TVET AUSTRALIA

NATIONAL QUALITY COUNCIL

GUIDE FOR DEVELOPING ASSESSMENT TOOLS

2009

Table of Contents

Table of Contents 2
List of Tables 3
Introduction 4
1. Tool Components 5
2. Ideal Characteristics 6
2.1 Portfolio 8
2.2 Observation Methods 10
2.3 Product Based Methods 13
2.4 Interview Methods 16
3. Quality Checks 18
Appendix 20
A.1 Assessment Tool: Self Assessment 21
A.2 Assessment Tool: Competency Mapping 23
Glossary of Terms 24

This work has been produced by the Work-based Education Research Centre of Victoria University in conjunction with Bateman and Giles Pty Ltd as part of a project commissioned by the National Quality Council in 2008 with funding provided through the Australian Government Department of Education Employment and Workplace Relations and state and territory governments.

© Commonwealth of Australia 2009

AEShareNet P PRESERVE INTEGRITY www.aesharenet.com.au/P4/

This document is available under a "Preserve Integrity" licence – see http://www.aesharenet.com.au/P4 for details. All other rights reserved. For licensing enquiries contact sales@tvetaustralia.com.au.

List of Tables

TABLE 1: IDEAL CHARACTERISTICS OF AN ASSESSMENT TOOL 7
TABLE 2: PORTFOLIO: EXEMPLAR ASSESSMENT TOOL FEATURES. 8
TABLE 3: WORKPLACE OBSERVATION: EXEMPLAR ASSESSMENT TOOL FEATURES. 10
TABLE 4: PRODUCT BASED METHODS: EXEMPLAR ASSESSMENT TOOL FEATURES. 13
TABLE 5: INTERVIEW: EXEMPLAR ASSESSMENT TOOL FEATURES. 16

Introduction

This Guide is a practical resource material for assessors and assessor trainers seeking technical guidance on how to develop and/or review assessment tools. The Guide is not intended to be mandatory, exhaustive or definitive but instead it is intended to be aspirational and educative in nature.

There are three sections to this Guide. **Section 1** explains what an assessment tool is, including its essential components.

Section 2 identifies a number of ideal characteristics of an assessment tool and provides four examples of how each of these characteristics can be built into the design for four methods of assessment: observation, interview, portfolio and product-based assessments. These four examples encapsulate methods that require candidates to either do (observation), say (interview), write (portfolio) or create (product) something. In fact, any assessment activity can be classified according to these four broad methods.

Section 3 provides an overview of three quality assurance processes (i.e. panelling, piloting and trialling) that could be undertaken prior to implementing a new assessment tool.

There is also an appendix that contains the following two exemplar templates for assessors:

- *Assessment Tool: Self Assessment:* A self assessment checklist for the assessor to check that s/he has included within his/her tool the administration, decision making, recording and reporting conditions of the tool. The self assessment could be subsequently used by the panel during the consensus meeting (if so, the checklist would need to be attached to the tool); and
- *Competency Mapping Tool:* A template to assist assessors with mapping the key components within their task to the Unit(s) of Competency to demonstrate content validity. This should be attached to the assessment tool for validation purposes. Note that multiple copies may need to be produced for each task within an assessment tool.

Finally, as a number of technical assessment concepts have been referred to throughout this Guide, a Glossary of Terms has been included.

1. Tool Components

According to the *AQTF Essential Standards for Registration,* an assessment tool is defined as.

The instrument(s) and procedures used to gather and interpret evidence of competence:

a) Instrument- the specific questions or activity used to assess competence by the assessment method selected. An assessment instrument may be supported by a profile of acceptable performance and the decision-making rules or guidelines to be used by the assessors.

b) Procedures – the information or instructions given to the candidate and the assessor about how the assessment is to be conducted and recorded.

In accordance with the *AQTF Essential Standards for Registration*, an assessment tool includes the following components:

- The learning or competency unit(s) to be assessed;
- The target group, 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;
- The evidence criteria used to judge the quality of performance (i.e. the assessment decision making rules); as well as
- The administration, recording and reporting requirements.

To assist with validation and/or moderation, the tool should also provide evidence of how validity and reliability have been tested and built into the design and use of the tool.

In some instances, all the components within the assessment tool may not necessarily be present within the same document. That is, it is not necessary that the hard copy tool holds all components. It may be that the tool makes reference to information in another document/material/tool held elsewhere. This would help avoid repetition across a number of tools (e.g. the context, as well as the recording and reporting requirements of the tool may be the same for a number of tools and therefore, may be just cited within one document but referred to within all tools).

The quality test of any assessment tool is the capacity for another assessor to use and replicate the assessment procedures without any need for further clarification by the tool developer. That is, it should be a stand-alone assessment tool.

2. Ideal Characteristics

A number of ideal characteristics of an assessment tool have been provided in Table 2. This is referred to hereon as the 'assessment tool framework.' The framework could be used by:

- Assessors during tool development (refer to Template A.1 in the appendix for an example of a self-assessment checklist); as well as
- Members of a Consensus Group during a validation and/or moderation meeting (refer to the Implementation Guide: Validation and Moderation for an example of how it could be used to review assessment tools post assessment).

Following Table 1, four examples have been included in this *Guide* to illustrate how the assessment tool framework could be applied to the development of assessment tools. It should be acknowledge that the examples provided are not assessment tools but instead, are guidance as to what key features should be in an assessment tool based on the specific assessment method. These four examples encapsulate methods that require candidates to either do (observation), say (interview), write (portfolio) and create (build) something. In fact, any assessment activity can be classified according to these four broad methods.

Table 1: Ideal Characteristics of an Assessment Tool

Component	Description
The context	The target group and purpose of the tool should be described. This should include a description of the background characteristics of the target group that may impact on the candidate performance (e.g. literacy and numeracy requirements, workplace experience, age, gender etc).
Competency Mapping	The components of the Unit(s) of Competency that the tool should cover should be described. This could be as simple as a mapping exercise between the components of the task (e.g. each structured interview question) and components within a Unit or cluster of Units of Competency. The mapping will help to determine the sufficiency of the evidence to be collected. An example of how this can be undertaken has been provided in Template A.2 (refer to the Appendix).
The information to be provided to the candidate	Outlines the task(s) to be provided to the candidate that will provide the opportunity for the candidate to demonstrate the competency. It should prompt them to say, do, write or create something.
The evidence to be collected from the candidate	Provides information on the evidence to be produced by the candidate in response to the task.
Decision making rules	The rules to be used to: • Check evidence quality (i.e. the rules of evidence); • Judge how well the candidate performed according to the standard expected (i.e. the evidence criteria); and • Synthesise evidence from multiple sources to make an overall judgement.
Range and conditions	Outlines any restriction or specific conditions for the assessment such as the location, time restrictions, assessor qualifications, currency of evidence (e.g. for portfolio based assessments), amount of supervision required to perform the task (i.e. which may assist with determining the authenticity of evidence) etc.
Materials/resources required	Describes access to materials, equipment etc that may be required to perform the task.
Assessor intervention	Defines the amount (if any) of support provided.
Reasonable adjustments (for enhancing fairness and flexibility)	This section should describe the guidelines for making reasonable adjustments to the way in which evidence of performance is gathered (e.g. in terms of the information to be provided to the candidate and the type of evidence to be collected from the candidate) without altering the expected performance standards (as outlined in the decision making rules).
Validity evidence	Evidence of validity (such as face, construct, predictive, concurrent, consequential and content) should be provided to support the use of the assessment evidence for the defined purpose and target group of the tool.
Reliability evidence	If using a performance based task that requires professional judgement of the assessor, evidence of reliability could include providing evidence of: • The level of agreement between two different assessors who have assessed the same evidence of performance for a particular candidate (i.e. inter-rater reliability); and • The level of agreement of the same assessor who has assessed the same evidence of performance of the candidate, but at a different time (i.e. intra-rater reliability). If using objective test items (e.g. multiple choice tests) than other forms of reliability should be considered such as the internal consistency of a test (i.e. internal reliability) as well as the equivalence of two alternative assessment tasks (i.e. parallel forms).
Recording requirements	The type of information that needs to be recorded and how it is to be recorded and stored, including duration.
Reporting requirements	For each key stakeholder, the reporting requirements should be specified and linked to the purpose of the assessment.

2.1 PORTFOLIO

Note a portfolio is defined here as a purposeful collection of samples of annotated and validated pieces of evidence (e.g. written documents, photographs, videos, audio tapes).

Table 2: Portfolio: Exemplar Assessment Tool Features.

The context	The purpose and target group should be described	The target group is XXX candidates undertaking the Certificate of XXX. This tool assists with assessing the candidate's application of knowledge and skills and will need to be assessed in conjunction with XXX (e.g. interview) to ensure adequate coverage of the entire Unit of Competency.
Competency Mapping	Map key components of task to the Units(s) of Competency (content validity) – refer to Template A.2	The assessment criteria used to evaluate the contents of the portfolio should be mapped directly against the Unit(s) of Competency. This will help to determine the sufficiency of the evidence to be collected and determine whether any other aspects of the Unit(s) of competency need to be collected elsewhere.
Information to candidate	Outline the task to be provided to the candidate that will provide the opportunity for the candidate to demonstrate the competency. It should prompt them to say, do, write or create something.	The tool should provide instructions to the candidate for how the portfolio should be put together. For example, the candidate may be instructed to: • Select the pieces of evidence to be included; • Provide explanations of each piece of evidence; • Include samples of evidence only if they take on new meaning within the context of other entries; • Include evidence of self-reflection; • Map each piece of evidence to the Unit(s) of Competency; • Include evidence of growth or development; and • Include a table of contents for ease of navigation.
Evidence from candidate	Provides information on the evidence to be produced by the candidate in response to the task.	The instructions for submitting the portfolio should be included here as well as a description of the evidence criteria that would be used to assess the portfolio.

Decision making rules	The rules to be used to:
	• check evidence quality (i.e. the rules of evidence)
	• judge how well the candidate performed according to the standard expected (i.e. the evidence criteria); and
	• synthesise evidence from multiple sources to make an overall judgement
		This section should outline the procedures for checking the appropriateness and trustworthiness of the evidence included within the portfolio such as the:
		• Currency of evidence within the portfolio - is the evidence relatively recent).The rules for determining currency would need to be specified here (e.g. less than five years);
		• Authenticity -is there evidence included within the portfolio that verifies that the evidence is that of the candidate and/or if part of a team contribution, what aspects were specific to the candidate (e.g. testimonial statements from colleagues, opportunity to verify qualifications with issuing body etc);
		• Sufficiency - is there enough evidence to demonstrate to the assessor competence against the entire unit of competency, including the critical aspects of evidence described in the Evidence Guide (e.g. evidence of consistency of performance across time and contexts);
		• Content Validity – does the evidence match the unit of competency (e.g. relevance of evidence and justification by candidate for inclusion, as well as annotations and reflections);
		• Once the evidence within the portfolio has been determined to be trustworthy and appropriate, the evidence will need to be judged against evidence criteria such as:
		  • Profile descriptions of varying levels of achievement (e.g. competent versus not yet competent performance (also referred to as standard referenced frameworks[1]);
		  • Behaviourally Anchored Rating Scales (BARS)[2] that describe typical performance from low to high (also referred to as analytical rubrics); and
		• The Unit of Competency presented in some form of a checklist.
		
		The outcomes of the portfolio assessment should be recorded and signed and

1 Standard referenced frameworks requires the development and use of scoring rubrics that are expressed in the form of ordered, transparent descriptions of quality performance that are specific to the Unit of Competency, underpinned by a theory of learning and are hierarchical and sequential.

[2] Behaviourally Anchored Rating Scales (BARS) are constructed by identifying examples of the types of activities or behaviour typically performed by individuals with varying levels of expertise. Each behaviour/activity is then ordered in terms of increasing proficiency and linked to a point on a rating scale, with typically no more than five points on the scale. Each behaviourally anchored rating scale can be treated as a separate item on the Observation Form in which each item requires the observer to select the statement that best describes the performance of the candidate's application of skills and knowledge in the workplace.

Component		
		dated by the assessor and the comment section should indicate where there are any gaps or further evidence required.
Range and conditions	Outlines any restriction or specific conditions for the assessment such as the location, time restrictions, assessor qualifications etc.	It should be explained to candidates (preferably in written format prior to the preparation of the portfolio) that the portfolio should not be just an overall collection of candidate's work, past assessment outcomes, checklists and of the information commonly kept in candidates' cumulative folders. The candidate should be instructed to include samples of work only if they take on new meaning within the context of other entries. Consideration of evidence across time and varying contexts should be emphasised to the candidate. Candidate should also be instructed to only include recent evidence (preferable less than five years) although more dated evidence can be used but should be defended for inclusion. Such information should be provided in written format to the candidate prior to preparing the portfolio.
Materials/resources required	Describes access to materials, equipment etc that may be required.	Materials to be provided to the candidate to assist with preparing his/her portfolio may include access to: photocopier, personal human resource files etc., if required.
Assessor intervention	Defines the amount (if any) of support provided.	Clarification of portfolio requirements permitted.
Reasonable adjustments	Guidelines for making reasonable adjustments to the way in which evidence of performance is gathered without altering the expected performance standards	An electronic and/or product based version of the portfolio may be prepared by the candidate. The portfolio may include videos, photographs etc.
Validity	Evidence of validity to support the use of the assessment evidence for the defined purpose and target group of the tool.	Evidence of the validity of the portfolio tool may include: • Detailed mapping of the evidence used to judged the portfolio with the Unit(s) of Competency (content validity); • Inclusion of documents produced within the workplace and/or has direct application to the workplace (face validity); • Evidence that the tool was panelled with subject matter experts (face and content validity); • The tool clearly specifying the purpose of the tool, the target population, the evidence to be collected, decision making rules, reporting requirements, as well as the boundaries and limitations of the tool (consequential validity); and • Evidence of how the literacy and numeracy requirements of the unit(s) of competency have been adhered to (construct validity).
Reliability	Evidence of the reliability of the tool should be included.	Evidence of the reliability of the portfolio tool may include: • Detailed scoring and/or evidence criteria for the content to be judged within the portfolio (inter-rater reliability); and • Recording sheet to record judgements in a consistent and methodical manner (intra-rater reliability).
Recording requirements	The type of information that needs to be recorded and how it is to be recorded and stored, including duration.	The following information should be recorded and maintained: • The Portfolio tool (for validation and/or moderation purposes); • Samples of candidate portfolios of varying levels of quality (for moderation purposes); and • Summary Results of each candidate performance on the portfolio as well as recommendations for future assessment and/or training etc in accordance with the organisation's record keeping policy. The outcomes of moderation and validation meetings should also be recorded in accordance with the organisation's requirements. The overall assessment result should be recorded electronically on the organisation's candidate record keeping management system.
Reporting requirements	For each key stakeholder, the reporting requirements should be specified and linked to the purpose of the assessment.	• *Candidate*: Overall decision and recommendations for any future training. Progress toward qualification and/or grades/competencies achieved; • *Trainer*: Recommendations for future training requirements; and • *Workplace Supervisor*: Assessment results and competencies achieved.

2.2 OBSERVATION METHODS

(e.g. Workplace Observation, Simulation Exercise, Third Party Report)

Table 3: Workplace Observation: Exemplar Assessment Tool Features.

The context	The purpose and target group should be described	The target group is XXX candidates undertaking the Certificate of XXX. The candidate should be able to demonstrate evidence within the boundaries of their workplace context. Evidence can be collected either on and/or off the job. The tool has been designed to be used to assess candidate's competency acquisition following training (e.g. summative) and/or may be used to demonstrate recognition of current competency. This tool assists with assessing the candidate's ability to apply skills and knowledge and will need to be assessed in conjunction with an interview to ensure adequate coverage of the entire Unit of Competency.
Competency Mapping	Map key components of task to the Units(s) of Competency (content validity) – refer to Template A.2	Evidence criteria needs to be established to judge the quality of the observed performance. Each evidence criterion could be presented as a separate item on an Observation Form. Each item on the Observation Form (i.e. the form to be used to record observations made by the assessor) should be mapped to the relevant sections within the Unit of Competency. This will help to determine the sufficiency of the evidence to be collected and determine whether any other aspects of the Unit(s) of competency need to be collected elsewhere.
Information to candidate	Outline the task to be provided to the candidate that will provide the opportunity for the candidate to demonstrate the competency. It should prompt them to say, do, write or create something.	This may be part of a real or simulated workplace activity. Prior to the assessment event, the candidate must be informed that they will be assessed against the Observation Form and should be provided with a copy of the Form. Details about the conditions of the assessment should also be communicated to the candidate as part of these instructions (e.g. announced versus unannounced observations, period of observation).
Evidence from candidate	Provides information on the evidence to be produced by the candidate in response to the task.	Observations of the candidate performing a series of tasks and activities as defined by the information provided to the candidate. The performance may be: • Part of his/her normal workplace activities; • A result of a structured activity set by the observer in the workplace setting; and • A result of a simulated activity set by the assessor/observer. This section should outline what evidence of performance the assessor should be looking for during the observation of the candidate. The evidence required should be documented and presented in an Observation Form.
Decision making rules	The rules to be used to: • check evidence quality (i.e, the rules of evidence); • judge how well the candidate performed according to the standard expected (i.e. the evidence criteria); and • synthesise evidence from multiple sources to make an overall judgement.	To enhance the inter-rater reliability of the observation (i.e. increasing the likelihood that another assessor would make the same judgement, based upon the same evidence, as the assessor), an Observation Form should be developed and used to judge and record candidate observations. The observer should record the assessors (or observers) observations of the candidate's performance directly onto the Observation Form. The observer should be instructed as to whether to record his/her observations on the Observation Form during and/or after the observation. The Observation Form may have a series of items in which each key component within the Unit of Competency is represented by a number of items. Each item would be the evidence criteria. Each item may be presented as: • a Behaviourally Anchored Rating Scale (BARS; • standard referenced frameworks (or profiles); • a checklist; and/or • open ended statements to record impressions/notes made by the observer. Instructions on how to make an overall judgement of the

Component		
		competence of the candidate would need to be documented (e.g. do all items have to be observed by the assessor?). The form should also provide the opportunity for the observer to record that s/he has not had the opportunity to observe the candidate applying these skills and knowledge. Again, instructions on how to treat not observed items on the checklist would need to be included within the tool. The form should also be designed to record the number of instances and/or period of observation (this will help determine the level of sufficiency of the evidence to be collected), as well as the signature of the observer; and the date of observation(s) (to authenticate the evidence and to determine the level of currency).
Range and conditions	Outlines any restriction or specific conditions for the assessment such as the location, time restrictions, assessor qualifications etc.	Assessors need to provide the necessary materials to the candidate, as well as explain or clarify any concerns/questions. The period of observation should be communicated to the observer and candidate and this would need to be negotiated and agreed to by workplace colleagues, to minimise interruptions to the everyday activities and functions of the workplace environment.
Materials/ resources required	Describes access to materials, equipment etc that may be required to perform the task.	The tool should also specify the materials required to record the candidate's performance. For example: • A copy of the Unit(s) of Competency; • The Observation Form; • Pencil/paper; and • Video camera. In addition, any specific equipment required by the candidate to perform the demonstration and/or simulation should be specified.
Assessor intervention	Defines the amount (if any) of support provided	In cases where observations are to be made by an internal staff member and are to be unannounced, the candidate needs to be warned that s/he will be observed over a period of time for purposes of formal assessment against the Unit(s) of Competency. If the observer is external to the workplace (e.g. teacher or trainer), s/he will need to ensure that the time and date of the visit to the candidate's workplace is confirmed and agreed to by the candidate and the workplace manager. The external observer will need to inform the candidate and his/her immediate supervisor of his/her presence on the worksite as soon as possible. At all times, the external observer will need to avoid hindering the activities of the workplace.
Reasonable adjustments	Guidelines for making reasonable adjustments to the way in which evidence of performance is gathered without altering the expected performance standards	If the candidate does not have access to the workplace, then suitable examples of simulated activities may be used. This section would outline any requirements and/or conditions for the simulated activity.
Validity	Evidence of validity should be included to support the use of the assessment evidence for the defined purpose and target group of the tool.	Evidence of the validity of the observation tool may include: • Detailed mapping of the Observation Form with the Unit(s) of Competency (content validity); • Direct relevance and/or use within a workplace setting (face validity); • A report of the outcomes of the panelling exercise with subject matter experts (face and content validity); • Observing a variety of performance over time (predictive validity); • The tool clearly specifying the purpose of the tool, the target population, the evidence to be collected, decision making rules, reporting requirements as well as the boundaries and limitations of the tool (consequential validity); and • Evidence of how the literacy and numeracy requirements of the Unit(s) of Competency have been adhered to (construct validity).
Reliability	Evidence of the reliability of the tool should be included.	Evidence of the reliability of the observation tool may include: • Detailed evidence criteria for each aspect of performance to be observed (inter-rater reliability); and • Recording sheet to record observations in a timely manner (intra-rater reliability).

Component		
Recording requirements	The type of information that needs to be recorded and how it is to be recorded and stored, including duration.	The following information should be recorded and maintained: • The Observation Form (for validation and/or moderation purposes); • Samples of completed forms of varying levels of quality (for moderation purposes); • Summary Results of each candidate performance on the Observation Forms as well as recommendations for future assessment and/or training etc in accordance with the organisation's record keeping policy; and • The outcomes of validation and moderation meetings should also be recorded in accordance with the organisation's requirements. The overall assessment result should be recorded electronically on the organisation's candidate record keeping management system.
Reporting requirements	For each key stakeholder, the reporting requirements should be specified and linked to the purpose of the assessment.	• *Candidate*: Overall decision and recommendations for any future training. Progress toward qualification and/or grades/competencies achieved; • *Trainer*: Recommendations for future training requirements; and • *Workplace Supervisor*: Assessment results and competencies achieved.

2.3 PRODUCT BASED METHODS

(e.g. Reports, Displays, Work Samples.)

Table 4: Product Based Methods: Exemplar Assessment Tool Features.

The context	The purpose and target group should be described	The target group is XXX candidates undertaking the Certificate of XXX. The candidate should be able to demonstrate evidence within the boundaries of their workplace context. Evidence can be collected either on and/or off the job. The tool has been designed to be used to assess candidate's competency acquisition following training (e.g. summative) and/or may be used to demonstrate recognition of current competency. This tool assists with assessing the candidate's ability to apply skills and knowledge and will need to be assessed in conjunction with XXX (e.g. interview) to ensure adequate coverage of the entire unit of competency.
Competency Mapping	Map key components of task to the Units(s) of Competency (content validity) – refer to Template A.2	Each key component of the activity should be mapped to the relevant sections within the Unit of Competency. For example, if the task is to produce a policy document, each key feature to be included in the policy document should be mapped to the Unit of Competency. This will help to determine the sufficiency of the evidence to be collected and determine whether any other aspects of the Unit(s) of Competency need to be collected elsewhere.
Information to candidate	Outline the task to be provided to the candidate that will provide the opportunity for the candidate to demonstrate the competency. It should prompt them to say, do, write or create something.	The instructions for building/creating the product need to be clearly specified and preferably provided to the candidate in writing prior to formal assessment. The evidence criteria to be applied to the product should also be clearly specified and communicated (preferably in writing) to the candidate prior to the commencement of the formal assessment. Details about the conditions of the assessment should also be communicated to the candidate as part of these instructions (e.g. access to equipment/resources, time restrictions, due date etc)
Evidence from candidate	Provides information on the evidence to be produced by the candidate in response to the task.	The tool needs to specify whether the product only will be assessed, or whether it will also include the process. If it is product based assessment only, then the candidate needs to be instructed on what to include in the product. The conditions for producing the product should be clearly specified in the 'information to be provided to the candidate'; which will directly influence the type of response to be produced by the candidate (e.g. whether they are to draw a design, produce a written policy document, build a roof etc). If the Tool also incorporates assessing the process of building the product, then the observations of the process would need to be also judged and recorded (refer to the *Tool Characteristic – Observation Methods* for guidance). In relation to product based assessment only, the candidate would need to be instructed on how to present his/her product for example: • Portfolio (possibly containing written documents, photos, videos etc); • Display or exhibition of work; and • Written document etc.
Decision making rules	The rules to be used to: • check evidence quality (i.e. the rules of evidence); • judge how well the candidate performed according to the standard expected (i.e. the evidence criteria); and • synthesise evidence from multiple sources to make an overall judgement.	This section should outline the procedures for checking the appropriateness and trustworthiness of the product evidence such as its: • *Currency* - is the product relatively recent. The rules for determining currency would need to be specified here (e.g. less than five years); • *Authenticity* - is there evidence included within the product that verifies that the product has been produced by the candidate and/or if part of a team contribution, what aspects were specific to the candidate (e.g. testimonial statements from colleagues); and • *Sufficiency* - is there enough evidence to demonstrate to the assessor competence against the entire Unit of Competency, including the critical aspects of evidence described in the Evidence Guide (e.g. evidence of consistency of performance

Component	Feature	Generic application
		across time and contexts). To enhance the inter-rater reliability of the assessment of the product, the criteria to be used to judge the quality of the product should be developed. Such criteria (referred hereon as evidence criteria) should be displayed in a Product Form to be completed by the assessor. The assessor should record his/her judgements of the product directly onto a Product Form. There are many different ways in which the form could be designed. For example, the form may have broken down the task into key components to be performed by the candidate to produce the product. Each key component may be assessed individually using analytical rubrics (also referred to as behaviourally anchored rating scales (BARS)), or the product overall may be compared to a holistic rubric describing varying levels of performance (also referred to as standard referenced frameworks or profiles) or it simply may be judged using a checklist approach. The candidate should be provided with the evidence criteria prior to commencing building his/her product.
Range and conditions	Outlines any restriction or specific conditions for the assessment such as the location, time restrictions, assessor qualifications etc.	Assessors need to provide the necessary materials to the candidate, as well as explain or clarify any concerns/questions. The time allowed to build the product should be communicated to the candidate and if there are any restrictions on where and when the product can be developed, this would also need to be clearly specified to the candidate.
Materials/ resources required	Describes access to materials, equipment etc that may be required to perform the task.	The tool should also specify the materials required by the candidate to build the product. It should also specify the materials required for the assessor to complete the form. For example: • A copy of the Unit(s) of Competency; • The Product Form; • Pencil/paper; and • Specific technical manuals/workplace documents etc.
Assessor intervention	Defines the amount (if any) of support provided.	The amount of support permitted by the assessor, workplace supervisor and/or trainers needs to be clearly documented.
Reasonable adjustments	Guidelines for making reasonable adjustments to the way in which evidence of performance is gathered without altering the expected performance standards (as outlined in the decision making rules).	If the creation of the product requires access to the workplace, then suitable examples of simulated activities may be used to produce the product if the candidate does not have access to the workplace.
Validity	Evidence of validity should be included to support the use of the assessment evidence for the defined purpose and target group of the tool.	Evidence of the validity of the product tool may include: • Detailed mapping of the key components within the task with the Unit(s) of Competency (content validity); • Direct relevance and application to the workplace (face validity); • A report of the outcomes of the panelling exercise with subject matter experts (face and content validity); • The tool clearly specifying the purpose of the tool, the target population, the evidence to be collected, decision making rules, reporting requirements as well as the boundaries and limitations of the tool (consequential validity); and • Evidence of how the literacy and numeracy requirements of the Unit(s) of Competency have been adhered to (construct validity).
Reliability	Evidence of reliability of the tool should be included.	Evidence of the reliability of the observation tool may include: • Detailed evidence criteria for each aspect of the product to be judged (inter-rater reliability); and • Recording sheet to record judgements in a consistent and methodical manner (intra-rater reliability).

Component		
Recording requirements	The type of information that needs to be recorded and how it is to be recorded and stored, including duration.	The following information should be recorded and maintained: • The Product Form (for validation and/or moderation purposes); • Samples of completed forms of varying levels of quality (for moderation purposes); and • Summary Results of each candidate performance on the Product Forms as well as recommendations for future assessment and/or training etc in accordance with the organisation's record keeping policy. The outcomes of validation and moderation meetings should also be recorded in accordance with the organisation's requirements. The overall assessment result should be recorded electronically on the organisation's candidate record keeping management system.
Reporting requirements	For each key stakeholder, the reporting requirements should be specified and linked to the purpose of the assessment.	• *Candidate*: Overall decision and recommendations for any future training. Progress toward qualification and/or grades/competencies achieved; • *Trainer*: Recommendations for future training requirements; and • *Workplace Supervisor*: Assessment results and competencies achieved.

2.4 INTERVIEW METHODS

(e.g. structured, semi-structured, unstructured interviews)

Table 5: Interview: Exemplar Assessment Tool Features.

The context	The purpose and target group should be described	The target group is XXX candidates undertaking the Certificate of XXX. This tool assists with assessing the candidate's knowledge and understanding and will need to be assessed in conjunction with XXX (e.g. an observation of performance and/or portfolio) to ensure adequate coverage of the entire Unit of Competency (i.e. sufficiency of evidence).
Competency Mapping	Map key components of task to the Units(s) of Competency (content validity) – refer to Template A.2	Each question within the interview schedule should be mapped to the relevant sections within the Unit of Competency. This will help to determine the sufficiency of the evidence to be collected and determine whether any other aspects of the Unit(s) of competency need to be collected elsewhere.
Information to candidate	Outline the task to be provided to the candidate that will provide the opportunity for the candidate to demonstrate the competency. It should prompt them to say, do, write or create something.	The interview schedule may be structured, semi-structured and unstructured. If using structured and/or semi-structured interview techniques, each question to be asked in the interview session should be listed and presented within the interview schedule. The type of questions that could be asked may include open ended; diagnostic; information seeking; challenge; action; prioritization, prediction; hypothetical; extension; and/or generalisation questions. When designing the interview schedule, the assessor will need to decide whether to: • Provide the candidate with the range of questions prior to the assessment period; • Provide the candidate with written copies of the questions during the interview; • Allow prompting; • Place restrictions on the number of attempts; • Allow access to materials etc throughout the interview period; and • Allow the candidate to select his/her preferred response format (e.g. oral versus written).
Evidence from candidate	Provides information on the evidence to be produced by the candidate in response to each question.	Instructions on how the candidate is expected to respond to each question (e.g. oral, written etc). This section should also outline how responses will be recorded (e.g. audio taped, written summaries by interviewer etc).
Decision making rules	The rules to be used to: • check evidence quality (i.e. the rules of evidence); • judge how well the candidate performed according to the standard expected (i.e. the evidence criteria); and • synthesise evidence from multiple sources to make an overall judgement.	Procedures for judging the quality and acceptability of the responses. For each question, the rubric may outline: • Typical, acceptable and/or model responses; and • BARS that describe typical responses of increasing cognitive sophistication that are linked to separate points on a rating scale (usually 3 to 4 points). The tool should outline the administration procedures for asking each question. For example, not all questions may need to be asked if they are purely an indication of what may be asked. In such circumstances, the schedule should specify whether an assessors needs to ask a certain number of questions per category (as determined in the mapping exercise (see competency targets). The tool should also provide guidelines to the assessor on how to combine the evidence against the interview with other forms of evidence to make an overall judgement of competence (to ensure sufficiency of evidence). As the interview is to be administered by the assessor and conducted in present time, there will be evidence of both currency and authenticity of the evidence. However, if the candidate within the interview refers to past activities etc that s/he has undertaken as evidence of competence, then decision making rules need to be established to check the currency and authenticity of such claims.
Range and conditions	Outlines any restriction or specific conditions for the assessment such as the location, time restrictions, assessor qualifications etc.	The tool should also specify any restrictions on the number of attempts to answer the interview questions and/or time restrictions (if applicable).

Component	Feature	Generic application
Materials/ resources required	Describes access to materials, equipment etc that may be required to perform the task.	The interview schedule should specify the type of materials provided to the candidate which may include: • Written copies of the questions prior to or during the assessment; • Access to materials (e.g. reference materials, policy documents, workplace documents) during the interview to refer to (see the Range of Variables for the specific Unit of Competency); and • Access to an interpreter/translator if the candidate is from a non English speaking background (NESB). The interview schedule should also specify the materials required by the interviewer to record the candidate's responses. For example, paper, pencil, video camera, audio tape etc.
Assessor intervention	Defines the amount (if any) of support provided.	The tool should specify the extent to which the assessor may assist the candidate to understand the questions.
Reasonable adjustments	This section would describe guidelines for making reasonable adjustments to the way in which evidence of performance is gathered without altering the expected performance standards (as outlined in the decision making rules).	Candidates may be given the option of responding to the interview questions in writing, as opposed to oral response. Access to an interpreter during the interview may also be permitted if the competency is not directly related to oral communication skills in English. Similarly, candidates from NESB may be provided with a copy of the interview schedule in their native language prior to the interview.
Validity	Evidence of validity included to support the use of the assessment tool for similar purposes and target groups.	Evidence of the validity of the interview tool may include: • Detailed mapping of the questions to be included within the interview schedule with the Unit(s) of Competency (content validity); • Direct relevance to the workplace setting (face validity); • Evidence of panelling the questions with industry representatives during the tool development phase (face validity); • The tool clearly specifying the purpose of the tool, the target population, the evidence to be collected, decision making rules, reporting requirements, as well as the boundaries and limitations of the tool (consequential validity); and • Evidence of how the literacy and numeracy requirements of the unit(s) of competency (construct validity) have been adhered to.
Reliability	Evidence of the reliability of the tool should be included.	Evidence of the reliability of the interview tool may include: • Detailed scoring and/or evidence criteria for each key question within the interview schedule (inter-rater reliability); • Recording sheet to record responses in a timely, consistent and methodical manner (intra-rater reliability); and • Audio taping responses and having them double marked blindly by another assessor (i.e. where each assessor is not privy to the judgements made by the other assessor) during the development phase of the tool (inter-rater reliability).
Recording requirements	The type of information that needs to be recorded and how it is to be recorded and stored, including duration.	The following information should be recorded and maintained: • The Interview Schedule (for validation and/or moderation purposes); • Samples of candidate responses to each item as well examples of varying levels of responses (for moderation purposes); and • Summary Results of each candidate performance on the interview as well as recommendations for future assessment and/or training etc in accordance with the organisation's record keeping policy. The outcomes of validation and moderation meetings should also be recorded in accordance with the organisation's requirements. The overall assessment result should be recorded electronically on the organisation's candidate record keeping management system.
Reporting requirements	For each key stakeholder, the reporting requirements should be specified and linked to the purpose of the assessment.	• *Candidate*: Overall decision and recommendations for any future training. Progress toward qualification and/or grades/competencies achieved; • *Trainer*: Recommendations for future training requirements; and • *Workplace Supervisor*: Assessment results and competencies achieved

3. Quality Checks

There are several checks that could be undertaken (as part of the quality assurance procedures of the organisation) prior to implementing a new assessment tool. For example, the tool could be:

- Panelled with subject matter experts (e.g. industry representatives and/or other colleagues with subject matter expertise) to examine the tool to ensure that the content of the tool is correct and relevant. The panellist should critique the tool for its:
  - Clarity;
  - Content accuracy;
  - Relevance;
  - Content validity (i.e. match to unit of competency and/or learning outcomes);
  - Avoidance of bias; and/or
  - Appropriateness of language for the target population.

- Panelled with colleagues who are not subject matter experts but have expertise in assessment tool development. Such individuals could review the tool to check that it has:
  - Clear instructions for completion by candidates;
  - Clear instructions for administration by assessors; and
  - Avoidance of bias.

- Piloted with a small number of individuals who have similar characteristics to the target population. Those piloting the tool should be encouraged to think out aloud when responding to the tool. The amount of time required to complete the tool should be recorded and feedback from the participants should be gathered about the clarity of the administration instructions, the appropriateness of its demands (i.e. whether it is too difficult or easy to perform), its perceived relevance to the workplace etc.

- Trialled with a group of individuals who also have similar characteristics to the target population. The trial should be treated as though it is a dress rehearsal for the 'real assessment'. It is important during the trial period that an appropriate sample size is employed and that the sample is representative of the expected levels of ability of the target population. The findings from the trial will help predict whether the tool would:
  - Be cost effective to implement;
  - Be engaging to potential candidates;
  - Produce valid and reliable evidence;
  - Be too difficult and/or too easy for the target population;
  - Possibly disadvantage some individuals;
  - Able to produce sufficient and adequate evidence to address the purpose of the assessment; as well as

- Satisfy the reporting needs of the key stakeholder groups.

This process may need to be repeated if the original conditions under which the assessment tool were developed have been altered such as the:

- Target group;
- Unit(s) of Competency and/or learning outcomes;
- Context (e.g. location, technology);
- Purpose of the assessment;
- Reporting requirements of the key stakeholder groups; and/or
- Legislative/regulatory changes.

A risk assessment will help determine whether it is necessary to undertake all three processes (i.e. panelling, piloting and trialling) for ensuring the quality of the assessment tool prior to use. If there is a high likelihood of unexpected and/or unfortunate consequences of making incorrect assessment judgements (in terms of safety, costs, equity etc), then it may be necessary to undertake all three processes. When the risks have been assessed as minimal, then it may only be necessary to undertake a panelling exercise with one's colleagues who are either subject matter experts and/or assessment experts.

Appendix

A.1 ASSESSMENT TOOL: SELF ASSESSMENT

The following self-assessment is useful for the assessor when reviewing the administration, scoring, recording and reporting components of an assessment tool.

Check to see that the tool has the following information documented to enable another assessor to implement the tool in a consistent manner.

Major component	Type of information
The Context	☐ The purpose of assessment (e.g. formative, summative) ☐ Target group (including a description of any background characteristics that may impact on performance) ☐ Unit(s) of Competency ☐ Selected methods ☐ Intended uses of the outcomes
Competency Mapping	☐ Mapping of key components of task to Unit(s) of Competency (see Template A.2)
Information to candidate	☐ The nature of the task to be performed (how). This component outlines the information to be provided to the candidate which may include: ☐ Standard instructions on what the assessor has to say or do to get the candidate to perform the task in a consistent manner (e.g. a listing of questions to be asked by the assessor). ☐ Required materials and equipment. ☐ Any reasonable adjustments allowed to the standard procedures ☐ Level of assistance permitted (if any) ☐ Ordering of the task(s)
Evidence from candidate	☐ Describe the response format – i.e. how will the candidate respond to the task (e.g. oral response, written response, creating a product and/or performance demonstration)
Decision making rules	☐ Instructions for making Competent/Not Yet Competent decisions (i.e. the evidence criteria) ☐ Scoring rules if grades and/or marks are to be reported (if applicable) ☐ Decision making rules for handling multiple sources of evidence across different methods and/or tasks ☐ Decision making rules for determining authenticity, currency and sufficiency of evidence.
Range and conditions	☐ Location (where) ☐ Time restrictions (when) ☐ Any specific assessor qualifications and/or training required to administer the tool.
Materials/resources required	☐ Resources required by candidate ☐ Resources required by the assessor to administer the tool
Assessor intervention	☐ Type and amount of intervention and/or support permitted
Reasonable adjustments	☐ Justification that the alternative procedures for collecting candidate evidence do not impact on the standard expected by the workplace, as expressed by the relevant Unit(s) of Competency.

Major component	Type of information
Evidence of validity	☐ The assessment tasks are based on or reflect work-based contexts and situations (i.e. face validity) ☐ The tool, as a whole, represents the full-range of skills and knowledge specified within the Unit(s) of Competency (i.e. content validity) ☐ The tool has been designed to assess a variety of evidence over time and contexts (i.e. predictive validity) ☐ The boundaries and limitations of the tool in accordance with the purpose and context for the assessment (i.e. consequential validity) ☐ The tool 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 (i.e. construct validity) ☐ The tool has been designed to adhere to the literacy and numeracy requirements of the Unit(s) of Competency (i.e. construct validity)
Evidence of reliability	☐ There is clear documentation of the required training, experience and/or qualifications of assessors to administer the tool (i.e. inter-rater reliability) ☐ The tool provides model responses and/or examples of performance at varying levels (e.g. competent/not yet competent) to guide assessors in their decision making (i.e. inter and intra-rater reliability) ☐ There is clear instructions on how to synthesis multiple sources of evidence to make overall judgement of performance (i.e. inter-rater reliability) ☐ If marks or grades are to be reported, there are clear procedures for scoring performance (e.g. marking guidelines, scoring rules and/or grading criteria) (i.e. inter-rater reliability)
Recording Requirements	☐ The type of information to be recorded ☐ How it is to be recorded and stored, including duration
Reporting requirements	☐ What will be reported and to whom? ☐ What are the stakes and consequences of the assessment outcomes?
Supplementary information	☐ Any other information that will assist the assessor in administering and judging the performance of the candidate

A.2 ASSESSMENT TOOL: COMPETENCY MAPPING

This form is to be completed by the assessor to demonstrate the content validity of his/her assessment tool. This should be attached to the assessment tool for validation purposes. Note that multiple copies may need to be produced for each task within an assessment tool.

		Component of Unit(s) of Competency			
Step	Component of Task	Elements/Performance Criteria	Required Skill and Knowledge	Range Statements	Evidence Guide
1					
2					
3					
4					
5					
6					
7					

Glossary of Terms

Accuracy of evidence	The extent to which the evidence gathered is free from error. If error is present, the assessor needs to determine whether the amount is tolerable.
Analytical Rubric	An analytical rubric looks at specific aspects of the performance assessment. Each critical aspect of the performance is judged independently and separate judgements are obtained for each aspect in addition to an overall judgement.
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 Guide, 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.
Behaviourally Anchored Rating Scales (BARS)	Behaviourally Anchored Rating Scales (BARS) are similar to rating scales (e.g. 1=Strongly Disagree, 2=Agree, 3=Disagree and 4=Strongly Disagree) but instead of numerical labels, each point on the rating scale has a behavioural description of what that scale point means (e.g 1=the technical terms validity and reliability are stated, 2= strategies for enhancing the content validity and inter-rater reliability have been built into the design of the tool, 3= evidence of how the tool has been designed to satisfy different forms of validity and reliability has been provided etc). They are typically constructed by identifying examples of the types of activities or behaviour typically performed by individuals with varying levels of expertise. Each behaviour/activity is then ordered in terms of increasing proficiency and linked to a point on a rating scale, with typically no more than five points on the scale.
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 performance against industry competency standards and/or learning outcomes.
Concurrent validity	A form of criterion validity which is concerned with comparability and consistency of a candidate's assessment outcomes with other related measures of competency. For example, evidence of high levels of performance on one task should be consistent with high levels of performance on a related task. This is the transfer of learning.
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.
Consequential validity	Concerned with the social and moral implications of the value-laden assumptions that are inherent in the use of a specific task, and its interpretation in a specific, local context.
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.
Criterion referencing	A means of interpreting candidate performance by making comparisons directly against pre-established criteria that have been ordered along a developmental continuum of proficiency.
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 meeting. At the completion of the meeting, 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.
Holistic rubric	A holistic rubric requires the assessor to consider the quality of evidence produced for each competency or learning area. The evidence produced for each competency is balanced to yield a single determination or classification (i.e. competent or not yet competent) of the overall quality of the evidence produced by the candidate.
Internal consistency	A type of reliability which is concerned with how well the items of tasks act together to elicit a consistent type of response, usually on a test.
Inter-rater reliability	A type of reliability which is concerned with determining the consistency of judgement across different assessors using the same assessment task and procedure.
Intra-rater reliability	A type of reliability which is concerned with determining the consistency of assessment judgements made 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 *Guide* 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.
Parallel forms of reliability	A type of reliability which is concerned with determining the equivalence of two alternative forms of a task.
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	Concerned with monitoring, and where necessary, making adjustments to decisions made by assessors prior to the finalisation of assessment results/outcomes. It is referred to as an active approach 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 yet 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 (eg potential danger to clients from an incorrect judgement), equity (eg. outcomes impacting on highly competitive selection procedures), human capacity (eg experience and expertise of assessors) etc.
Rubrics	They are formally defined as scoring guides, consisting of specific pre-established performance indicators, used in judging the quality of candidate work on performance assessments. They tend to be designed using behaviourally anchored rating scales in which each point on the rating scale is accompanied by a description of increasing levels of proficiency along a developmental continuum of competence.
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 validation and/or moderation.
Split half reliability	Type of reliability which is concerned with the internal consistency of a test, where the candidate sits the one test, which is subsequently split into two tests during the scoring process.
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.
Standard Referenced Frameworks	It is a subset of criterion referencing which requires the development and use of scoring rubrics that are expressed in the form of ordered, transparent descriptions of quality performance that are specific to the unit(s) of competency; underpinned by a theory of learning; and are hierarchical and sequential. Subject matter experts unpack the unit(s) of competency to develop the frameworks where levels of performance are defined along a developmental continuum of increasing proficiency and used for interpretative purposes to infer a competency decision. The developmental continuum describes the typical patterns of skills and knowledge displayed by individuals as they progress from novice to expert in a specific area. Along this developmental continuum, a series of cut-points can be made for determining grades (e.g. A, B, C or D etc) as well as the cut-point for making competent/not yet competent decisions.
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.
Thresholds	The cut point between varying levels of achievement. For example, the point in which performance crosses over from a 'competent' performance to a 'not yet competent' performance.
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 tools[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.
Validator	In this *Guide* a validator refers to a member of the validation panel who is responsible for carrying out validation processes. 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 (i.e. predictive and concurrent), construct and consequential. In

[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.

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.