TVET
AUSTRALIA

NATIONAL QUALITY COUNCIL

A CODE OF PROFESSIONAL PRACTICE
FOR VALIDATION AND MODERATION
2009

Table of Contents

1. About this Code

1.1 Use of the Code

The Code of Professional Practice contains a set of high level principles as general guidance on how to conduct assessment validation and moderation within a vocational education and training (VET) setting.

The *Code* is intended to complement Elements 1.1 and 1.5 of the Australian Quality Training Framework (AQTF) Essential Standards for Registration and be consistent with the TAA04 Training and Assessment Training Package. The *Code* is not intended to be mandatory, exhaustive or definitive, and may not be applicable to every situation. Instead, the *Code* is intended to be aspirational and educative in nature.

This *Code* firstly explains the rationale behind its development. Secondly it explains the technical terms validation and moderation and provides a clear distinction between the two terms. It also contains a set of principles that should underpin assessment validation and moderation in the VET sector. It is recommended that professional judgement be used to apply the principles to the various situations an organisation may face when conducting assessment validation and/or moderation.

1.2 Rationale for a Code of Professional Practice

In recent times, some key stakeholders have raised concerns with the quality and consistency of assessments being undertaken by Registered Training Organisations (RTOs). That is, there are some concerns that assessment standards in the VET sector are often not comparable. Ensuring the comparability of standards[1] has become particularly pertinent in the VET sector, as assessments can now be made across a range of contexts (e.g. vocational education, educational and industrial contexts) by a diverse range of assessors using highly contextualised performance based tasks that require professional judgement by assessors.

There are a number of different quality management processes that could be used to help achieve national comparability of standards whilst still maintaining sufficient

---

[1] Comparability of standards are said to be achieved when the performance levels expected (e.g. competent/not yet competent decisions) for a unit (or cluster of units) of competency are similar between assessors assessing the same unit(s) in a given RTO and between assessors assessing the same unit(s) across RTOs.

flexibility at the RTO level to conduct assessments. Typically, there are three major components to quality management of educational assessments: quality assurance; quality control; and quality review[2].

A quality assurance approach attempts to **assure** quality of assessment in VET through focusing on the procedures used in the assessment process. Such an approach is based upon the assumption that the introduction of products and processes such as policies, competency standards, professional development support materials and training within the sector can improve the quality of assessments. Hence, it is referred to as an 'input approach' to quality management.

The second approach to quality management, referred to as 'quality control' focuses on monitoring, and where necessary making adjustments to judgements made by assessors prior to the finalisation of assessment results/outcomes. This approach therefore involves the direct management of assessment judgements to **ensure** consistency in the interpretation and application of the competency standards. As it occurs prior to the finalisation of the result, in which alterations can be made to assessor judgements, it is referred to as an 'active process'.

The third approach to quality management is referred to as 'quality review' as it involves the review of the assessment procedures and outcomes for the sole purpose of **improving** assessment processes and procedures for future use. It is referred to as a retrospective approach as the outcomes of the review are aimed at making recommendations for future improvements. The outcomes of the review have no direct impact on any current or past assessments.

A number of mechanisms or potential mechanisms for enhancing quality assurance, quality control and quality review within the Australian VET sector can be identified. These have been displayed in Table 1.

---

[2] See Maxwell, G.S. (2001). *Moderation of assessments in vocational education and training*, QLD Dept of Employment and Training, Brisbane.

Table 1: Examples of processes for enhancing quality assurance, quality control and quality review of assessments in vocational education and training contexts.

Assessment Quality Management		
Quality Assurance (Input approach)	**Quality Control (Outcome approach)**	**Quality Review (Retrospective approach)**
Examples include: • Industry competency standards as the benchmarks for assessment • National assessment principles • Minimum qualifications for assessors (i.e. TAA40404) • Development of a Professional Code of Practice • Standardisation of reporting formats • Assessment Guidelines and Policy Documents • Benchmark examples of varying levels of performances • Assessment tool banks • Common assessment tasks • Exemplar assessment tools • Panelling, Piloting and/or Trialling of assessment tools. • Professional development programs/workshops for assessors	Examples include: • ***Moderation*** in which adjustments to assessor judgements are made to overcome differences in the difficulty of the assessment tool and/or severity of the judgement	Examples Include: • Monitoring and auditing of registered training organisations • Review and ***validation*** of assessment tools, processes and outcomes to identify future improvements • Follow-up surveys with key stakeholders (e.g. student destination surveys, employer feedback on how well the assessment outcomes predicted workplace performance)

It can be seen in Table 1 that there are a number of different quality management processes that could be used to help achieve national comparability of standards, whilst still maintaining flexibility at the RTO level to design and conduct assessments. One such approach is the development of a *Code of Professional Practice* for guiding assessors when conducting assessment validation and moderation. In terms of a quality management approach, this would be classified as a 'quality assurance approach' as it attempts to provide guidance to assessors on how to conduct assessment validation and moderation through the establishment of a set of principles. There is currently no such *Code of Professional Practice* in the Australian VET sector despite RTOs being encouraged to implement validation as a means of supporting continuous improvement (i.e. refer to the AQTF User's Guide to the Essential Standards for Registration).

From a government policy perspective, it is hoped that such a *Code of Professional Practice* will enhance consistency of judgements across assessors within the VET sector, and ultimately improve comparability of standards across RTOs nationally.

2. Validation and Moderation

The two terms validation and moderation have been used interchangeably in the VET sector; and whilst each are based on similar processes, there are a number of distinctive features. These have been outlined in Table 2.

Table 2: The distinctive features of validation and moderation.

Features	Validation	Moderation
Assessment Quality Management Type	Quality Review	Quality Control
Primary Purpose	Continuous improvement	Bring judgements and standards into alignment
Timing	On-going	Prior to the finalisation of candidate results
Focus	Assessment Tools; and Candidate Evidence (including assessor judgements) (desirable only)	Assessment tools; and Candidate Evidence, including assessor judgements (mandatory)
Type of Approaches	Assessor Partnerships Consensus Meetings External (validators or panels)	Consensus Meetings External (moderators or panels) Statistical
Outcomes	Recommendations for future improvements	Recommendations for future improvements; and Adjustments to assessor judgements (if required)

2.1 Assessment quality management type

Validation: Quality Review

Validation is a quality review process. It involves checking that the assessment tool[3] produced valid, reliable, sufficient, current and authentic evidence to enable reasonable judgements to be made as to whether the requirements of the relevant aspects of the Training Package or accredited course had been met. It includes reviewing and making recommendations for future improvements to the assessment tool, process and/or outcomes.

---

[3] An assessment tool includes the following components: the context and conditions for the assessment, the tasks to be administered to the candidate, an outline of the evidence to be gathered from the candidate and the criteria used for judging the quality of performance (i.e. the assessment decision making rules). It also includes the administration, recording and reporting requirements.

Moderation: Quality Control

Moderation is the process of bringing assessment judgements and standards into alignment. It is a process that ensures the same standards are applied to all assessment results within the same Unit(s) of Competency. It is an active process in the sense that adjustments to assessor judgements are made to overcome differences in the difficulty of the tool and/or the severity of judgements.

2.2 Primary Purpose

Validation: Continuous Improvement

The Australian Quality Training Framework has a strong emphasis on continuous improvement of assessment practices, processes, tools and judgements. Under the AQTF, validation is a vital aspect of the quality review processes of any training organisation. It ensures that there is continuous improvement in the assessment undertaken by a training organisation. It also provides valuable opportunities for assessors to share ideas, experiences and tools with other assessors.

Moderation: Bringing judgements and standards into alignment

When there are high stakes associated with the assessment outcomes, moderation can help ensure comparability of standards by bringing judgements into alignment. High stakes refers to situations where the consequences of making a wrong judgement are high (e.g. assessing someone as competent, when in actual fact they are not yet competent[4]). Such consequences may be associated with the safety, reputation, equity and/or financial considerations of the candidate, assessor, RTO and/or employer. If the consequences are high, then it is more likely that some form of quality control on assessment outcomes may need to be implemented.

For example, the AQTF gives assessors flexibility in deciding what assessment tools they will use to assess Units of Competency specified within the relevant industry Training Package, as well as the autonomy to design their own tasks and tools. As a result, the assessments from two different assessors will often be based on two different sets of assessment tools, although they will be assessing the same Unit(s) of Competency. In some cases, the assessment tools set by one assessor may be easier than those set by another assessor. This is referred to as differences in the *difficulty of the tool*. Differences can also be present in the *severity of assessor judgements*. For example, an assessor may be stricter or more lenient in his/her expectations of the standard to be achieved than his/her peers, despite the fact that both assessments have been designed

---

[4] This is referred to a false positive assessment outcome. A false negative assessment outcome occurs when someone is assessed as not yet competent, but in actual fact, they are competent. In many instances, an appeals process will detect false negative assessment outcomes. Moderation can help minimise the risk of false positive assessment outcomes.

to measure the same Unit of Competency[5]. In an organisation where assessment tools are easier and/or assessor judgements more lenient, candidates would receive a different result for the same standard of work than they would in another organisation where the tools and/or judgements were harder. If differences in the difficulty of tools and/or the severity of judgements are not taken into account when finalising candidates' results, some candidates would be treated unfairly. Moderation is therefore a quality control mechanism that can be used to help achieve comparability of standards, which in turn, can help to achieve fairness for all candidates.

2.3 Timing

Validation: Ongoing

Whilst validation is an ongoing activity, it can be most powerful when there is evidence available for review concerning the extent to which the assessment tool produced valid, reliable, sufficient, current and authentic evidence across contexts and over time.[6.]

Moderation: Prior to the finalisation of candidate results

Moderation usually occurs toward the end of the assessment process when judgements of candidate evidence have been made, but prior to the final recording and reporting of results.

2.4 Focus

Validation: Assessment Tools and Candidate Evidence including assessor judgements (if available)

Validation typically focuses on reviewing assessment tools and, if available, candidate evidence to make recommendations for future improvements. To ensure the process is manageable and cost-effective, validation is often based upon samples of assessment materials such as:

- Assessment tools and/or judged candidate evidence of a specific cohort of assessors;
- Assessment tools and/or judged candidate evidence of a selection of units/qualifications within a Training Package; and/or
- A selection of judged candidate evidence at varying levels of achievement (e.g. borderline cases) within a qualification/Training Package.

---

[5]As competency based assessments tend to be based upon performance tasks that rely on the professional judgement by the assessor (in terms of the interpretation of the evidence collected), differences in the severity of the judgements can occur even when the same assessment tools are used.

[6] Within this Code, the review of assessment tools (via panelling, piloting and/or trialling) in advance of the actual assessment has been classified as a 'quality assurance' approach to assessment quality management as it is primarily concerned with establishing appropriate circumstances for assessment to take place.

Validation may include checking whether the tool:

- Has clear, documented evidence of the procedures for collecting, synthesising, judging and recording outcomes (i.e. to help improve the consistency of assessments across assessors [inter-rater reliability]);
- Has evidence of content validity (i.e. whether the assessment task(s) as a whole, represents the full range of knowledge and skills specified within the Unit(s) of Competency);
- Reflect work-based contexts, specific enterprise language and job-tasks and meets industry requirements (i.e. face validity);
- Adheres to the literacy and numeracy requirements of the Unit(s) of Competency (construct validity);
- Has been designed to assess a variety of evidence over time and contexts (predictive validity);
- Has been designed to minimise the influence of extraneous factors (i.e. factors that are not related to the unit of competency) on candidate performance (construct validity);
- Have clear decision making rules to ensure consistency of judgements across assessors (inter-rater reliability) as well as consistency of judgements within an assessor (intra-rater reliability);
- Has a clear instruction on how to synthesise multiple sources of evidence to make an overall judgement of performance (inter-rater reliability);
- Has evidence that the principles of fairness and flexibility have been adhered to;
- Has been designed to produce sufficient, current and authentic evidence;
- Is appropriate in terms of the level of difficulty of the task(s) to be performed in relation to the skills and knowledge specified within the relevant unit(s) of competency;
- Has outlined appropriate reasonable adjustments that could be made to the gathering of assessment evidence for specific individuals and/or groups; and
- Has adhered to the relevant organisation assessment policy.

Validation may also include checking the appropriateness of assessor judgements using samples of evidence of candidate performance. This would require checking whether the judgement was too harsh or too lenient by reviewing the evidence provided against the:

- Requirements set out in the Unit(s) of Competency;
- Benchmark samples of candidate evidence at varying levels of achievement (including borderline cases); and the
- Assessment decision making rules specified within the assessment tools.

Moderation: Assessment Tools, Candidate Evidence and Assessor Judgements (mandatory)

Whilst moderation can include the same processes outlined above, the major difference between the two approaches is that moderation *must* include samples of judged candidate evidence; whereas with validation, it is only desirable.

2.5 Type of Approach

There are a number of different approaches to conducting validation and/or moderation, which differ in complexity and level of scrutiny. Four broad approaches are considered next.

Assessor Partnerships (Validation only)

Assessor partnerships involve the sharing of assessment tools and outcomes within a small group of assessors, possibly even just two assessors. Often this type of approach to quality review is informal and self-managed. The focus is on collegiality, mutual assistance and confirmation. The partnership may involve:

- Sharing and discussing one another's assessment tools, processes and outcomes;
- Providing mutual support for reviewing one another's assessment tools;
- Assisting one another in resolving any problems and/or issues (e.g. appeals); and
- Checking one another's judgements of candidate performance against the Unit(s) of competency and/or decision making rules specified within the assessment tool.

A major benefit of assessor partnerships is that they can be locally organised and tend to have minimal implementation and maintenance costs. They can also be personally empowering to participants and can help build confidence and expertise of less experienced assessors. However, as partnerships tend to be locally self-managed, other quality review mechanisms *may* also be required to ensure continuous improvement of assessment practices. This is because there is a possibility that some assessor partnerships may simply reinforce each others' misconceptions and mistakes if there are no other quality review processes available.

Consensus Meetings (Validation and Moderation)

Typically consensus meetings involve assessors reviewing their own and their colleagues' assessment tools and outcomes as part of a group. It can occur within and/or across organisations. It is typically based on agreement within a group on the appropriateness of the assessment tools and assessor judgements for a particular unit(s) of competency. A major strength of consensus meetings is that assessors are directly involved in all aspects of assessment and gain professionally by learning not only how and what to assess, but what standards to expect from their candidates. It also enables

assessors to develop strong networks and promotes collegiality. Another benefit from consensus meetings is that it provides opportunity for sharing materials/resources among assessors. If used for moderation purposes, consensus meetings however provide less quality control than external and statistical approaches as again, they can be influenced by local values and expectations.

External Approaches (Validation and Moderation)

There are various external approaches to assessment validation and moderation. One approach would be for an external person (or a panel of people) to visit the organisation to judge the way in which candidates' evidence were collected and judged against the Unit(s) of Competency. Differences between the local and external assessment judgements could then be either:

- Discussed and reconciled accordingly (i.e. if conducted for moderation purposes); and/or
- Discussed to identify ways in which improvements to future assessment practices could be undertaken (i.e. if conducted for validation purposes).

An alternative external approach would be for samples of assessment tools and/or judged candidate evidence to be sent to a central location for specialist assessors to review directly against the Unit(s) of Competency. The specialist external assessors could be representatives of the relevant national Industry Skills Council (ISC) and/or the relevant state/territory registering bodies. Again, differences between the organisation and the external-based assessments could then be discussed (e.g. for validation) and/or reconciled (e.g. for moderation) at a distance.

There are a number of benefits from using external moderators/validators. These include the potential to:

- Offer authoritative interpretations of the standards specified within Units of Competency;
- Improve consistency of the standards across locations by identifying local bias and/or misconceptions (if any);
- Offer advice to organisations and assessors on assessment approaches and procedures; and
- Observe actual assessment processes in real time as opposed to simply reviewing assessment products (if site visits are included).

In relation to moderation, although external approaches have greater quality control over the assessment processes and outcomes than consensus meetings, they have less quality control than statistical approaches.

Statistical (moderation)

Although yet to be pursued at the national level within the VET Sector, statistical moderation could be used to ensure that RTO based assessments are comparable throughout the nation, particularly if grades or marks are to be reported. However, to implement this moderation process, some form of a common assessment task(s) would need to be introduced at a national level in the VET sector (e.g. external exam or standardised assessment tools) to moderate the organisation-based assessments.

If a common assessment task was used to statistically moderate organisation-based assessments, the statistical moderation process would maintain the rank order of the candidates' scores (as determined by the assessor/organisation) but it would bring the distributions of scores across groups of candidates (from other organisations or assessors) within the same units within a qualification into alignment. That is, statistical moderation adjusts the organisation-based assessments in accordance with candidates' performances on common external tasks. It should be acknowledged that any adjustment to a candidate's scores is determined by the external scores for the whole organisation's cohort, not by the candidate's own external score. It is also important to note that statistical moderation does not change the rank order of candidates, as determined by the organisation's scores. A candidate given the top score for an assessment task by his/her organisation would have the top score after statistical moderation, no matter how they performed on the external task.

The process recognises that organisations are in the best position to make comparative judgements about the performance of their candidates and these comparative judgements are not changed as a result of the statistical moderation.

Statistical moderation entails adjusting the level and spread of each organisation's assessments of its candidates in a particular qualification, to match the level and spread of the same candidates' scores on a common external task. If a common assessment task was to be completed by all candidates across the nation or within an industry area, it could become the common standard against which organisation's assessments could be compared. At a national level, the organisation-based assessments could be statistically moderated using:

- A common exam across all qualifications based on measuring generic/employability skills;
- Qualification specific national exams (similar to those used for licensing purposes); and
- National common assessment tools within each qualification that would need to be judged centrally.

The major benefit of statistical moderation is that it provides the strongest form of quality control over organisation-based assessments. It can also be less expensive to implement and maintain (if paper-based) than external moderation processes. It would however require the introduction of some form of common assessment task(s) at the national level. If the common assessment task was paper-based (as has been typically implemented in other educational sectors due to reduced costs associated with the implementation and scoring procedures), then any adjustments to candidate results would be limited to estimates of candidates' cognitive skills (i.e. knowledge and understanding); and therefore may have limited face and content validity within the VET sector.

2.6 Outcomes

Validation: Recommendations for future improvements.

The outcomes of a validation process generally result in recommendations for improvements to the assessment tool, process and outcomes. This may include making recommendations for changes to the:

- Context and conditions for the assessment;
- Task(s) to be administered to the candidates;
- Administration instructions;
- Criteria used for judging the quality of performance (e.g. the decision making rules, evidence requirements etc);
- Guidelines for making reasonable adjustments to the way in which the evidence of performance was gathered to ensure that the expected standard of performance specified within the Unit(s) of Competency has not been altered; and the
- Recording and reporting requirements.

Each recommendation should include some form of justification.

Moderation: Recommendations for future improvement and adjustments to assessor judgements (if required).

In addition to making recommendations for improvement to the assessment tools (as outlined above), consensus or external moderation may also include:

- Actioning the assessor/RTO to adjust the results of a specific cohort of candidates prior to the finalisation of results; and
- Requesting copies of final candidate assessment results in accordance with recommended actions.

In relation to statistical moderation, candidate results are automatically adjusted (where required) to bring standards into alignment. The final [adjusted] results are then reported to the key stakeholders (e.g. assessor, candidate, RTO etc).

2.7 Summary

In summary, the major distinguishing features between validation and moderation are that:

- Validation is concerned with quality review whilst moderation is concerned with quality control;
- The primary purpose of moderation is to help achieve comparability of standards across organisations whilst validation is primarily concerned with continuous improvement of assessment practices and outcomes;
- Whilst validation and moderation can both focus on assessment tools, moderation requires access to judged (or scored) candidate evidence. The latter is only desirable for validation;
- Both consensus and external approaches to validation and moderation are possible. Moderation can also be based upon statistical procedures whilst validation can include less formal arrangements such as assessor partnerships; and
- The outcomes of validation are in terms of recommendations for future improvement to the assessment tools and/or processes; whereas moderation may also include making adjustments to assessor judgements to bring standards into alignment, where determined necessary.

3. Principles

The following principles should underpin validation and/or moderation within the VET sector.

3.1 Transparent

The purpose, process and implications of validation and/or moderation should be transparent to all relevant stakeholders.

This principle can be enhanced if:

- It is made explicit to assessors the purpose, approach and potential outcomes;
- The approach to be implemented is clearly delineated and communicated to relevant stakeholders; and
- The justification for the outcomes recommended (validation) and/or imposed (moderation) are clearly documented and made available to assessors.

3.2 Representative

It is not possible or necessary to validate and/or moderate every possible assessment tool or piece of candidate evidence within an RTO at one time. A representative sample should therefore be used to validate and moderate assessment tools and judgements. A properly selected representative sample can identify any issues with assessment practices and decisions.

This principle can be enhanced if:

- A sampling framework is designed in which risk indicators are identified that may impact on the assessment process and/or outcomes, and such indicators are targeted for selection; and
- There is an element of random selection.

3.3 Confidential

Information regarding individuals (i.e. assessors and candidates) and providers must be treated with sensitivity and discretion. Confidentiality should be observed in relation to the identity of the assessors (i.e. those who developed the assessment tools and/or made the judgements) and candidates (i.e. those whose evidence is submitted in the process). This allows the validation and/or moderation process to focus on the quality of the assessment tools and the assessment judgements rather than the individuals involved.

This principle can be enhanced if:
- De-identified samples of candidates' work and assessors' tools are used; and
- The outcomes of the process are given in a private, supportive environment.

3.4 Educative

Validation and/or moderation should form an integral rather than separate part of the assessment process. It should provide constructive feedback, which leads to continuous improvement across the organisation.

This principle can be enhanced if:
- The process is supportive and positive for assessors, validators and/or moderators;
- The process and outcomes provide the basis for individuals as well as organisations to monitor and reflect on their own practice;
- The rationales behind recommendations for alterations and/or adjustments are made explicit to assessors;
- Recommendations for improvement to the assessment tool and/or decision making process are succinct, constructive and explicit; and
- Professional development support is available for assessors.

3.5 Equitable

Validation and/or moderation must be demonstrably fair, equitably applied and unbiased.

This principle can be enhanced if:
- There are clear and effective policies and mechanisms for the appeal or review of moderation outcomes by key stakeholders, in circumstances in which an appeal or review is appropriate; and
- Confidentiality of evidence can be assured.
- The process is sensitive to assessor and candidate diversity and has no inherent biases.

3.6 Tolerable

Any assessment includes a margin of error. The way in which evidence is gathered and interpreted against the standards will vary. The challenge is to limit the variation to acceptable proportions. Validation and/or moderation enables the variation to be identified and limited to what is tolerable.

This principle can be enhanced if:
- Benchmark samples of borderline cases are used as points of reference;
- Exemplar tools are made available to assessors as well as validators/moderators; and
- A risk assessment has been undertaken of the implications of a false positive judgement (i.e. assessing someone as competent when in actual fact they are not yet competent) and a false negative judgement (i.e. assessing someone as not yet competent when in actual fact the person is competent).

4. Glossary of Terms

Assessment quality management	Processes that could be used to help achieve comparability of standards. Typically, there are three major components to quality management of assessments: quality assurance, quality control and quality review.
Assessment tool	An assessment tool includes the following components: the context and conditions for the assessment, the tasks to be administered to the candidate, an outline of the evidence to be gathered from the candidate and the evidence criteria used to judge the quality of performance (i.e. the assessment decision making rules). It also includes the administration, recording and reporting requirements.
Assessor	In this *Code*, an assessor means an individual or organisation responsible for the assessment of Units of Competency in accordance with the Australian Quality Training Framework.
Authenticity	One of the rules of evidence. To accept evidence as authentic, an assessor must be assured that the evidence presented for assessment is the candidate's own work.
Benchmark	Benchmarks are a point of reference used to clarify standards in assessment. They are agreed good examples of particular levels of achievement which arise from the moderation process. Benchmarks help clarify the standards expected within the qualification, and illustrate how they can be demonstrated and assessed. They can also identify new ways of demonstrating the competency.
Comparability of standards	Comparability of standards are said to be achieved when the performance levels expected (e.g. competent/not yet competent decisions) for a unit (or cluster of units) of competency are similar between assessors assessing the same unit(s) in a given RTO and between assessors assessing the same unit(s) across RTOs.
Competency based Assessment	Competency based assessment is a purposeful process of systematically gathering, interpreting, recording and communicating to stakeholders, information on candidate development against industry competency standards and/or learning outcomes.
Consensus Meetings	Typically consensus meetings involve assessors reviewing their own and their colleagues' assessment tools and outcomes as part of a group. It can occur within and/or across organisations. It is typically based on agreement within a group on the appropriateness of the assessment tools and assessor judgements for a particular unit(s) of competency.
Consistency of evidence	The evidence gathered needs to be evaluated for its consistency with other assessments of the candidate's performance, including the candidate's usual performance levels.
Construct validity	The extent to which certain explanatory concepts or constructs account for the performance on a task. It is concerned with the degree to which the evidence collected can be used to infer competence in the intended area, without being influenced by other non-related factors (eg literacy levels).
Content validity	The match between the required knowledge and skills specified in the competency standards and the assessment tool's capacity to collect such evidence.
Continuous Improvement	A planned and ongoing process that enables an RTO to systematically review and improve its policies, procedures, services or products to generate better outcomes for clients and to meet changing needs. It allows the RTO to constantly review its performance against the *AQTF 2007 Essential Standards for Registration* and to plan ongoing improvements. Continuous improvement involves collecting, analysing and acting on relevant information from clients and other interested parties, including the RTO's staff.
Currency	One of the rules of evidence. In assessment, currency relates to the age of the evidence presented by the candidate to demonstrate that they are still competent. Competency requires demonstration of current performance, so the evidence must be from either the present or the very recent past.

Decision making rules	The rules to be used to make judgements as to whether competency has been achieved (note that if grades or scores are also to be reported, the scoring rules should outline how performance is to be scored). Such rules should be specified for each assessment tool. There should also be rules for synthesising multiple sources of evidence to make overall judgements of performance.
De-identified samples	This is a reversible process in which identifiers are removed and replaced by a code prior to the validation/moderation session. At the completion of the session, the codes can be used to link back to the original identifiers and identify the individual to whom the sample of evidence relates.
Face validity	The extent to which the assessment tasks reflect real work-based activities.
Fairness	One of the principles of assessment. Fairness in assessment requires consideration of the individual candidate's needs and characteristics, and any reasonable adjustments that need to be applied to take account of them. It requires clear communication between the assessor and the candidate to ensure that the candidate is fully informed about, understands and is able to participate in, the assessment process, and agrees that the process is appropriate. It also includes an opportunity for the person being assessed to challenge the result of the assessment and to be reassessed if necessary.
Flexibility	One of the principles of assessment. To be flexible, assessment should reflect the candidate's needs; provide for recognition of competencies no matter how, where or when they have been acquired; draw on a range of methods appropriate to the context, competency and the candidate; and support continuous competency development.
Inter-rater reliability	A type of reliability which is concerned with determining consistency of judgement across different assessors using the same assessment task and procedure.
Intra-rater reliability	A type of reliability concerned with determining the consistency of assessment judgements by the same assessor. That is, the consistency of judgements across time and location, and using the same assessment task administered by the same assessor.
Moderation	Moderation is the process of bringing assessment judgements and standards into alignment. It is a process that ensures the same standards are applied to all assessment results within the same Unit(s) of Competency. It is an active process in the sense that adjustments to assessor judgements are made to overcome differences in the difficulty of the tool and/or the severity of judgements.
Moderator	In this *Code* moderator means a person responsible for carrying out moderation processes. A moderator may be external or internal to the organisation.
Panelling of assessment tools	A quality assurance process for checking the relevance and clarity of the tool prior to use with other colleagues (i.e. who have expertise within the Units of Competency and/or assessment tool development). This may involve examining whether the content of the tool is correct and relevant to industry, the unit(s) f; the instructions are clear for candidates and assessors and that there is not potential bias within the design of the tool.
Piloting of assessment tools	A quality assurance process for checking the appropriateness of the tool with representatives from the target group This may involve administering the tool with a small number of individuals (who are representative of the target group) and gathering feedback on both their performance and perceptions of the task. Piloting can help determine the appropriateness of the amount of time to complete the task, the clarity of the instructions, the task demands (i.e. whether it is too difficult or easy to perform) and its perceived relevance to the workplace.
Predictive validity	A form of criterion validity concerned with the ability of the assessment outcomes to accurately predict the future performance of the candidate.
Principles of assessment	To ensure quality outcomes, assessments should be: ▪ Fair ▪ Flexible ▪ Valid ▪ Reliable

- Sufficient.

Quality assurance	Concerned with establishing appropriate circumstances for assessment to take place. It is an input approach to assessment quality management.
Quality control	Focuses on monitoring, and where necessary making adjustments to judgements made by assessors prior to the finalisation of assessment results/outcomes. This approach therefore involves the direct management of assessment judgements to *ensure* consistency in the interpretation and application of the competency standards. As it occurs prior to the finalisation of the result, in which alterations can be made to assessor judgements, it is referred to as an 'active process' to assessment quality management.
Quality review	Concerned with the review of the assessment tools, procedure and outcomes to make improvements for future use. It is referred to as a retrospective approach to assessment quality management.
Reasonable adjustments	Adjustments that can be made to the way in which evidence of candidate performance can be collected. Whilst reasonable adjustments can be made in terms of the way in which evidence of performance is gathered, the evidence criteria for making competent/not yet competent decisions [and/or awarding grades] should not be altered in any way. That is, the standards expected should be the same irrespective of the group and/or individual being assessed, otherwise comparability of standards will be compromised.
Reliability	One of the principles of assessment. There are five types of reliability: internal consistency, parallel forms, split-half, inter-rater and intra rater. In general, reliability is an estimate of how accurate or precise the task is as a measurement instrument. Reliability is concerned with how much error is included in the evidence.
Risk Assessment	Concerned with gauging the likelihood of unexpected and/or unfortunate consequences. For example, determining the level of risk (e.g. in terms of safety, costs, equity etc) of assessing someone as competent when in actual fact they are not competent, and or vice versa.
Risk Indicators	The potential factors that may increase the risk associated with the assessment. These factors should be considered when selecting a representative sample for validation and/or moderation. Risk factors may include safety (e.g. potential danger to clients from an incorrect judgement), equity (e.g. outcomes impacting on highly competitive selection procedures), human capacity (e.g. experience and expertise of assessors) etc.
Rules of evidence	These are closely related to the principles of assessment and provide guidance on the collection of evidence to ensure that it is valid, sufficient, authentic and current.
Sampling	Sampling is the process of selecting material to use in the validation and/or moderation.
Stakeholders	Individuals or organisations affected by, or who may influence, the assessment outcomes. These may include candidates, assessors, employers, other RTOs etc. Each stakeholder group will have their own reporting needs in relation to the outcomes of the assessment.
Statistical Moderation	Is a process for adjusting organisation-based assessments to the same standard, while maintaining the candidates' rank order given by an organisation. To implement this moderation process, some form of common assessment tasks would need to be introduced into the VET sector.
Sufficiency	One of the principles of assessment and also one of the rules of evidence. Sufficiency relates to the quality and quantity of evidence assessed. It requires collection of enough appropriate evidence to ensure that all aspects of competency have been satisfied and that competency can be demonstrated repeatedly. Supplementary sources of evidence may be necessary. The specific evidence requirements of each Unit of Competency provide advice on sufficiency.
Target group	This refers to the group of individuals that the assessment tool has been designed for. The description of the target group could include any background characteristics of the group (such as literacy and numeracy) that may assist other assessors to determine whether the tool could be applied to other similar groups of individuals.
Trialling of assessment tools	A quality assurance process for checking that the assessment tool will produce valid and reliable evidence to satisfy the purpose of the assessment and the reporting needs of the key stakeholder groups. A trial is often referred to as a

'dress rehearsal' in which the tool is administered to a group of individuals who are representative of the target group. The information gathered from the trial can be used to determine the cost-effectiveness, fairness, flexibility, validity and reliability of the assessment prior to use.

Unit of Competency

Specification of industry knowledge and skill and the application of that knowledge and skill to the standard of performance expected in the workplace.

Validation

Validation is a quality review process. It involves checking that the assessment tool[7] produced valid, reliable, sufficient, current and authentic evidence to enable reasonable judgements to be made as to whether the requirements of the relevant aspects of the Training Package or accredited course had been met. It includes reviewing and making recommendations for future improvements to the assessment tool, process and/or outcomes.

Validator

In this *Code* a validator refers to a person responsible for carrying out validation. The validator may be internal or external to the organisation.

Validity

One of the principles of assessment. There are five major types of validity: face, content, criterion (predictive and concurrent), construct and consequential. In general, validity is concerned with the appropriateness of the inferences, use and consequences that result from the assessment. In simple terms, it is concerned with the extent to which an assessment decision about a candidate (e.g. competent/not yet competent, a grade and/or a mark), based on the evidence of performance by the candidate, is justified. It requires determining conditions that weaken the truthfulness of the decision, exploring alternative explanations for good or poor performance, and feeding them back into the assessment process to reduce errors when making inferences about competence. Unlike reliability, validity is not simply a property of the assessment tool. As such, an assessment tool designed for a particular purpose and target group may not necessarily lead to valid interpretations of performance and assessment decisions if the tool was used for a different purpose and/or target group.

---

[7] An assessment tool includes the following components: the context and conditions for the assessment, the tasks to be administered to the candidate, an outline of the evidence to be gathered from the candidate and the evidence criteria used to judge the quality of performance (i.e. the assessment decision making rules). It also includes the administration, recording and reporting requirements.