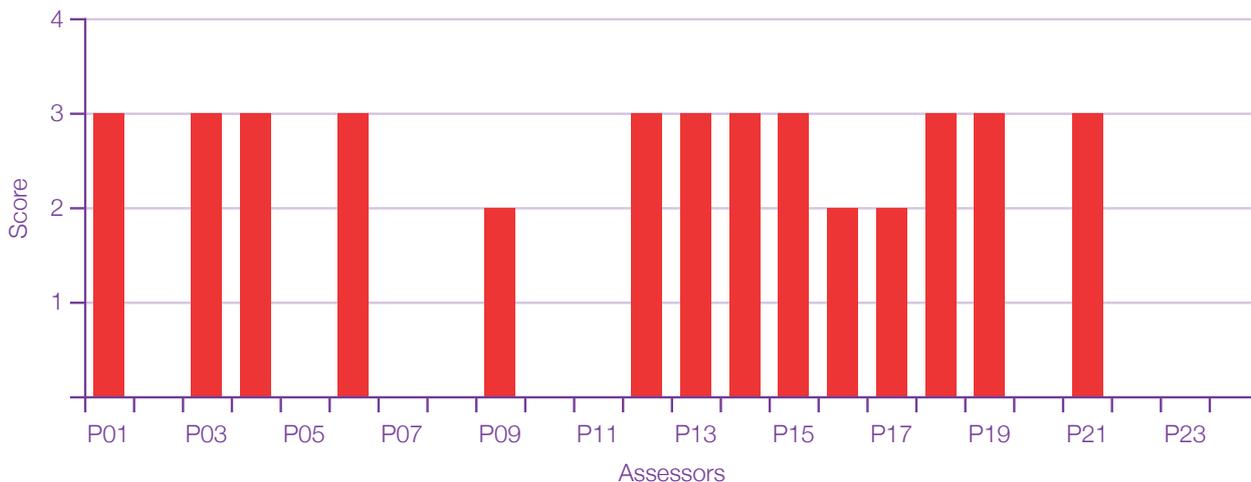
As a final step before making their final judgement, and to minimise bias, assessors were also required to consider questions such as:

Does the entry still meet the certification level, even if the approach used by the teacher is not the one you would have chosen to use yourself?

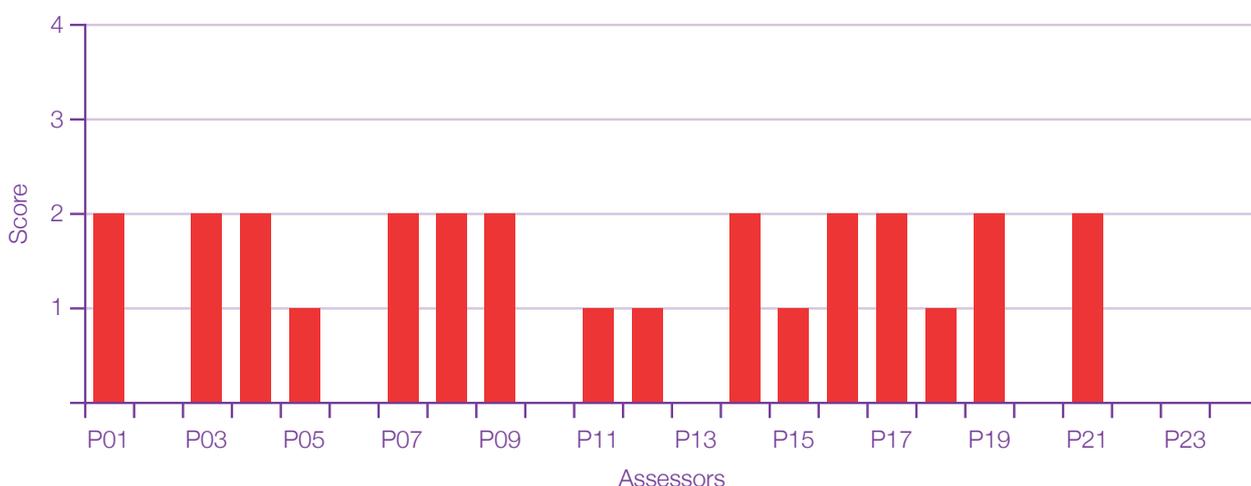
### Identifying benchmark portfolio entries

Following training, assessors began judging portfolio entries independently and submitting their assessment record forms. Figure 1 (p. 64) shows, for example, that 11 out of 14 assessors gave Writing Entry P1004, an entry on writing from a primary teacher, a score of 3, which meant that most assessors thought the teacher's performance was at the certification level.

<sup>4</sup> <https://portfolio.acer.org/guidelines>



**Figure 1:** Writing Entry P1004



**Figure 2:** Writing Entry P1033

Figure 2 shows that 11 out of 16 assessors gave Writing Entry P1033, another teacher’s writing entry, a score of 2.

A ‘benchmark’ is an example of what the Standards looks like in practice. An important aim of the Portfolio Project was to investigate whether it was possible to identify benchmark entries. A portfolio entry could be labelled as a ‘benchmark portfolio entry’ if there was a high level of agreement among assessors about the level of performance it represented.

The level of agreement about scores for Writing Entry P1004 means that most assessors judged it to be a clear example of a performance that met the certification level. It almost warranted being labelled a benchmark 3; an example of entry with a score of 3. All assessors agreed that Writing Entry P1033 did not quite meet the certification level of performance. Eleven out of 16 gave it a score of 2, meaning that it provided limited evidence of a performance at the certification level. However, the level of agreement is not high enough to warrant using the entry as a benchmark.

Assessors went on to assess entries that primary teachers had submitted in the other three categories, mathematics discourse, inquiry skills, and engagement with colleagues, with similarly high levels of agreement.

Indications were, therefore, that that portfolio entries could be assessed reliably. However, trials with much greater numbers of portfolio entries and assessors would be needed to substantiate claims in this direction. The training also indicated that benchmark entries representing performance at different score levels could be identified, though higher levels of agreement would be desirable (which may point also to the need to refine or clarify some aspects of the portfolio task guidelines themselves).

Further trials providing more entries will be needed to build a larger ‘stock’ of benchmark entries representing each score level. This will be essential for later and more thorough training of other assessors. Trainers use benchmark entries to sharpen assessors’ abilities to discriminate between portfolio entries that represent

different levels of performance. They also use them to show assessors that, although different in approach, portfolio entries may nevertheless represent the same level of performance in relation to the Standards.

## Assessor's views of the portfolio tasks and the assessment process

Following the training, assessors were asked to complete a survey similar to that completed by the field test teachers. Assessors also rated the validity and fairness of the tasks highly, but indicated the need for more work on the clarity of the guidelines. Assessors felt increasingly confident about the consistency of their assessments and found the assessment process a valuable professional learning experience.

## Conclusion

The Portfolio Project aimed to provide teachers applying for certification with practical and valid means by which they could show how they met the Standards. The structured nature of the portfolio tasks, with clear guideline prompts and word limits, meant that teachers found them feasible and that assessors were able to assess portfolio entries consistently.

Each portfolio task provided teachers with a clear idea of what they were being asked to demonstrate and how it would be assessed. This is a fundamental requirement for assessment of performance in any field, especially in the professions. Trial teachers clearly appreciated this structure.

In contrast with the AITSL Guide that asks a teacher to 'evidence' the Standards by collecting or gathering existing evidence, the portfolio tasks invite a teacher to create or produce coherent examples of their teaching performance, by initiating and documenting, analysing and reflecting on units of work that they have taught to particular classes. This meant that that the process of preparing an entry was also a vehicle for promoting professional learning, with spin-off benefits for their schools, especially if groups of teachers were preparing entries for certification together,

There is clearly more work to be done before the assessment framework and portfolio tasks can claim to provide a valid approach to differentiating teachers who have attained the Standards of performance at the Highly Accomplished level from those who as yet have not. The same applies to methods be used currently by

certifying authorities in each state and territory, and in each sector.

However, an important outcome of the Portfolio Project so far is that it does provide teachers with tasks that they can undertake with some confidence in their validity and, when completed, submit as part of their evidence in applying for certification. Our hope is that, if our assessments prove to be sound, responsible authorities might also encourage teachers to use them as evidence for certification purposes. We also believe the Portfolio Project points the way to a more economically affordable, administratively feasible and legally defensible certification system, major long-term considerations if a certification system is to 'go to scale'.

Australia's current ambitions to establish a respected and nationally consistent certification system for Highly Accomplished and Lead Teachers will live or die depending on the trust the profession, the public and employing authorities are willing to place in the validity, reliability and fairness of the assessment methods that underpin certification decisions (Ingvarson, 2013). Without such a guarantee, the original objective that a nationally consistent certification system would lift the status of teaching, provide stronger incentives for professional learning, reward accomplished teaching more appropriately, and thereby enable teaching to compete more effectively with other professions for our ablest graduates, will not be realised (Ingvarson, 2014).

## References

- Ingvarson, L. (2014). Standards-based professional learning and certification: By the profession, for the profession. In L. Martin (Ed.), *The handbook of professional development in education, successful models and practices, pre K-12* (pp. 384-411). New York, NY: Guilford Publishing.
- Ingvarson, L. C. (2013). Reforming career paths for Australian teachers. In M. Akiba (Ed.), *Teacher reforms around the world: Implementations and outcomes* (pp. 237-273). Bingley, UK: Emerald Publishing.
- Ingvarson, L. C., & Hattie, J. (Eds.). (2008). *Assessing teachers for professional certification: The first decade of the national board for professional teaching standards*. Volume 11, *Advances in Program Evaluation*. Amsterdam, Netherlands: Elsevier Press.