Competency International Ltd

defining the standard

ENSURING CONSISTENCY OF QUALIFICATION OUTCOMES – A DISCUSSION PAPER

JANUARY 2011

Table of Contents

1. Introduction

New Zealand Qualifications Authority (NZQA) has been tasked by the government to undertake a targeted review of qualifications focussed on certificates and diplomas at levels 1 to 6 with the intention of unifying and simplifying New Zealand's qualification system. The review has determined that qualifications will be outcomes-based, allowing for the development of different programmes of study or training leading to the award of a qualification.

One of the challenges in implementing this change is ensuring consistency of assessment judgements of the graduate profile outcomes across a range of delivery and assessment approaches by different providers and workplace assessors.

NZQA has contracted Competency International Ltd (CIL) to provide a practically based discussion paper identifying high level strategies to ensure this consistency occurs.

Samples of combined graduate profiles had already been developed and these have provided input into the discussion paper.   The paper also references relevant international practice and links to existing moderation practice have been drawn and where relevant examples of current moderation practice have been included. An analytical workshop with educational specialists was held to discuss these issues and information sought from international experts in this area. Information has also been included from discussions over a number of years with a range of quality assurance and assessment practitioners including ITO quality systems managers, discussions within NZQA unit standard review panels, ITP academic staff, ITP tutors, registered assessors, moderators from a range of industries and PTE tutors and assessors.

2. What will be the requirements for a qualification to be listed on the NZQf?

NZQA is currently determining the requirements for listing a qualification on the NZQf. Whilst the exact details are not yet available four principles will apply.

A qualification will:

◆ Be needs based

◆ Focus on outcomes

◆ Be flexible

◆ Require both trust and accountability

For the purpose of this report the key terms will be defined as follows:

◆ Qualification developer - an organisation accepted by NZQA as a legal entity, that develops qualifications which meet the NZQF listing requirements. Qualification developers are responsible for reviewing the qualifications they develop and for monitoring and reporting on the consistency of achievement

◆ Programme owner – an organisation that develops programmes leading to qualifications on the NZQF that are approved by NZQA

◆ Programme provider – tertiary education providers accredited by NZQA to deliver programmes leading to qualifications on the NZQF

♦ Qualification graduate profile outcomes – outline the capabilities of a graduate as a result of achieving a specified programme of study or training leading to the award of a qualification. Graduate profile outcomes define the minimum level of skills, knowledge, understanding and attributes a graduate awarded a qualification can demonstrate.

3. Good practice in assuring consistency in the assessment of qualification outcomes

3.1   What is consistency in assessment?
Consistency in assessment involves the achievement of comparable outcomes. An assessment process would be considered to deliver consistent outcomes if assessors assessing learners against the same specification of outcomes in different contexts made comparable assessment decisions.

Various jurisdictions call this process by different names. In New Zealand, to date it is called 'moderation'. The NZQA website describes this as a process of ensuring an organisation's assessment activities are fair, valid, and consistent with the required standard across a number of assessors or assessing organisations.

3.2   Why is it important?
The key to the drive to achieve consistency is the requirement for confidence in an assessment process. Any assessment system requires formal assurance processes to verify that its assessment judgements are credible and defensible as stakeholders need to be confident that programme owners' assessment processes deliver comparable outcomes from assessments carried out by different assessors. Similarly, they need to be confident that the outcomes of assessments conducted by assessors within one organisation are comparable with assessments conducted by assessors in other organisations or workplaces.

3.3   What do we know about how it works?
Interestingly there is very little international literature on achieving consistency in assessor decision making and some of it is extrapolated from discussions on assessment rather than moderation/consistency analysis itself.

For some considerable time in the initial stages of the development of assessment standards[1] it was believed that the clear and detailed specification of the standards would automatically lead to consistency of judgements between assessors. Bowen-Clewley and Strachan (1997) note that in New Zealand

> the solution chosen, in the first instance, to hold the standards in the detail of the performance criteria, range statements and special notes has not led to the unequivocal (consistent) interpretation of standards. Clearly, as has been found in other areas, the interaction of professional educators (including the moderation system) and the development of exemplars will be the key to agreement on the interpretation of the standards.

They cite Coogan (1996) providing a further example of this in reviewing implementation in the UK where he contrasts

> widely disparate interpretation of seemingly unambiguous performance criteria with high comparability of judgements in the "woolly" communications standards because tutors formed a clear network…

Greatorex & Shannon (2003) further explored how to standardise assessor judgements in National Vocational Qualification (NVQ) assessment. They found that assessors tended to believe that standardisation was undertaken by ensuring that all assessors followed the same assessment procedure and that such standardisation ensured that assessors made reliable judgements.

---

[1] These are known variously in different jurisdictions

During the standardisation exercises the assessors found that their judgements were not necessarily always consistent, questioning this belief. The limitations of the standardisation exercises were discussed and they concluded that

> *the belief that standardising procedures, paperwork and/or practice will standardise assessment decisions needs to be challenged. It is likely that standardising processes, procedures and practice will make the NVQ system fairer if assessors judge in a similar way, but it will not necessarily ensure that consistent assessment decisions are made.*

In an attempt to explain the different interpretations made by assessors a number of writers focused on the concept of assessors' mental models. Their discussions were often underpinned by Argyris & Schön's (1974) argument that people have mental maps with regard to how to act in situations. This involves the way they plan, implement and review their actions. Furthermore, they assert that it is these maps that guide people's actions rather than the theories they explicitly espouse. They explore this view further in the concept of double loop learning.

Figure 1 Single and Double Loop Learning (Argyris &Schön 1974)
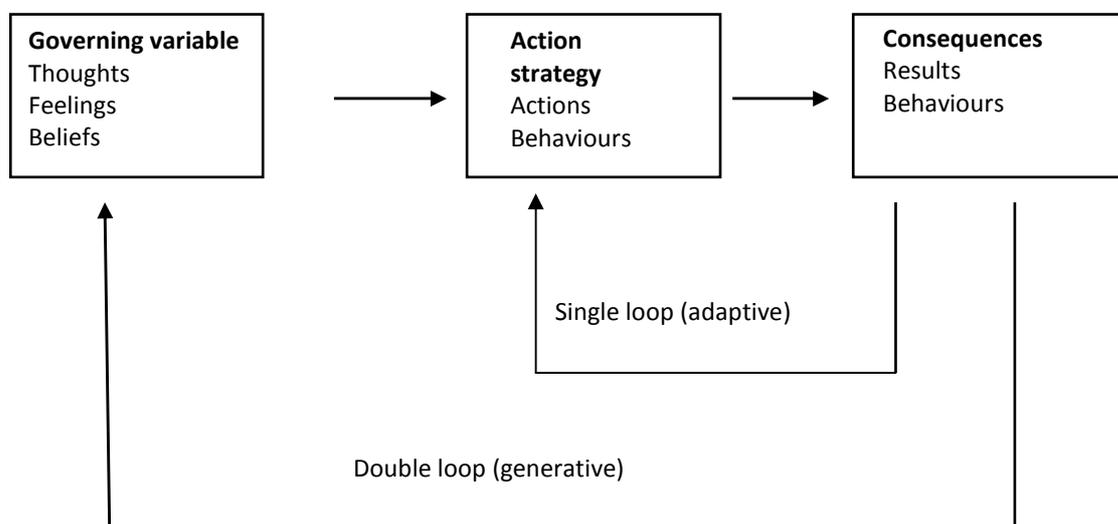
Wolf (1995) built on this research in terms of assessment and explored these differing mental models referring to the need for networks and exemplars as mechanisms to enhance consistency of assessors' judgements. She notes

> *all the research evidence that we have on assessors behaviour, emphasises the very active role their own concepts and interpretations play … They operate in terms of an internalised, holistic set of concepts about what an assessment ought to show, and about how, and how far, they can take account of the context of the performance.*

In an attempt to increase consistency Wolf (1995) suggested the use of networks and exemplars

> *we can reach quite high levels of standardisation with decentralised assessment systems of the type which competence demands. The key requirements are exemplars and networks of assessors – plus a good deal of realism of what can be claimed and achieved*

Kennie & Green (2001) studied the Royal Institution of Chartered Surveyors (RICS) (UK) and assisted them to develop a framework that helped clarify and made explicit the criteria and standards used by assessors when judging what constitutes 'professional competence'.

They too investigated the different 'mental models' used by professionals to define what was meant by 'being professionally competent', and tested these findings with assessors to come to some common understanding of this concept.

They found that no single, uniform definition of the concept existed, and developed (based on previous frameworks) a framework of competence for assessors to use that combines:

♦ Knowledge competence

♦ Cognitive/problem solving competence

♦ Business competence

♦ Ethical/personal behavioural competence

They believed that this framework enabled assessors to easily compare and contrast their judgement of the balance that is appropriate between the elements.

A pilot programme using the framework found the following:

♦ Variability did exist in the minds of assessors and a discussion between assessors about the appropriate balance of competence (across the four components) was of critical importance

♦ The importance of integrating judgements of ethical and personal behavioural competence with those associated with technical competence, rather than assessing them as disconnected activities

♦ The importance of having much clearer processes for judging learners who were of a borderline nature

♦ The need for judgements of technical competence to be more fully integrated with the holistic judgement of professional competence.

As a consequence they recommended a standard judgement format be used focusing on integrating the high level outcomes, regardless of context.

In reviewing how assessors reach judgements Johnson's (2008) research on assessors of a health and social care qualification identified four key values that appeared to inform assessor practice: a sympathetic and contextualised view of the whole learner; respect for supportive and positive relationships; valuing professional trust; and a commitment to "care". These are all far more difficult practices and positions for assessors if they are not also the trainers or teachers because in such cases they would not have relationships with the learners.[2]

Johnson suggests

> … how experts make judgements about learners' performances since this might help to make the assessment system more transparent and able to justify claims of fairness. This is particularly the case for vocationally related holistic assessments where assessors might attend to a variety of factors in different contexts when forming a judgement … the challenge for assessment agencies dealing with large numbers of assessors with differing experiential backgrounds across different contexts is to implement measures that structure, or perhaps calibrate, their judgements so that in the face of uncertainty they are basing their decisions on common comparators.

---

[2] Alternatively it could be argued that these very relationships may lead to poor judgements based on personalities.

Sadler (1996) goes further and suggests rather than just using end point moderation, assessment quality depends on three integrated components:

♦ *well written broadly based standards*

♦ *trained and experienced assessors*

♦ *availability of quality exemplars [3]*

It is apparent then that the training of assessors is vital if they are to actively partake in a community of practice to assure that consistency in qualifications outcomes is achieved. They need to be prepared to stay updated with developments in what counts as knowledge and performance and how knowledge and performance can be evidenced. More broadly assessors need to understand what kinds of things are now being demanded of people in workplaces.

Hase & Saenger's (2004) study showed that many assessors struggled with the idea that assessment should take many forms, involve a range of evidence and be more holistic, because it required a broader view of assessment and greater involvement of the assessors. In other words, the requirements of assessment can get beyond the capabilities and knowledge of the assessors.

What emerges from the discussion above is an interesting, although not always stated distinction between technical competence and a holistic view of competent performance. In New Zealand's current system the focus tends to be on the components of performance which are aggregated to form a qualification where the sum of components is deemed to equal the notion of competence. It may be easier to assess and moderate against technical components and more difficult to assess and moderate the actual holistic or professional outcomes but this is what the new qualification design will require with its graduate profile outcomes statements.

It has become clear that achieving consistency of outcomes starts with understanding by a range of people including assessment designers, assessors and moderators of what is being assessed, what evidence is needed and how the assessment decisions will be made. It must also be recognised that assessors have their own internalised mental models of what achievement looks like. These models appear to be influenced by exemplars and discussions with peers rather than by standardised assessment processes alone. However, while having a shared understanding is critical to the achievement of consistent outcomes, this does not imply that assessment arrangements and processes must be uniform. Learners and settings do vary and assessors must be able to employ assessment approaches that suit the learner and the assessment context.

This has significant implications for the design of consistency strategies for graduate profile outcomes as well as for assessment standards.

4. What are the processes currently being used to maintain consistency both within assessment standards and qualifications?

As indicated above, in New Zealand moderation is the term used to describe the quality management process that is designed to ensure there is uniform interpretation and application of assessment criteria to ensure assessment judgments are consistent both across unit standard assessment and programme outcomes.

---

[3] Personal interview, November 1996

The function of these quality assurance activities will usually be:

- To verify that assessments are fair, valid and consistent

- To identify where there is any need for redesigning assessment activities, or for reassessing applicants

- To adjust interpretations of the assessment criteria for the future

- To provide a mechanism for handling appealed assessment

- To provide feedback on the quality of the outcomes/assessment criteria

- To provide feedback to assessors on the quality of their judgements.

This discussion paper explores a range of ways in which all the five elements can be incorporated into ensuring consistency of graduate profile outcomes.

It needs to be noted that there is considerably more information provided on moderation of standards registered on the NQF than on moderation of qualification and programme outcomes. This is because requirements are specified at a national level for the standards but qualification and programme outcomes are determined in accordance with their programme owners' and their institutions' quality management processes.

4.1   *Moderation of provider qualifications and programmes*
There are a number of programmes that are offered by Tertiary Education Organisations (TEOs) leading to qualifications. A range of the following processes are currently being used to ensure consistency of qualification outcomes and component outcomes across a range of providers:

- Use of common assessment tasks

- Use of a major project designed holistically across a number of course outcomes and verified by industry placements

- Capstone assessments[4] assessed by a panel of assessors

- Standardisation meetings for assessors including check marking

- Panel assessment of learner work samples

- Peer review of assessment samples

- Inter-institutional visits to review assessments of specific programmes and/or confirm internal self evaluation processes

- Appointment of a cross institution monitor

---

[4] Capstone assessments draw together a learner's knowledge, skills and attributes into a final summative assessment. They take a number of forms such as projects , observed on job performance or a performance assessment under examination conditions

- Technology generated rating scales that convert qualitative judgements into standardised quantitative measures for subsequent statistical analysis and auditing. These often incorporate the concept of consistency of performance, application or similar constructs and use a variety of referencing systems for the award of marks according to selected patterns of distribution[5]

- Occupational licensing/professional registration that confirms quality of qualification outcomes

These approaches may be used in combination, with some viewed as more rigorous and effective than others. For example the use of common assessment tasks, in itself provides no guarantee of consistency because of the 'mental model' issues described above. Some, for example panel assessment of capstone events, are only practical where smaller numbers of learners are involved and cost becomes a major impediment as numbers of learners grow.

However many of these approaches may be useful in determining the consistency of graduate profile outcomes.

4.2 *Moderation of standards registered on the National Qualifications Framework (NQF)*

Moderation requirements are specified in the assessment and moderation action plans that must be provided as part of the registration of standards on the NQF. The details of these vary between standard setting bodies but always focus on strategies to determine consistency in assessor judgements. These may include:

- Roles and responsibilities for moderation

- Pre-assessment moderation requirements

- Post-assessment moderation requirements

- Sampling requirements including:

  - Moderation visits

  - Non compliance procedures

There is no doubt that methods to ensure consistency of assessor judgements need to be implemented and good practices are occurring particularly where there is a focus on:

- Quality training of assessment designers, assessors and moderators

- Use of effective internal and external moderation systems

- Use of assessor networks to peer review assessment of outcomes in order to reach consistent assessment decisions

- Holistic judgement across integrated assessment and whole task performance.

However there is a view expressed in a range of forums[6] and reported by members of the analytical workshop established to support this discussion paper that the current model has come to focus on compliance rather than quality assurance. Anecdotal reports suggest that there tends to be an over emphasis on compliance in assessing the minutiae of a unit standard and its component parts rather than focusing on assessment at element/outcome level.

---

5 An example of this approach is the Talent Centre's Affirmative Assessment System™ (www.talent-centre.com/Affirmative.aspx)
[6] These include discussions over a number of years with a range of quality assurance and assessment practitioners including ITO quality systems managers, discussions within NZQA unit standard review panels, ITP academic and teaching staff, ITP tutors, registered assessors, moderators from a range of industries, PTE tutors and assessors and discussions within the ITP Quality Board.

This is not a new concern. In 1996 Coutts & McAlpine conducted the first independent research on NQF moderation systems. This focused on the 1995 trials of mathematics and geography. They reported a number of positive aspects and identified areas in which improvement could be made in the school sector as follows:

♦ The need for moderation systems to allow flexibility in delivery and not to constrain the types of assessment used

♦ The need for evidence to show improving consistency and validity as a result of the moderation systems

♦ Difficulties in getting moderators to accept integrated assessment activities, collection of evidence approaches and different formats for assessment task design and submission for moderation

♦ Compliance costs, increased workload and manageability of participating in several different moderation systems where programmes assess standards from a variety of standard setting bodies

♦ The workload for moderators who undertake the position in addition to their full-time employment

♦ The tension between sharing of material and commercial opportunities created by taking part in moderation cluster groups

♦ The need to use new technology to enhance efficiency

♦ Clarification of moderators' role

♦ The need for moderators who are well trained and working with standards themselves.

New Zealand is not alone. In a recent paper Sadler (2007) was critical of this atomised approach that is apparent in a number of countries and notes with concern

> *I believe that a blinkered conceptualisation of curriculum, the strong trend towards fine-grained prescription, atomised assessment, the accumulation of little 'credits' like grains of sand, and intensive coaching towards short-term objectives, are a long call from the production of truly integrated knowledge and skill ... Too often the focus has been on a multitude of discrete competencies, rather than on competence.*

In May 2001, a New Zealand based consortium produced a document called *Best Practice Moderation.* The joint working party comprised representatives for NZQA, Skill New Zealand, Association of Polytechnics in New Zealand and the Industry Training Federation. They recommended the following principles to guide moderation practice:

♦ **Principle 1 Partnership**

   Best practice moderation occurs when it is based on an effective partnership between assessors and between assessors and moderators.

♦ **Principle 2 Continuity**

   Best practice moderation occurs when it is an ongoing process that underpins relationships between assessors and between assessors and moderators.

♦ **Principle 3 Evolution**

   Best practice moderation occurs when it is an evolving and educative process for assessors and moderators

The project also included the provision of exemplars of current good practice and in the report identifying exemplars and reporting on trends around the submissions, Bowen-Clewley (2002) reported the following:

♦ *Emphasis on Performance Criteria*
There continues to be too much emphasis placed on the performance criteria in many of the assessment documents – in particular, the aligning of judgement statements with each criterion. Whilst best practice requires that performance criteria and range statements are considered by the assessor – they are statements of what would constitute a judgement or decision that could result in the award of credit – the decision should be made at element level. Moreover in the best submissions there is provision for evidence to be cross-referenced across a number of elements and their related performance criteria.

♦ *The requirements of level*
It became apparent in evaluating the submissions that care needs to be taken to ensure that the level of a unit standard is taken into consideration when interpreting the requirements of the elements and performance criteria into assessment and evidence guide design. NZQA level descriptors provide a guide to the expectations of the various levels and it is important when determining evidence sufficiency requirements that these be borne in mind. For example, some material submitted requires more evidence for a Level 1 unit standard than another requires at Level 5. This is particularly true of assessment design for generic unit standards, where Principle 3 of the Best Practice Protocols is seemingly ignored. Industry based standards seem to be less prone to this problem. This is probably because assessors in the industry have a common understanding of the meaning and requirements of the unit standards. With generic unit standards there is rarely a shared community of practice to inform interpretation of the requirements, although this should be mitigated by moderation activities.

♦ *The impact of moderation*
It is very clear from discussions with submitting organisations and other feedback, particularly from PTEs, that moderation systems and processes can have a limiting factor on the approach to learning and its assessment adopted by providers. Whilst the consortium has published the Best Practice Principles in Moderation and Assessment, many moderators and moderation cluster groups are still insisting on seeing specific evidence that each performance criterion and range statement has been assessed and this leads to a proliferation of evidence guides that provide this approach. As one provider said "we'd like to use a holistic, element based approach to evidence and evidence guides but it causes so much hassle in moderation that it's not worth the trouble".

It seems unlikely that assessment practice will change substantially until moderators and moderation systems become familiar with and adopt best practice principles in day-to-day moderation operations.

Similar problems identified above were reported in discussions for this paper[7]. It appears little has changed except the level of frustration at 'black box' moderation where there is no discussion with moderators and moderation reports focus on atomised performance criteria and range statements and are often seen to add little value to assessment . As one experienced assessor said 'If people have a stick hanging over them they lose sight of the fact that the moderation process can add value to their work and thus to the learners'.

---

[7] These discussions have been held with a range of quality assurance and assessment practitioners including discussions within a recent NZQA unit standard review panel, ITP academic and teaching staff, registered assessors and moderators from a range of industries.

This is further exacerbated by the minimal training undergone by many assessors and moderators. Initial assessor training ranges from half a day to five days and for moderators it is usually a day. There is no requirement for training in adult education and many assessors just follow prescribed checklists. Some assessors and moderators receive on-going professional development but this is not a normal practice in most contexts and is usually spasmodic.

A Symposium on Workplace Assessment at University[8] of Cambridge in 1999 participants concluded that:

♦ Attention to upfront matters such as training of assessors and assessment designers, pre moderation of assessment tasks and careful accreditation processes will yield better quality assurance than compliance based end point audit

♦ Quality assurance should be designed to squeeze out variance and take place as close to workplace as possible

♦ The best quality assurance processes involve a buy-in from all parties (including the assessor, the quality assurance person, learners, etc.)

♦ Because competence equals the capacity to put things together in the right way for the context, high quality assessment will be sensitive to particular and changing workplace circumstances ('fitness for purpose').

4.3 The role of evaluative quality assurance in determining consistency of graduate profile outcomes

External Evaluation and Review (EER) provides an independent judgement of the educational performance and capability in self-assessment of all TEOs that is focused on the outcomes of tertiary education and the key processes that contribute to these outcomes. A key difference from traditional approaches is that while evaluative quality assurance also looks at processes, it does so from the perspective of the utility or impact of these processes on what is done and achieved - the "valued outcomes" of tertiary education.

Moreover it operates in an environment of 'high trust and high accountability' throughout the tertiary sector. The relationship between government agencies and TEOs is based on good communication and collaboration. Thus through periodic external evaluation and review TEOs are held accountable to their students, employers, funders, quality assurance bodies and other interested parties. The review process also provides information to support improvement across the tertiary education sector. A full summary of this approach can be found on www.nzqa.govt.nz

Its relationship to determining the consistency of graduate profile outcomes lies in the self assessment questions that are of this process. These questions are linked to self assessment and depending on the type of organisation focus on:

♦ How well the organisation understands and meets the needs of industry, learners and government

♦ How well learners achieve

♦ The value of the outcomes for key stakeholders including learners

While this information is not directly related to individual programmes it provides a broad overview of practice within programme providers' institutions and has the potential to be linked to specific programmes if required.

---

[8] From notes of the meeting attended by the author

4.4   Conclusions on this section

Given that the principles underpinning moderation do not appear particularly contentious, it is interesting that it is an area of considerable debate in New Zealand, particularly when referring to the moderation of standards registered on the NQF. This debate has been included above to provide a context in which the new approach to qualification design and the possible ways of ensuring consistency of graduate profile outcomes is situated. It is very important that discussion of current moderation practices provides input into strategies for the future rather than a re-litigation of the pros and cons of the system.

5.   What needs to be addressed to ensure consistency of graduate profile outcomes?

5.1   The key considerations

- ◆ The fundamental principles underpinning the system?

- ◆ The roles of qualification developers, programme owners and providers and NZQA in this process?

- ◆ The methodologies by which consistency of graduate profile outcomes will be determined including:

    - ◆ What evidence will be required to demonstrate consistency of outcomes?

    - ◆ How will evidence be collected and by whom?

    - ◆ How often will outcomes be monitored for consistency?

    - ◆ What records are to be kept of any monitoring?

    - ◆ How will problems identified in consistency of outcomes be dealt with?

In order to operationalise these considerations, the following are suggested questions that could be used in determining the consistency of graduate profile outcomes. It is a version of questions used by a number of moderation systems using a peer review method.

Table 1 – Suggested questions for determining consistency of graduate profile outcomes

• What was the valid evidence produced?
• How was the evidence verified/authenticated?
• Did the evidence ensure that all critical outcomes were assessed and met?
• Was the evidence sufficient to show consistency of graduate profile outcomes across a range of learners in the programme?

• Do you agree with the judgment that the outcomes have been met?
    o If yes, why?
    o If no, why not?

• What advice would you give to improve consistency of graduate profile outcomes?
• Are there any comments that should be considered?

6. Suggested strategies for maintaining consistency of qualification outcomes

6.1 Underpinning principles

At the analytical workshop the following principles were agreed on as needing to underpin the processes designed for maintaining consistency of qualifications' outcomes:

- ◆ Flexibility. Any approach to ensuring consistency of graduate profile outcomes must ensure that a qualification is achievable in different contexts relevant to different learners and stakeholder needs. It must recognise that alternative programmes of training or study allow achievement of the qualification in ways most suited to learners' educational, work or cultural needs and aspirations

- ◆ Decisions on consistency are made on an evidential base

- ◆ The processes for determining consistency of graduate profile outcomes are transparent and provide assurance of a quality product that is valued by all stakeholders and allows for confidence in the NZQf

- ◆ In any process for determining consistency of graduate profile outcomes the focus is on graduate profile outcomes and at the highest level that provides stakeholder confidence

- ◆ Processes for determining consistency of graduate profile outcomes support strengthening of Māori as a people by enhancing and advancing mātauranga Māori

- ◆ Processes allow for sound judgements about the consistency of graduate profile outcomes to be made, are cost effective and manageable and add value in a partnership between key stakeholders that ensures improvement on an ongoing basis

- ◆ Processes for determining consistency of graduate profile outcomes are connected to other quality assurance processes and do not operate in a standalone environment

- ◆ People operating systems are well trained, competent and confident

6.2 Possible strategies for ensuring consistency of graduate profile outcomes

In reviewing the material in the sections above and from discussions at the analytical workshop, two key concepts emerged that should underpin any method/s for ensuring consistency of graduate profile outcomes:

- ◆ The foundation concept encapsulates the importance of the integration of all the inputs into assuring that consistency. It reflects that the combination of good graduate profile outcomes and assessment design and training of the relevant people, supported by quality assurance systems and determination of the consistency of graduate profile outcomes will lead to the assurance of the consistency of graduate profile outcomes across a range of pathways. It suggests that inputs at the front end of the process will lead to fewer requirements for sampling at the post assessment end. This is graphically represented in Figure 2.

This is supported by

- ◆ The triangulation concept that focuses on programme providers' quality assurance processes and is linked to their quality assured status. It suggests that differential requirements may be used in the depth and frequency of determining consistency of outcomes based on evidence that the programme provider demonstrated capability in a discipline area combined with external evaluations of quality. This may be identified by the qualification developer, NZQA or a combination of both parties in discussion with the programme provider. This is graphically represented in Figure 3.

Figure 2 Foundation concept

QUALIFICATION DEVELOPMENT	QUALIFICATION ASSURANCE	CONSISTENCY REVIEW
Training and credentialling of: • standards and qualifications designers • assessment designers • assessors • consistency reviewers + • on-going professional development • performance review	Internal quality assurance systems of programme owners ↓ EER	Provision of evidence of achievement of outcomes

Good qualification design rules →

Less

More

= Assurance of the consistency of outcomes across a range of pathways

Figure 3 Triangulation concept

QA Status of organisation

HIGH

HIGH

Assessment Capability in discipline area

Moderation/ consistency requirements

QA Status of organisation

LOW

LOW

Assessment Capability in discipline area

Moderation/ consistency requirements
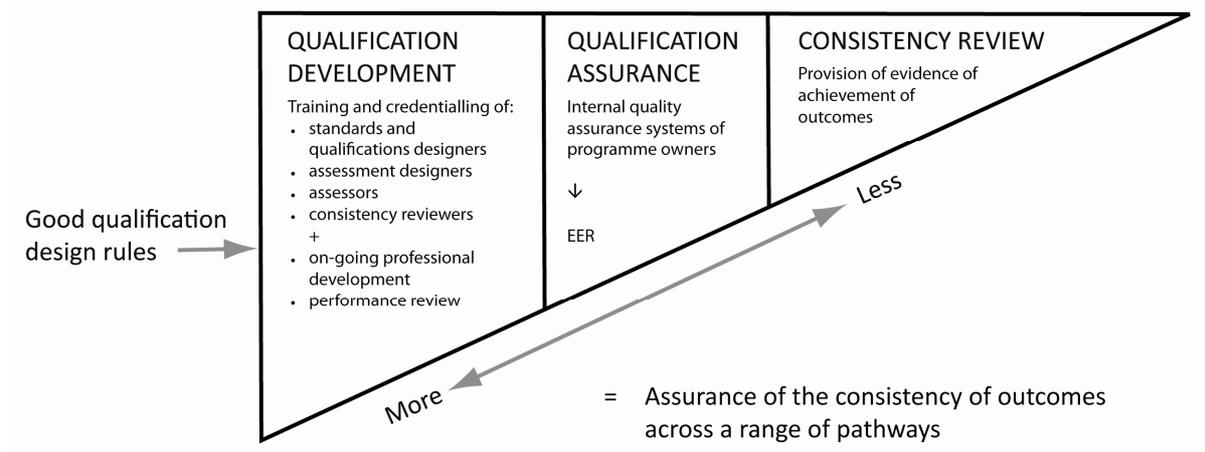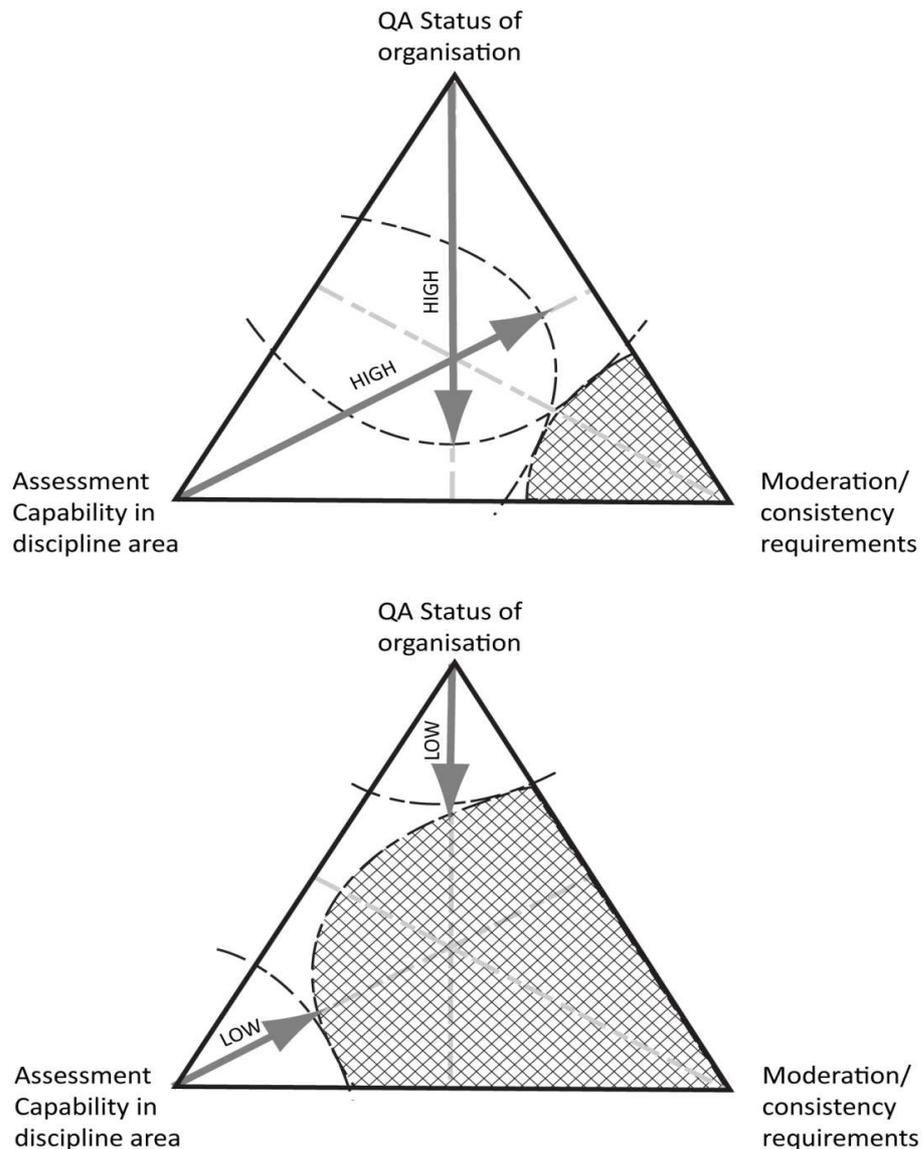
The critical discussion point for this paper lies in how consistency of graduate profile outcomes will be determined. The following are four possible implementation strategies that incorporate the principles embedded in both concept diagrams. These implementation strategies are not mutually exclusive and may be used in conjunction with each other depending on the needs and context in which they are to operate.

All options are provided with a level of confidence rating based on the Likert 5 point scale with 1 = very limited confidence and 5 = high level of confidence. This rating refers to the ability for the method to determine consistency of graduate profile outcomes.[9]

6.2.1 Light sampling method
This method is a directed approach from the qualification developer and/or NZQA. It may include:

♦ The qualification developer identifying sampling requirements of learner work which demonstrate achievement of specified graduate profile outcomes from programme providers. This may be undertaken initially as a baseline requirement for all programme providers and then on a differential basis identified through the track record of programme providers

♦ Requiring a range of samples that reflect outstanding, achieved and marginal learner performance

♦ Evidence of 'next user' feedback may be required. This may be from learners' employers or follow-on providers

♦ Feedback from learners conducted by qualification developers

Advantages and Disadvantages of the model

Advantages	Disadvantages
Programme providers know in advance what will be reviewed	Tends towards a compliance model. It involves the qualification developer taking responsibility for the specification of samples and other evidence
Processes transparent for all stakeholders so increases credibility of the process	Potential for fraudulent or ethically compromised data in order to meet specifications
Allows for specific comparison of generic components across programmes by specification of required outcomes	May limit self reflection and professional growth through review of own performance and feedback from the qualification developer as it takes a directive approach
Can target areas of weakness or gaps against specific outcomes	
Provides for data triangulation that increases defensibility of judgements	

Level of confidence provided by the model   3.5
The level of confidence of this model is directly related to the perceptions of it held by programme providers. Because it tends towards a directive approach, it potentially takes the responsibility from programme providers for assuring consistency of graduate profile outcomes to the qualification developer and/or NZQA.

6.2.2 Programme provider portfolios
This method requires the programme provider to determine the evidence they wish to present that verifies that the learners in their programme have met the graduate profile outcomes.

---

[9] The ratings have been determined by CIL.

This may include the following collected by programme providers:

- Samples of learner performance at a variety of performance or outcomes levels

- Stakeholder and 'next user' feedback

- Destination surveys

- Completion data

- Programme reviews

- Internal quality assurance processes and reports

- Programme provider publications

- Results of component moderation where applicable

- EER reports

- Benchmarking exercises with other programme providers for generic components

- Attestations

- Validation visit reports

- Any other evidence that the programme provider believes supports that their learners have met the graduate profile outcomes eg iwi or community attestation

Advantages and Disadvantages of the model

Advantages	Disadvantages
Programme providers are in control of the process and must be responsible for collecting the evidence that verifies the achievement of graduate profile outcomes	The qualification developer has less control on what is submitted which may lead to workload issues
Provides for reflection and professional growth through review of own performance and allows the programme provider to identify and remediate any problems	May limit specific comparison of generic components across programmes by specification of required outcomes although this could be addressed by the use of a generic outcomes lead reviewer approach as outlined in the next section
Provides for data triangulation that increases defensibility of judgements	Processes may be less transparent for all stakeholders so may decrease credibility of the process
Supports the evaluative quality assurance approach	
Minimises potential for fraudulent or ethically compromised data	
Minimises potential for double compliance issues (component and graduate profile outcomes)	
Much of this information is already collected by programme providers	

Level of confidence provided by the model   4

The level of confidence of this model is directly related to the potential for 'buy in' by programme providers and the fact that much of this evidence is already collected. It links to the 'high trust, high accountability' of evaluative quality assurance and also provides the opportunity for the determination of a level of monitoring by qualification developers/NZQA that reflects the programme provider's on-going performance in maintaining consistency of assessment in a variety of contexts.

6.2.3 Common critical outcome assessment

This method requires the qualification developer, in conjunction with programme providers to identify critical outcomes and to design common assessment activities and/or evidence requirements and related judgement statements for these outcomes. Activities and/or evidence may include:

- ♦ Completion of specified activities and the related evidence for the specified outcomes

- ♦ Stakeholder and 'next user' feedback

- ♦ End product

- ♦ Reflective activities

Advantages and Disadvantages of the model

Advantages	Disadvantages
Programme providers are given clear specification of requirements	The programme provider has less control on how learners are assessed in their programmes
Provides a simple way to collect evidence for judgement of consistency across programme providers to undertake	Unless careful will adopt a one size fits all approach
Provides for data triangulation that increases defensibility of judgements	May tend towards atomisation of evidence for individual outcomes instead of viewing performance holistically
Processes are transparent for all stakeholders - leads to credibility of the process	May revert to assessors being technicians rather than exercising professional judgement
May be useful for specific comparison of generic components across programmes by specification of required outcomes	May constrain learning and learners to a 'normalised' pathway instead of encouraging creativity and innovation

Level of confidence provided by the model   4

Whist this approach has a high level of confidence it has the potential to return to the 'common assessment task' regime that atomised learning and assessment and constrained learners and programme providers from identifying and meeting varying needs and contexts. It could provide a useful approach for the assessment of cross qualification generics but taking these out of context for a programme or pathway could lead away from whole task, holistic assessment. It may also provide a practical model for low level, high uptake qualifications.

6.2.4 'Next user' method

This is a method that relies on feedback from next users of a learner's knowledge and skills. It may include the qualification developer:

- ♦ Using destination data from programme providers to survey next users of the qualification achievers

- Using sector forums and advisory groups to benchmark performance from programme providers in terms of usability of the graduates

- Offering feedback forums for assessors and programme providers

- Collecting formal survey data from learners and next users

- Implementing remedial processes with a programme provider where a problem is identified

Advantages and Disadvantages of the model

Advantages	Disadvantages
Allows programme providers to be responsible for their own quality assurance and only requires the qualification developer/NZQA to deal with identified issues of non-compliance	The qualification developer has no direct evidence on which to make a judgement of programme provider performance
The qualification developer can focus on providing support and information for programme providers and assessors	It relies on a complaints based approach, is retrospective and may miss emerging issues and the opportunity to provide timely feedback
Involves key stakeholders of the qualifications system in providing real feedback on the quality and relevance of holders of the qualifications rather than relying on programme provider perceptions	No ability to compare specific comparison of generic components across programmes by specification of required outcomes.
No double compliance requirements	Lacks transparency for all stakeholders so may compromise credibility of the process

Level of confidence provided by the model   2
This model appeals in its simplicity and its focus on resolving identified problems rather than requiring information from all programme owners. However it is risky in that problems may not be identified at an emerging stage and may have already have disadvantaged learners before they are identified. It also relies on potentially differing views of stakeholders on what constitutes quality outcomes. It is also notoriously difficult to track graduates of lower level qualifications, and difficult to identify whether it is the quality of the provider's programme or the quality of the graduate that determines the graduate's performance in employment. It may be useful in specialised areas where there is a small community of practice and for determining the consistency of assessment judgements around pre-entry qualifications

Whichever model is chosen it is critical to ensure that:

- Qualification and assessment designers are well trained and understand the purpose, meaning and context of their work

- Assessors are qualified in assessment and have experience in undertaking the assessments against mandated outcomes

- Assessors are provided with access to professional development activities and professional networks which facilitate structured discussion and exchanges of ideas about the implementation and interpretation of outcomes

- Formalised processes exist for assuring the quality of assessment tools

- Professional support is provided for assessment designers and assessors which promotes critical self-reflection on practice, assists in identifying strengths and weaknesses, offers up-to-date information on assessment approaches and materials, and provides assistance in dealing with difficult or problematic assessments

- Sound and defensible reporting and tracking processes to monitor assessments exist and are understood by all stakeholders

- Client and industry feedback is regularly sought about the assessment process and credibility of its outcomes.

6.3   Possible ways of implementing these strategies

The ensuing section provides an outline of three possible adjuncts to implementing the strategies outlined in 6.2. Once again these are not mutually exclusive and a combination of them may be used to meet contextual requirements.

6.3.1   Assessment panels

The assessment panel may include industry representatives and regulators in addition to assessors and administrators drawn from across organisations by the qualification developer

The focus of the assessment panel will be on ensuring the consistency of outcomes across a range of programme providers and pathways.

An assessment panel may be constituted from existing bodies that have been drawn together for another purpose but which have a keen interest and experience in the assessment process, such as:

- subject/industry groups
- assessor network groups
- standing industry association committees.

Assessment panels may also be developed specifically to support the assessment process. The membership of the assessment panel is critical and its value will rest on the bringing together of a mix of appropriately skilled and experienced people.

The assessment panel may be empowered to:

- Critique and authorise assessment tools and approaches

- Conduct an expert review of consistency requesting samples from programme providers

- Other activities as determined by the qualification developer to ensure consistency of qualification outcomes

The key issues associated with the implementation of this approach are:

- Having clear terms of reference

- Securing appropriate membership

- Accessing adequate resources including time and money

- Ensuring the work of the panel is purposeful and focused on ensuring consistency in assessment of graduate profile outcomes process

- Protection of confidential materials and intellectual property

- Ensuring the commitment of qualification developer's staff to facilitating this process.

6.3.2   Lead assessor/reviewer

A lead assessor is appointed by the qualification developer/NZQA to work with programme providers to develop a consistent approach to assessment. The lead assessor has responsibility for developing, managing and monitoring strategies that ensure consistency in assessment of graduate profile outcomes. The lead assessor will have a position with the qualification developer with professional responsibility for reviewing consistency in the assessment process and outcomes but without direct supervisory responsibility for the assessors.

The lead assessor may:

- Focus on a qualification or a group of qualifications

   and/or

- Focus on a set of generic standards that run across a range of qualifications

The key issues associated with the implementation of this approach are:

- Ensuring that this is a clear definition of the role and functions of the lead assessor

- Securing an appropriately skilled and qualified person to undertake the role

- Accessing appropriate resources including time and money

- Establishing appropriate administrative structures to support the model

- Ensuring the commitment of qualification developer's staff to facilitating this process.

6.3.3   Peer review meetings

Peer review is a process which involves assessors, assessment designers and programme providers in discussing and reaching agreement about assessment processes and outcomes in a particular industry or industry sector. In so doing, assessors develop a shared understanding of the requirements of specific graduate profile outcomes including the nature of evidence, how evidence is collected and the basis on which assessment decisions are made. They should also:

- Foster consistency in assessment processes and outcomes either within or across programme providers operating in a specific industry or industry sector

- Facilitate assessor professional development through providing opportunities for assessors to share information, resources and views on assessment practices within a particular industry or industry sector

- Facilitate the involvement of key stakeholders, such as enterprises, industry organisations, unions and employer associations, in the assessment process

♦ Promote a shared understanding of the basis on which assessment decisions are made.

Peer review meetings will be convened by the qualification's developer

The key issues associated with the implementation of this approach are:

♦ The cost and funding of the peer review process

♦ Ensuring that peer review meetings are purposeful and focused on ensuring consistency in assessment of graduate profile outcomes process

♦ Induction of assessors into the peer review process

♦ Management of the peer review process, including overall coordination, identification of people to lead the peer review process, establishment of goals and evaluation of the effectiveness of the peer review process

♦ Responsibility for coordination of the meetings

♦ Appropriate administrative structures to support the process

♦ Protection of confidential materials and intellectual property

♦ Maintaining assessor involvement in the peer review process

♦ Gaining stakeholder support for the process, including enterprise, industry and programme providers

7. Conclusions

There are significant challenges ahead in implementing the new way of designing NZQf qualifications which allows multiple programmes and pathways towards their achievement and assuring the consistency of their graduate profile outcomes across these programmes and pathways. However this is a goal worth pursuing and one that will significantly benefit learners in assuring the parity of esteem of national qualifications.

There are a number of strategies that will assist in providing this assurance for all stakeholders. These include:

♦ Operating from a principles based approach that recognises the need to be cautious about focusing on the easy to compare rather than the important

♦ Clarifying the extent of the role that NZQA will play in this system. As the country's Quality Assurance Body for these qualifications, NZQA must have a role in monitoring any methods that are adopted and must be confident that consistency of graduate profile outcomes is being achieved across all qualifications and programmes.

♦ Writing qualifications outcomes that are holistic, meet sector and cultural requirements and are not a collection of isolated tasks. There is a danger that the titles or content headings from current registered standards and course prescriptions will simply be collected together to form the graduate profile outcomes and much assessment and component moderation will stay the same

- Resolving the role and function of component moderation within this new system. There are significant dangers in developing a double compliance system – component moderation and review of qualification outcomes consistency. It seems likely that there will still be a requirement for some component moderation particularly while there is assessment of these in standalone contexts. However as the implementation of the new system progresses this should become the responsibility of programme providers through their internal moderation processes with a level of monitoring by qualification developers that reflects the programme provider's on-going performance in maintaining consistency of assessment in a variety of contexts

- Determining how the use of generic unit standards in multiple qualifications will be managed to assure consistency of assessment of these standards across qualifications as well as within them. It may be that the use of Lead Assessors for the generics as described in the Section 6.3 may be a strategy to address this

- Recognising the role that EER can play in providing an overview of an institution's quality assurance practices and the potential to include this information as part of the evidence that can support the determination of consistency of graduate profile outcomes

- Actively facilitating the paradigm shift that will be needed by all stakeholders to looking at qualification outcomes holistically and designing assessment accordingly rather than as a series of independent components

- Requiring quality training and continuing professional development for qualification and assessment designers, assessors and those personnel involved in monitoring of the consistency of assessment. It is time for those roles to be professionalised and credentialing of these key players will enhance what hitherto has been very patchy performance in all these areas.

- Avoiding the temptation to redesign a model that is familiar and thus requires little change management but that has compliance requirements that provide a disincentive for qualification development for those who are responsible for ensuring consistency and overburden those who have to be involved in the process

- Avoiding throwing out the baby with the bathwater. There are good practices happening in the current system. It is important to recognise these and build on them

- Acknowledging that we are dealing with qualifications from Level 1-6 and that one size may not fit all programme providers. There may be a need to use differential systems that reflect cultural needs and the variability of performance of all the players

- Ensuring the system is manageable, usable and credible with a focus on whether there is a higher level at which the comparison for consistency can be done that is easier and with less compliance cost

- Effectively managing the transition between the old and new approaches. The new approach should only be required for new qualifications and qualifications that have had a mandatory review. This should not prevent stakeholders of existing qualifications being early adopters if they wish to move to the new formats

- Valuing and acknowledging the expertise, goodwill and commitment of all the people who have been involved in implementing and operating the NQF to date and who have much to offer in this new environment.

This paper does not pretend to suggest all the solutions to determining consistency of graduate profile outcomes but it is intended that the material provided and the suggested strategies will initiate debate and discussion that will lead to robust and credible processes that enhance the value of qualifications to all stakeholders.

8. Acknowledgements

CIL wishes to acknowledge the contribution to this paper of the following people:

- ◆ Analytical workshop participants and peer reviewers

  - ◆ Jack Doherty   Associate, Competency International Ltd

  - ◆ Janet Calder   Quality Improvement Manager, Te Wananga o Aotearoa

  - ◆ Judith Brown   Executive Director, ITP Quality

  - ◆ Jeremy Baker   Executive Director, ITF

  - ◆ Peter Coolbear   Director, Ako Aotearoa

  - ◆ Shirley Wilson   Workstream Leader, NZQA

  - ◆ Linda Glogau   Manager, Qualifications Development and Tertiary Moderation, NZQA

  - ◆ Anna Williams   Assessment and Moderation Facilitator, Tertiary Assessment and Moderation, NZQA

- ◆ International commentators

  - ◆ Professor Royce Sadler   Griffith University, Australia

  - ◆ Brent Richardson   Independent consultant, Namibia

- ◆ CIL contributors and peer reviewers

  - ◆ Greg Clewley

  - ◆ Jenny Grainger

  - ◆ Melita Farley

  - ◆ Michelle Hoskin

  - ◆ Wendy Rapana

  - ◆ Keith Pyne

  - ◆ Lauri Russell

- ◆ ITP staff

  - ◆ Colin Mcdonnell, Tai Poutini Polytechnic

  - ◆ Denise Holling, CPIT

- ♦ A wide range of other people too numerous to name who have discussed this project with me and shared their views

9. Bibliography

Argyris, C., & Schön, D.A. (1974). *Theories in practice: Increasing professional effectiveness*. San Francisco: Jossey-Bass.

Australian National Training Authority. (2001). *Strategies for ensuring consistency in assessment.* Melbourne: ANTA.

Bowen-Clewley, L., & Strachan, J. (1997). *A report on an assessment stocktake undertaken by the New Zealand Qualifications Authority*. Wellington: New Zealand Qualifications Authority.

Bowen-Clewley, L. (2002). *Best practice in assessment and moderation*. Wellington: Competency International Ltd.

Coogan, P. (1996) *Standards-Based Assessment: Lessons for New Zealand from the United Kingdom* Research Project; Institute of Education, University of London

Coutts, C., & McAlpine, D. (1996). *Monitoring the qualifications framework trials in school: Evaluation of the external moderation system.* Research and Development Centre: Massey University.

Greatorex, J., & Shannon, M. (2003). *How can NVQ assessors' judgements be standardised?* Paper presented at the annual conference of the British Educational Research Association, Heriott-Watt University: Edinburgh.

Hase, S., & Saenger, H. (2004). When quality counts: Implementing competency assessment in the workplace. *International Journal of Training Research, 2* (2), 75-88.

Johnson, M. (2008). Exploring assessor consistency in a health and social care qualification using a sociocultural perspective. *Journal of Vocational Education and Training, 60*(2), 173-187.

Kennie, T., & Green, M. (2001). *Assessing professional competence and enhancing consistency of assessment within the surveying profession*. Paper presented at the annual meeting of the Centre for Education in the Built Environment, Cardiff: Ranmore Retrieved from http://search2.openobjects.com/kbroker/hea/cebe/search/search.lsim?qt=kennie&sr=0&sb=0&nh=10&cs=iso-8859-1&sc=heaportal-cebe&oq=&mt=0&ha=3022.

New Zealand Qualification Authority. (2001). *Best practice moderation.* Wellington: NZQA.

Sadler, D. R. (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education, 14*(3), 387-392.

Wolf, A. (1995). *Competence-based assessment*. Buckingham, Philadelphia: Open University Press.

Background reading

Argyris, C. (1982). *Reasoning, learning and action: Individual and organisational*. San Francisco: Jossey-Bass.

Bierema, L., & Eraut, M. (1994). Workplace focused learning: Perspective on continuing professional education and human resource development. *Developing Human Resources, 6* (1), 52-68.

Bowen-Clewley, E. (2000). Assessing against competency standards in the workplace: A New Zealand perspective. In A. Arguelles & A. Gonczi (Eds.), *Competency based education and training: A world perspective* (pp. 207-225). Balderas, Mexico: Noriega.

Bowen-Clewley, L., & Devereux, C. (1999). *South meets north: Love at first sight: Exploring workplace assessment methodologies that examine whether the marriage, and relationships with in-laws, will survive*. Paper presented at the annual conference of IAEA, Slovenia.

Brittain, S., & Ryder, P. (1999). Get complex. *People Management, 13,* 12-14.

Capper, P. (2000). Understanding competence in complex work contexts. In A. Arguelles & A. Gonczi (Eds.), *Competency based education and training: A world perspective* (pp. 147-172). Balderas, Mexico: Noriega.

Cheetham, G., & Chivers, G. (1998). The reflective (and competent) practitioner: a model of professional competence. *Journal of European Industrial Training, 22* (7), 267-276.

Devereux, C. M. (1997). *Rigour without rigidity: Assessing customer service NVQ/SVQ: A guide for assessors and candidates*. Cheam, Surrey: WA Consultants.

Draves, W. A. (1997). *How to teach adults* (2nd ed.). Manhattan,   KS: Learning Resource Network.

Elley, W. B. (1992). *Is standards-based assessment feasible? – some blunt words about our blunt instruments.* Paper presented at the national workshop on professional development in educational assessment. Ministry of Education: Wellington

Eraut, M., & Cole, G. (1993). *Assessing competence in the professions*. Brighton: University of Sussex.

Eraut, M. (1994). *Developing professional knowledge and competence*. London: Falmer Press.

Eraut, M. (2000). Non formal learning, implicit learning and tacit knowledge in professional work. In Coffield, F. (Ed.), *The necessity of informal learning* (pp. 37-49). Bristol: The Policy Press.

Garrick, J., & Boud, D. (Eds.). (2002). *Understanding learning at work*. London: Routledge.

Gonczi, A. (1993). Integrated approaches to competency based assessment. *Training Agenda, 1*(2), 9-11.

Gonczi, A. (2000). Review of international trends and developments in competency based education and training. In A. Arguelles & A. Gonczi (Eds.), *Competency based education and training: A world perspective* (pp. 15-40). Balderas, Mexico: Noriega.

Graf, A. (2004). Screening and training inter-cultural competencies: Evaluating the impact of national culture on inter-cultural competencies. *Human Resource Management, 15* (6), 10-14.

Greatorex, J. (2005). Assessing the evidence: Different types of NVQ evidence and their impact on reliability and fairness. *Journal of Vocational Education and Training, 57* (2), 149-164.

Hager, P. (1998). Learning in the workplace. In C. Robinson & P. Thomson (Eds.), *Readings in Australian vocational education and training research* (123-145). Adelaide: National Centre for Vocational Education Research.

Hager, P. (2000). Judgement and the law society of NSW specialist accreditation scheme. In A. Gonczi & A. Arguelles (Eds.), *Competency based education and training* (pp. 173-186). Grupo Noriega: Mexico.

Harlen, W. (Ed.). (1994). *Enhancing quality in assessment*. London: Chapman.

Harris, R., Guthrie, H., Hobart, B., & Lundberg, D. (1995). *Competency based education and training: Between a rock and a whirlpool*. Melbourne: MacMillan Education.

Kerka, S. (2001). *Capstone experiences in career and technical education*. Ohio: Eric Publications.

Knowles, M. (1984). *The adult learner: A neglected species*. Houston, TX: Gulf Publishing.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lipman, M. (1991). *Thinking in education*. Cambridge: Cambridge University Press.

Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education, 13,* 191-209.

Schön, D. (1983). *The reflective practitioner*. San Francisco: Jossey-Bass.

Schön, D. (1987). *Educating reflective practitioners*. San Francisco: Jossey-Bass.

Strachan, J. (1995a) Educational Assessment Quality in New Zealand - from Control to Assurance, Second Asia-Pacific Conference on Educational Assessment and Research, Chiang Mai.

Strachan, J. (1995b) Moderation of Assessments in the National Qualifications Framework, IAEA Conference, Montreal.

Strachan, J. (1996) Moderation - Swan or Ugly Duckling, NZARE Conference, Nelson.

Tennant, M. (1997). *Psychology and adult learning* (2nd ed.). London: Routledge.