

Australian Qualifications *Framework*

AQF COUNCIL

Overview of the Empirical Validation
of the Strengthened Australian
Qualifications Framework

September 2010

*Australian Qualifications Framework Council
A Committee of the Ministerial Council for Tertiary
Education and Employment*

www.aqf.edu.au

Background to the research

The Australian Qualifications Framework Council is currently undertaking a project to strengthen the Australian Qualifications Framework (AQF).¹

In October 2009, after extensive national consultations, a paper entitled *Strengthening the AQF: An Architecture for Australia's Qualifications* was released by the AQF Council for comment. The paper described a new AQF architecture: a proposed levels-based structure consisting of ten levels, each with attributes and criteria, revised qualification type descriptors for each of the existing qualification types (and a proposed additional type) and an indication of the notional duration of student learning for each qualification. The levels criteria and qualification type descriptors are defined by learning outcomes and are expressed in terms of knowledge, skills and the application of knowledge and skills.

After further consultations and workshops with stakeholders and other interested parties, the draft levels criteria and qualification type descriptors were further revised. The AQF Council has since agreed to the draft levels criteria and qualification type descriptors and the AQF terminology – in effect, to a new version of the Australian Qualifications Framework (Version 6).

In early 2010 objective research was commissioned to validate the 'strengthened' AQF. The aim of the objective testing component was to confirm the validity of the consolidated draft levels, the criteria and qualification type descriptors, as well as the relationships between them. An online questionnaire was developed and completed by users of the AQF. The results were analysed using an item response modelling technique, designed by a consortium of researchers led by Victoria University. The objective testing research also presented a timely opportunity for identifying potential improvements to the revised framework.

It is anticipated that the results from the testing phase and the implementation of suggested improvements will contribute to greater certainty and trust in the strengthened framework, at the same time encouraging meaningful engagement in its ongoing development by a significant number of stakeholders nationally, while ensuring international credibility in Australia's qualifications structure.

This paper provides an overview of the findings of the testing and the response of the AQF Council to the results, including measures the Council has taken to fine-tune the strengthened levels criteria and qualification type descriptors.

Overview of the research project

The aim of this study was to undertake an empirical analysis of the revised design of the strengthened Australian Qualifications Framework. In particular, four elements of the revised framework were to be examined:

- The levels structure, with ten levels expressed as learning outcomes (referred to as 'levels criteria')

¹ For more information about the Strengthening the AQF project, see <http://www.aqf.edu.au>.

- Revised descriptors for each of the existing 14 qualification types (and two kinds²) expressed as learning outcomes (referred to as 'qualification type descriptors')
- The relationship between the qualification types and the levels structure
- An estimate of the notional duration of student learning for each qualification type.

In the development of the strengthened framework a number of assumptions underpinned the development of the levels criteria and qualification type descriptors. These were:

- The overall framework was designed to present a pathway through various stages/levels of learning; the levels criteria should therefore be strictly hierarchical and cumulative. This same principle should apply to the qualification type descriptors, although to a lesser extent.
- Each set of levels criteria should clearly demonstrate increasing complexity from one level to the next and, along with qualification type descriptors, clearly and explicitly signal this increasing complexity.
- Neither the levels criteria nor qualification type descriptors should identify the learning or workplace context, or educational sector where the qualification type is typically delivered.
- The levels structure and the qualification types should be underpinned by three dimensions (knowledge, skills and application):
 - Within each dimension, the criteria and descriptors should represent a wide range of levels of complexity (that is, across all qualifications from Certificate I to Doctoral Degree).
 - Each dimension should be internally coherent, in that the set of statements³ should represent increasing complexity.
 - Each of the dimensions should contribute to some unique aspect of the measurement of complexity of learning outcomes.
- Each qualification type should be described by a set of qualification type descriptors that will capture the desired complexity of the learning outcomes of each.
- Individual descriptors within one qualification type could be used in another. It is the unique combination of descriptors in a set that makes the description of the qualification type unique, not the individual descriptors themselves.
- More than one qualification type could be positioned at the same level on the framework (see Table 1).

The major aims of the empirical validation were to:

- Estimate the complexity of the criteria for each of the levels, and for each set, compare the estimates with the proposed 10-level structure.
- Estimate the complexity of each qualification type descriptor for each of the 14 qualification types.
- Identify any potentially redundant and non-discriminating levels criteria and/or qualification type descriptors.
- Determine where each qualification type is typically positioned within the proposed 10-level structure.
- Investigate the adequacy of the suggested duration for each qualification type.

² The Master's and Doctoral Degree qualifications types had two kinds: other and research.

³ The term 'statement' is used within this document to refer to the levels criteria and/or qualification type descriptors.

The AQF Council had already proposed a relationship between the levels structure and qualification types, which is shown in Table 1. An aim of this study was therefore to validate empirically the proposed relationship between the qualification types and the 10-level structure depicted in Table 1.

Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Level 9	Level 10
Certificate I	Certificate II	Certificate III Senior Secondary Certificate of Education	Certificate IV	Diploma	Advanced Diploma Associate Degree	Bachelor Degree	Bachelor Honours Degree Graduate Certificate Graduate Diploma	Masters Degree (Other and Research kinds)	Doctoral Degree (Other and Research Kinds)

❖ **Table 1: The proposed relationship between the levels structure and the qualifications type**

For example, the AQF Council proposed that the Bachelor Honours Degree and to the Graduate Certificate and Graduate Diploma be aligned at Level 8. Similarly, the Senior Secondary Certificate and the Certificate III were both expected to be aligned with Level 3.

How the research was conducted

The empirical analysis of the strengthened AQF was undertaken through four stages:

- survey design
- pilot study
- data collection
- data analysis and reporting.

Survey design

There were 186 statements that required validation in the strengthened AQF, although some of these statements had been used to describe more than one qualification type. Once the duplicated statements had been identified, 50 levels criteria and 109 qualification type descriptors remained, which were to be presented as unique statements in the survey.

The online survey was designed so that minimal work was required from respondents, but that sufficient data on all 159 statements could be collected. It was also considered that the time taken to complete the survey should not be any more than 20 minutes. To this end five alternative forms were designed, with some common statements across the forms. This meant that each form was linked to another, which enabled all statements to be positioned on a single measurement scale. Each form also contained statements that were within three or four levels of the qualification type being rated. Five alternative survey forms meant that each respondent was required to complete only around 50 items.

The online questionnaire was designed to:

- ◆ Present items randomly to ensure that the order of the items *could not* be determined by the respondent using cues unrelated to the actual content of the items, and to avoid an item-positioning effect.⁴
- ◆ Randomly present a form for those qualification types which were common across some forms. This would help minimise the likelihood of some items being rated against just a few specific qualification types and also ensure that the forms were adequately linked for the purposes of a single measurement scale.
- ◆ Restrict respondents from moving onto the next item until the previous item had been completed, to minimise missing data on some items that may be more difficult to rate.
- ◆ Enable a respondent to return to an incomplete form at a later time.
- ◆ Provide easy reference to a glossary of terms to assist with interpretation of the language used to describe the levels criteria and qualification type descriptors.

Each respondent was required to supply background information as well as select a qualification type that would become the focus of his/her responses to the questionnaire. Each respondent was then required to rate whether a particular statement was 'too low', 'at this level' or 'too high', in terms of the learning outcomes expected of the selected qualification type.

Pilot study

A small pilot study was undertaken to examine the usability and functionality of the online survey prior to the launch on the website. The findings informed improvements to the questionnaire prior to data collection.

Data collection

The target population for the survey were individuals who used the AQF to develop, accredit, deliver or assess Australian senior secondary, vocational education and training (VET) or higher education qualifications. To ensure that there were enough data across all qualification types, a sample size of at least 700 respondents was sought (that is, 50 respondents per qualification type).

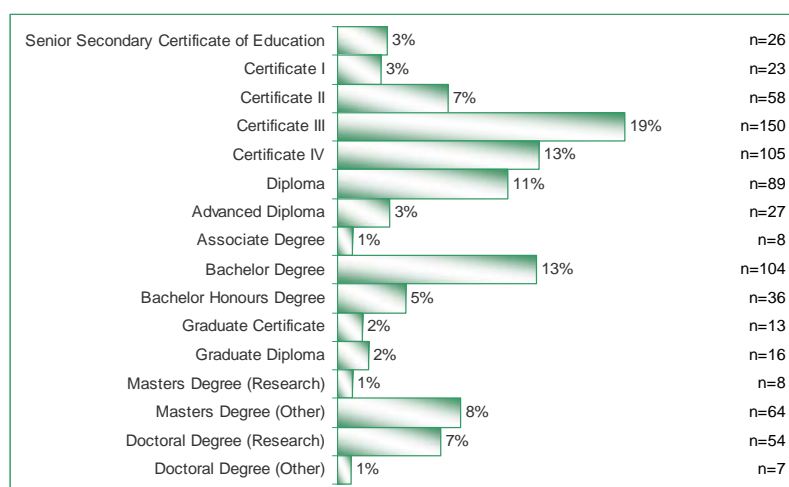
Data were collected over a six-week period via an online survey, accessed through the AQF Council's website. During and prior to the data-collection period key stakeholder groups were informed about the survey and asked to urge their members and employees to participate.

Respondents

Individuals from a wide range of fields of study participated in the national online survey. Of the 788 respondents, 39% represented the higher education sector, 52%, the vocational education and training sector, with 4% from the senior secondary education sector. All states and territories were represented in the sample, with the majority of respondents from New South Wales (29%) and Victoria (27%); South Australia and Western Australia had equal representation at 11%, while very few respondents were located in the Northern Territory (1%), Tasmania (1%) and the Australian Capital Territory (4%).

⁴ An item-positioning effect refers to a tendency for those items located toward the end of the questionnaire to have a higher proportion of missing data or estimated responses because of respondent fatigue and/or boredom

Although a minimum of 50 respondents per qualification type was the desired sample size, some qualification types did not meet this quota, while others exceeded it. Figure 1 displays the percentage (%) and number of respondents (n) for each of the 14 qualification types (and two kinds).



❖ **Figure 1: Qualification type selected**

As the figure demonstrates, the quota was reached for seven of the 14 qualification types, namely Certificates II to IV, Diploma, Bachelor's Degree, Master's Degree (Other and Research combined) and Doctoral Degree (Other and Research combined).

Each survey form comprised statements drawn from a range of qualification types and levels. Table 2 displays the minimum number of respondents rating the set of descriptors for each qualification type.

Qualification type descriptor set	Minimum number of respondents
Senior Secondary Certificate of Education	409
Certificate I	223
Certificate II	223
Certificate III	409
Certificate IV	409
Diploma	281
Advanced Diploma	181
Associate Degree	181
Bachelor Degree	181
Bachelor Honours Degree	284
Graduate Certificate/Graduate Diploma	284
Masters Degree (Research and Other)	198
Doctoral Degree (Research and Other)	198

❖ **Table 2: Minimum number of respondents rating the set of descriptors for each qualification type**

At a minimum, any one statement was rated by at least 181 respondents and, in some cases, as many as 409 respondents.

Data analysis

The data were analysed using Item Response Theory (IRT). This technique enabled the complexity of the statements and qualification types to be estimated according to the same scale of measurement, meaning that the complexity estimate of one statement and/or qualification type could be directly compared with the estimate of another statement and/or qualification type.

Because the ratings of a particular qualification type may vary between respondents (as people who mark essays may differ in making subjective judgements of a piece of student work), this study recruited multiple respondents to rate each qualification type against multiple statements of criteria and descriptors. The ratings from all respondents for each qualification type were then aggregated to provide a more reliable measure for each type. Similarly, each statement was rated not only by multiple respondents, but also rated against multiple qualification types, so that the statements could be calibrated in relation to one another and also to the qualification types.

Item Response Theory also provided a helpful tool for 'equating' the five survey forms, in which the respondents were presented with different sets of questions. Since there were common statements that linked the different questionnaire forms, the process of equating became possible.

The outcome of the analyses was that the complexity of each statement and each qualification type was estimated on a single measurement scale. The scale ranged from a minimum score of 100 (indicating the statement with the lowest complexity), to a maximum of 200 (indicating the statement with the highest complexity estimate on the scale).

In addition to the complexity estimates, the performance of each statement was also reviewed according to its 'discrimination' level. Low discrimination indicated that the statement was unable to separate qualification types of varying complexity, while statements with high discrimination signified the opposite – the capacity to identify qualification types of varying complexity. Consequently, a statement with higher discrimination is considered superior to one of lower discrimination.

What the research found

The levels criteria

The levels structure, as anticipated, was hierarchical and cumulative and covered a wide range of complexity. However, the Item Response Theory analyses revealed that two levels on the revised framework (namely, Levels 2 and 7) would benefit from minor modifications to some of the individual criteria to increase the overall complexity of each set. Increasing the complexity of the sets of levels criteria for Levels 2 and 7 and possibly reviewing a knowledge criterion within Level 9 would improve the gradual progression of complexity across the ten band levels.

Qualification types

The Item Response Theory analyses revealed that the sets of descriptors for each qualification type had levels of complexity which generally accorded with the expectations of the AQF Council (see Table 1). The exception to this was the Associate Degree, which had a set of descriptors that were, on average,

lower in complexity than those for the Advanced Diploma and Diploma. Furthermore, the set of descriptors for the Advanced Diploma were, on average, similar in complexity to the Bachelor Degree and Bachelor Honours Degree.

The analyses revealed that qualification type complexity for each of the 14 qualification types was generally as expected. Furthermore, it was closely aligned to the average descriptor complexity and was within the descriptor complexity range for all qualification types. These findings suggest that the qualification type descriptors adequately described the complexity of the qualification types currently being used.

Relationship between the levels structure and qualification types

Examining the proposed relationship between the qualification types and the 10-level structure, as depicted in Table 1, was a major objective of this research. The Item Response Theory analyses revealed that the average complexity estimates (that is, qualification type complexity) for the Associate Degree, Bachelor Honours Degree and Graduate Certificate/Graduate Diploma were lower than the range of complexity expected for their proposed level (that is, the level complexity range). Furthermore, if the recommended changes to the criteria within Levels 2, 7 and 9 were to be implemented, then the average complexity of the Certificate II, Bachelor Degree and Master's Degree would also possibly be below their expected level complexity range.

The descriptor complexity range closely matched the proposed level complexity range for the Senior Secondary Certificate of Education and Certificates I to IV. That is, there was a good match between the complexity of the descriptors for these five qualification types and their expected level on the revised framework. At the higher levels, however, it was found the qualification types tended to have one or more descriptors that were below their expected level complexity range. There was also a tendency at the higher end for the levels criteria to be higher in complexity than the qualification type descriptors.

Notional duration of student learning

The study found that 76% of the respondents supported the suggested duration of student learning specified for the qualification type selected. However, the findings for Certificate III were inconclusive, with approximately half of the respondents in agreement (that is, agreeing or strongly agreeing) over the time frame specified, while the other half disagreed (disagree or strongly disagree). Although qualitative feedback was gathered to investigate respondent disagreement further, this issue remains unresolved.

Individual statements for review

Complexity estimates and discrimination values were used to identify individual statements for further consideration by the AQF Council. As explained earlier in the data analysis section, the complexity estimate of each statement can be used to identify criteria/descriptors considerably higher or lower than the complexity of the level/qualification type they purport to describe. The discrimination of a statement provides additional information indicating 'how well' a statement was able to separate qualification types of varying complexity. Any statement that was either considerably higher or lower in complexity than

expected, as well as having relatively low discrimination values, was considered to be a high priority for review.

The analyses revealed that four statements fell within this category, all of which were within the application dimension. Two of these four statements were located in the set of criteria for Levels 7 and 8, while the remaining two were from the set of descriptors within the Associate Degree and Doctoral Degree.

The empirical testing also identified statements that demonstrated average/high discrimination, but had a considerably higher or lower than expected complexity estimate, indicating the need for review. Although these statements were found to be adequate in separating qualification types of varying complexity, the discrepancy between expected and observed complexity signals that they may be better aligned to different levels/qualification types than originally intended. These statements were identified as medium priority for review.

Fifty-two of the 159 statements investigated in this study were classified as medium priority (around one-third of the statements in the framework). The majority of these were associated with levels/qualification types at the higher levels of the revised AQF.

In some cases a statement displayed a low discrimination value, but it also had the complexity estimate generally expected. It is recommended that statements such as these, although of a low priority, be reviewed, since they did not separate qualification types (in terms of varying levels of complexity) as adequately as other statements with higher discrimination. Statements in this category could be considered redundant or could be reviewed to improve their discrimination.

Finally, four descriptors were classified as low priority for review. Two of these descriptors were from Certificate I, with one each from the Associate Degree and Bachelor Degree.

AQF Council response

The findings of the objective research indicated that only minor modifications were needed to achieve clearer levels criteria and qualifications type descriptors and improved alignments, and that if the suggested refinements were made there would no need to extensively re-test the framework. The technical report produced by the research consortium clearly outlined where these modifications were required and why. As a consequence, the criteria and the descriptors were refined to improve the overall workability and relevance of the strengthened AQF architecture.

The refinement work was undertaken with input from qualifications and accreditation experts and representatives from the relevant sectors. Specifically, the AQF Secretariat and a qualifications expert reviewed and modified the levels criteria identified in the report as problematic. This was a relatively simple task as the results of the testing provided a clear guide to items that were relatively too high or low or were redundant or non-discriminating. Because the research indicated little change to the qualification type descriptors up to and including the Advanced Diploma, the Secretariat worked with the same qualifications expert to undertake the modifications required to refine these descriptors.

The changes identified in the report for the higher education qualification types were more complex and required academic expertise. A technical working party of representatives of higher education providers

and accrediting authorities was convened to consider the modifications identified in the report. The group was also asked to validate the modifications made to the levels criteria for Levels 6 to 10. A change from three to five years to three to four years was recommended to the notional duration of student learning for the Doctoral Degree to better reflect expectations in the sector.

The research showed that the descriptor used for the Associate Degree was not a match with Level 6. A separate working group was convened to refine the Associate Degree qualification type descriptor to better align with Level 6. The group included people with current experience with the Associate Degree.

The levels criteria, qualification type descriptors and notional duration of student learning refined in response to the testing were released as part of the *Strengthening the AQF: A Framework for Australia's Qualifications* consultation paper in July 2010. Following the consultation the AQF Council agreed to a strengthened AQF comprising the structural elements of the refined levels criteria and qualification type descriptors with a volume of learning (replacing the term 'notional duration of student learning'), supported by a set of policies.

Further details

The empirical validation of the strengthened Australian Qualifications Framework was carried out by a consortium led by Victoria University. The principal researchers on the project were Dr Shelley Gillis from Victoria University, Dr Margaret Wu, Mark Dulhunty and Leanne Calvitto from Education Measurement Solutions, and Andrea Bateman from Bateman & Giles Pty Ltd.

Further information on the Strengthening the AQF Project, including the consultation papers, can be found on the Australian Qualifications Framework website <http://www.aqf.edu.au>.

Definitions used in research

Qualification terms

AQF level	An AQF level is an indication of the relative complexity and/or depth of achievement and the autonomy required to demonstrate that achievement. There are ten levels of complexity. AQF Level 1 has the lowest complexity. AQF Level 10 has the highest complexity.
Levels criteria (also referred to as 'criteria')	Levels criteria describe the relative complexity and/or depth of achievement and the autonomy required to demonstrate that achievement for each AQF Level 1 to 10. Each AQF level is described by a set of levels criteria.
Notional duration of student learning	Notional duration of student learning is the estimated time it takes a student, on average, to complete all learning and assessment activities required for the achievement of a qualification.
Qualification type	A qualification type refers to the term used for a qualification and which is specific to the AQF, such as Certificate III, Bachelor Degree, or Vocational Graduate Diploma.
Qualification type descriptors (also referred to as 'descriptors')	A qualification type descriptor is the statement that describes the learning outcomes of each AQF qualification type. Each qualification type is described by a set of qualification type descriptors.
Statements	A statement is the collective term used in this report to describe the levels criteria and/or qualification type descriptors.

Technical terms

Item Response Theory	Item Response Theory (IRT) is the methodology used in this research. In this context it uses a probability function to describe the likelihood of a respondent giving a particular rating to a qualification type against a particular statement. Through the use of the probability function, conclusions can be drawn about the relative standing of qualification types and statements, as both are calibrated on the same measurement scale.
Complexity estimates	Complexity estimates have been determined for each statement and qualification type, based upon the findings from the Item Response Theory analyses. A linear transformation was applied to the Item Response Theory scores, whereby the statement estimates ranged from a low of 100 (lowest complexity) to a maximum of 200 (highest complexity).
Discrimination	Discrimination refers to the extent to which a statement discriminate between qualification types of varying complexity.
Qualification type complexity	Qualification type complexity refers to the average complexity estimate of a qualification type.
Average descriptor complexity	The average descriptor complexity refers to the average complexity estimate of the set of descriptors used to describe a particular qualification type.
Level complexity range	Level complexity range refers to the difference between the maximum and minimum complexity estimate within the set of levels criteria.
Descriptor complexity range	Descriptor complexity range refers to the difference between the maximum and minimum complexity estimate within the set of qualification type descriptors.



Level 4, 11 Waymouth Street
Adelaide SA 5001

GPO Box 320
Adelaide SA 5001
Australia

Tel: +61 8 8226 2775

Email: aqfc@saugov.sa.gov.au
www.aqf.edu.au