Guide to Best Practice in

'Direct Entry' Programs , Supplement in Validation & Moderation

2017

Supplement Writers:

Juliana Kendi, Director of Studies, ELICOS Programs, La Trobe Melbourne

Dianne McKeagney, General Manager, VU English, Victoria University

Nick Philippou, Director of Studies - General Programs, Hawthorn-Melbourne

ENGLISH AUSTRALIA

QUALITY | SUPPORT | ASSURANCE

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Contents	Page

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Introduction

Industry context

With the rapid rise in the number of international students coming to Australia from non-English speaking backgrounds over the past few years, concerns have been raised by government and education provider representatives, academics and the media wanting assurance that students undertaking further study in Australia have the necessary English language skills to succeed in their chosen courses.

Along with evidence of their academic ability, international students wishing to further their education in Australia need to provide evidence of their English language ability before they are accepted into a vocational or higher education course. Most do this either by getting the required score in a standardised, government-approved, industry-accepted English proficiency test (currently IELTS, TOEFL, Cambridge English: Advanced (CAE) or PTE Academic), or by successfully undertaking a 'Direct Entry' Program (DEP) at an ELICOS college.

There is a difference between these two options. While an approved test result provides a snapshot of a student's English language proficiency, thus indicating the probability that they will be able to manage the spoken and written language requirements of their chosen course, DEPs aim to prepare a student more comprehensively and holistically for the study environment that they are about to enter by covering not only the different skill areas of reading, writing, speaking and listening, but also 'academic literacy' and the culture of study in Australia. Direct Entry Programs aim to prepare students for their future study by teaching skills to build a smooth transition to the next phase of the international student's journey.

Despite a considerable amount of research having been carried out in this area, there is no conclusive evidence of the reliability of any test of English language proficiency or competency to indicate success in further study, as Arkoudis, Baik & Richardson outline (2012, pp 17-36). However, institutions utilising best practice principles in their DEPs, and incorporating sound practices in validation and moderation of programs and assessment enhance stakeholders' confidence in the efficacy of their programs and the readiness of students for further study.

Aim of the document

This document identifies features that characterise best practice in validation and moderation in Direct Entry Programs in ELICOS and is a supplement to the English Australia *Guide to Best Practice in Direct Entry Programs in ELICOS* (2017). It suggests practical processes and procedures to assist ELICOS providers in developing systems of validation and moderation that enhance the quality of their DEPs, ensuring that graduating students are able to meet the demands of further education in Australia.

Validation and moderation

Assessment

Academic staff at ELICOS colleges rely on **assessment** to inform their decision-making as to which students have the skills to continue on to further education. Staff need to be confident that the assessment tasks that students undertake have been designed to assess course learning outcomes in a way that is valid, reliable, fair and flexible. The process of ensuring this is known as **validation** and can be carried out both internally, i.e. within the college itself, and externally, by others outside the college such as specialist consultants, academics or staff from other colleges.

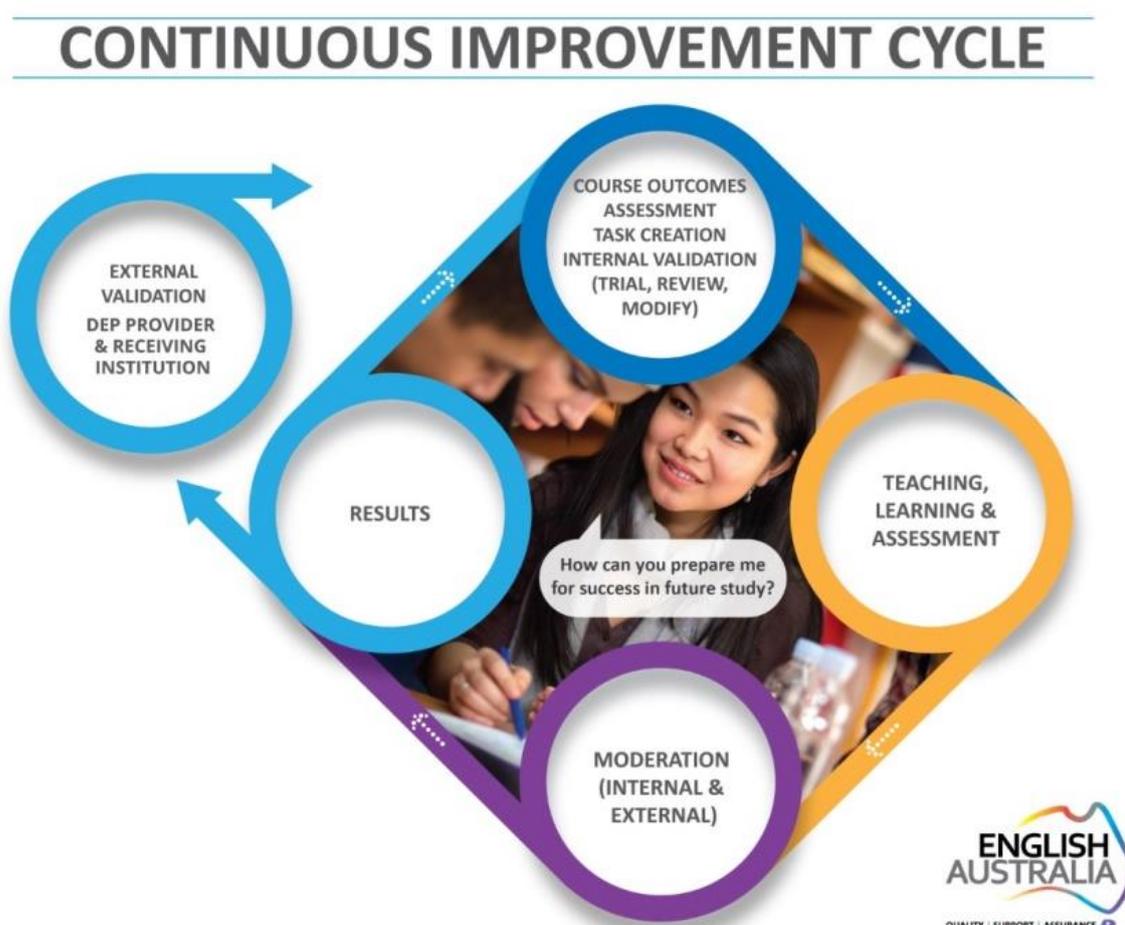# Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Once assessment tasks have been developed and validated it is essential that the decision making processes concerning student achievement levels are applied consistently to each student's work. This process is known as **moderation**, and is generally carried out by comparing decisions of those marking assessments.

Validation of programs

Alongside the need to validate assessments, there is a need to more broadly validate Direct Entry *Programs*, including course content, and this is most successfully done collaboratively between ELICOS providers and receiving institutions. Methodologies can include outcomes benchmarking, input by experts and student feedback.

Continuous improvement

Moderation and validation sit within the context of continuous improvement within the curriculum, as depicted in the diagram below. Colleges and receiving institutions are key partners in ensuring effective quality assurance mechanisms in moderation and validation for successful student outcomes.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Development of this resource

This resource is a supplement to the English Australia *Guide to Best Practice in Direct Entry Programs in ELICOS* (2017), which includes support for ELICOS DEP providers to offer best practice in student placement, progress and achievement. English Australia, in collaboration with the two ELICOS regulatory bodies, TEQSA and ASQA[1], identified the need for DEP providers to focus specifically on moderation and validation to further strengthen confidence in course outcomes. At the 2015 English Australia Conference in Brisbane, a workshop was conducted to identify best practice in moderation and validation in ELICOS colleges in Australia, and the outcomes of this workshop form the basis of this document. Over the following twelve months, further industry input was sought, gathered and consolidated into this supplement, supported by relevant theory and research.

---

[1] The Tertiary Education Quality and Standards Agency (TEQSA) regulates ELICOS colleges that are owned by or work with the higher education sector. The Australian Skills Quality Authority (ASQA) regulates all other ELICOS colleges.

Assessment task validation

Definition

The aim of assessment is to ensure that students have achieved course learning outcomes. Validation is the process of ensuring that assessment tasks or tools are designed to assess the learning outcomes that are being addressed. Each assessment task must be validated to ensure it meets the **core principles of assessment** (**face validity, reliability, fairness and flexibility**) and the **rules of evidence** (**content validity, sufficiency, authenticity and currency**).

An elaboration of these terms and others in bold throughout this supplement can be found in the glossary on page 29.

Best practice in assessment task validation

1.   Develop processes and procedures

Having a clear formalised process is integral to ensuring effective and efficient assessment task construction and validation. Clearly defined processes and procedures ensure that assessment tool development is efficient and outcomes-driven, and that the best possible use is made of resources. An ad hoc or unsystematic approach is not recommended and may result in tasks that are not appropriate to use in assessment. Processes may differ at individual centres but all should contain elements of the following stages:
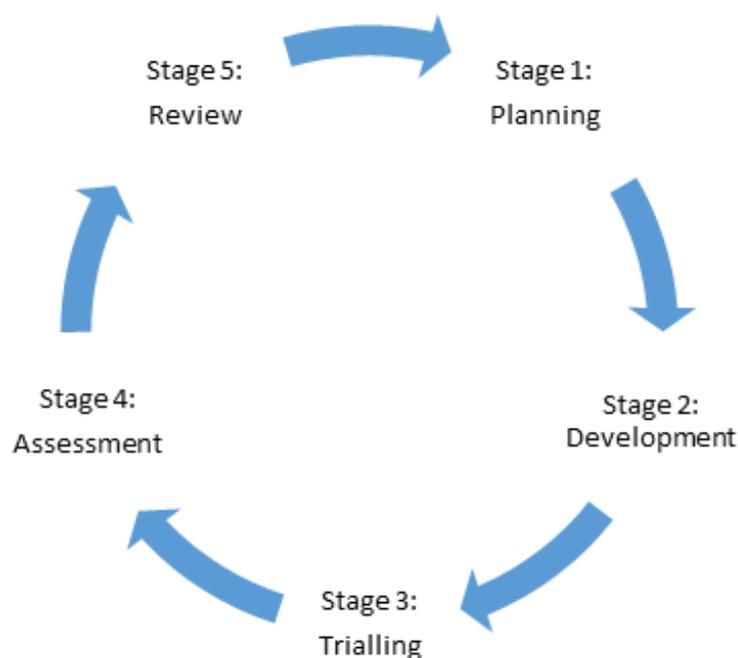

Figure 1: Process of task development

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Stage 1: Planning

All **assessment tools** (e.g. tests, essays, oral presentations) should have clearly-defined specifications that can guide task design.

Specifications should contain advice on how material should be selected and presented. These can include:

- the focus of the test/task

- outcomes being assessed

- suitable topic areas

- grammatical/lexical range being assessed

- number and length of sections (as appropriate)

- number of items in each section (as appropriate)

- timing of each section (as appropriate)

- format & length of tasks

- question types (T/F, open responses, multiple choice etc.)

- marks given for each item and total marks available

- details of weighting

- assessment criteria for essays, oral presentations, reports.

Stage 2: Task development

Multiple versions of assessment tools should be developed in relation to specifications for test security and to protect the integrity of the tools.

It is recommended that a core group of staff be involved in all stages of development and validation to evaluate whether principles of assessment and rules of evidence have been met.

The following additional points relate to the development of reading and/or listening test development:

Text selection

One effective approach is when developers search for 2-3 texts on the same subject. They then choose the appropriate text that lends itself to test development. Ideally, texts should require little modification in terms of relevance to the general reader and appropriate grading of language.

Item Writing

Developers increase efficiency by writing more items than are required. For example, if they need ten items per text according to specifications, they aim to write fifteen items as five of these may eventually not be suitable. If all items are suitable, extra items can be reused in another version of the test.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Useful resources on test planning and development

English Profile: http://www.englishprofile.org/wordlists

The English Vocabulary Profile contains information about phrases, words, idioms and collocations that are known and used by learners at each level of the Common European Framework (CEF).

Association of English Language Teachers in Europe (ALTE): http://www.alte.org/

Materials for the Guidance of Test Item Writers:
http://www.alte.org/attachments/files/item_writer_guidelines.pdf

Materials for the Guidance of Test Item Writers contains information relating to test construction. It provides many useful checklists to assist with validation. It also contains guidelines on how to conduct **item analysis** and useful criteria to assist with text selection.

Examples of best practice

• One centre has a core team of five teachers involved in assessment tool development. The centre manager acknowledged that although all teachers can be involved in the process it is essential to have a core group of select teachers involved who can follow the process from start to finish. Other centre staff can be involved in the next stages of task development. The centre manager emphasised that assessment tasks need to be 'shown to many eyes' and 'pulled apart' throughout development.

• Another centre's formalised process includes regular minuted curriculum meetings with **assessment validation** as a standing agenda item. There is a core group involved in assessment task validation but all teachers are invited to attend meetings. When an assessment task is undergoing validation, an Assessment Task Validation Report is completed. This is a checklist that meeting participants go through to review tasks and determine if the assessment instrument reflects the principles of assessment and meets the rules of evidence. (See Appendix A for examples of validation checklists.)

Stage 3: Trialling

Once an assessment task has been developed it needs to be piloted with students to make sure the instructions are clear and that the items can be completed without any problems being encountered. Results of the pilot test should then be evaluated by a core group of teachers and coordinators and any necessary adjustments can be made to the task. Time also needs to be devoted to constructing comprehensive answer keys and model responses. Note that the results of any assessments being trialled with students should not influence their course results unless the trial is later deemed a success.

Examples of best practice

• One centre pilots any new curriculum and related assessment with one class before it rolls out the new program to the entire cohort. This enables modifications to be made to assessment tools prior to formal delivery.

• Another centre uses item analysis software, such as SPSS, to analyse the degree of difficulty of each item and item discrimination. However, the centre manager emphasised that test responses can be analysed by simple statistics using Microsoft Excel. (See Appendix B for an example of a simple method used by one centre to analyse student responses on a test.)

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Stage 4: Assessment

The assessment task must be carried out under the same conditions every time it is undertaken. All teachers should be familiar with these conditions. Students should be informed of assessment conditions prior to an assessment taking place.

Centres demonstrating best practice in this area have formal guidelines outlining assessment/test delivery. These include areas such as;

• how desks/tables and seats should be arranged in a test room;

• whether support materials such as dictionaries may be used;

• how long the test should take and if breaks are allowed.

Stage 5: Review

Although review is ongoing throughout the entire process, to be fully effective, there should be a formalised review process after assessment has been carried out.

Examples of best practice

• Many centres reported that assessment tasks are reviewed at 'level meetings'. Any task-related issues identified by teachers are minuted and actioned. At one centre, teachers are required to give formal feedback to a level coordinator via an online curriculum log. A coordinator reviews the logs and highlights issues, which are then reviewed at the end of each intake and discussed and actioned at curriculum meetings.

• At some centres, teachers meet with level coordinators to review curriculum and assessment. Formal minutes are taken and key recommendations are provided to senior management. These recommendations are then prioritised and key items are acted upon.

• One centre has dedicated professional development days scheduled throughout the year during non-teaching periods, which allow academic managers to focus on key issues of assessment.

• A number of centres reported that they have to present results to an official English Assessment Committee made up of representatives from the Direct Entry Provider and the Receiving Institution. Results are reviewed and if there are any anomalies, actions are proposed with clear timelines for resolution.

• Student feedback collected via student evaluations at the end of each teaching period is also an important component of the review process. Many centres include specific questions relating to the assessment process and assessment tasks in formal student evaluations.

2.   Allocate resources

Assessment task creation and validation is very time consuming and, with the many demands that come with regular intakes, the challenge for ELICOS centres is to ensure that there is enough time allocated for academic staff to develop assessment tasks and take them through a systematic validation process.

Example of best practice

• Curriculum and assessment development is driven by senior management who are willing to commit resources to assessment and curriculum development as well as staff training.

3.   Allocate responsibilities

Duties relating to validation and assessment need to be well defined and allocated to specific people or roles so that an institution can be confident that all duties have been covered and everyone is aware of those responsible. Duties include scheduling validation meetings, creating and maintaining currency of assessment tasks and assessment task specifications, maintaining validation records, version control, maintaining currency of assessment task databases and conducting professional development activities for teachers.

Example of best practice

- A number of centres have validation as a regular item on curriculum meeting agendas.

- One large university language centre has four Assistant Directors of Studies, each with a separate area of responsibility. One of these positions has recently been reclassified as Manager of Assessment Validation. The reassignment of roles has given this person the space to specifically focus on assessment development and validation.

4.   Engage teachers in the process

It is extremely important to ensure that teachers are engaged in the process of assessment task construction and institutions have a responsibility to identify and build expertise in this area. As the Manager of Assessment Validation at one centre stated, 'It takes years of doing it to get very good at it.'

Examples of best practice

- One centre identified staff interested in assessment task design through performance management. These staff were given time out of the classroom to develop tasks, in a project that was managed by the Director of Studies. Staff had an initial briefing as well as regular opportunities to discuss their progress with others, including the project manager and colleagues.

- Many centres have organised an external consultant to run staff training on assessment development and validation.

- One centre employed a teacher who has training as a test writer. This person supports other teachers involved in test development.

5.   Benchmark tests

Different versions of the same assessment need to be benchmarked to ensure that they are equivalent in terms of both standard (i.e. degree of difficulty) and range of results.

Examples of best practice

- A number of centres regularly trial different versions of reading and listening tests with the same group of students to ensure that results are reliable. This can be done outside of scheduled class time. Some centres found that offering an incentive to students was necessary in order to encourage participation in trials. Incentives were wide ranging and included certificates of participation, coffee vouchers and feedback on their performance.

6.   Ensure test security

Assessment security is vital and without it, it is impossible to make confident decisions regarding student performance. Different versions of assessment tasks are essential if colleges need to select a number of assignments or reading and listening tests that can be given to students. If the same assessment task is used over and over again, test security can be compromised.

Example of best practice

•   Numerous centres store assessment materials (tests, answer keys, model responses) on a secure drive. Assessment instruments are not emailed, put on a USB or passed around in any form.

7.   Keep accurate records

Record keeping is an integral part of the formalised process. It is important to have clearly defined responsibilities with regards to record keeping.

Examples of best practice

Validation checklists

•   Several centres use checklists to evaluate assessment tools. These include an evaluation of features such as cultural appropriateness, level of difficulty, clarity of instructions and timing.

Assessment task database

•   A number of centres keep databases that allow developers to keep track of task/item usage. These typically record the tasks/test items that have been used for each intake and allow task creators to 'mix and match' different components of assessment instruments to create multiple versions. (See Appendix C: Sample Assessment Tool Database used at one centre to monitor assessment task usage.)

Version control

•   At one centre, responsibility for maintaining validation records and ensuring version control rests with a senior staff member. An Assessment Task Validation Record is maintained to keep track of the validation process.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS:
Supplement in Validation and Moderation

Moderation

Definition

Moderation in ELICOS DEPs is a process that aims to ensure that students' work is assessed accurately, consistently and fairly in relation to specific performance criteria. It focusses on particular examples of individual assessment in which there may be a significant degree of subjectivity in the assessment process (typically writing and speaking). Moderation is undertaken to identify and resolve differences in the way students' performance in assessments is judged, and allows for teachers to adjust their marks accordingly. It is carried out *after* an assessment has been conducted and *before* results are finalised, and is undertaken by staff at a particular centre (internally) and sometimes with the input of people with expertise in assessment from outside the centre (externally).

Standardisation differs from moderation in that the samples of student work under scrutiny are benchmark assessments usually taken from previous courses. The purpose of reviewing these assessments is not, therefore, to consider whether the marks assigned to them should be retained or altered (since they have already been finalised), but to assist in preparing assessors to mark to the appropriate standard before they are expected to mark current or future students' work. Standardisation and moderation are ideally undertaken as complementary processes, though some centres do one without the other.

While moderation is primarily a quality assurance mechanism concerned with establishing **intra-/inter-rater reliability** and minimising bias in marking, it may also:

- highlight possible refinements needed in the assessment task itself;

- assist in identifying teachers' skills, knowledge and professional development needs;

- alert teachers to suspected student plagiarism or collusion;

- support teachers by sharing the responsibility of marking;

- ensure the type and amount of feedback given to students is consistent; and

- contribute to the confidence that students, language centres, receiving institutions and other stakeholders have in the integrity of assessment results.

Best practice in moderation of assessment

1.   Set the standard

Effective standardisation provides teachers with an understanding of what is expected in terms of student performance before they participate in moderation. To do this, teachers need access to sample 'benchmark' assessments with assessment results and comments that they can refer to when they are marking. Ideally, students will have access to these too, so they are aware of the criteria they must meet to be successful.

The most common method of standardisation involves teachers using the assessment task performance criteria to evaluate a number of benchmark samples, then discussing their judgements in relation to the results and comments. This can be done with samples from either a previous round of assessment or from the current assessment, but done prior to moderation. If this is done with a previous assessment, then teachers are provided with samples that have been benchmarked previously. The experience of

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

standardisation can be very effective professional development, as teachers shape their understanding of assessment through discussion with colleagues and academic managers.

Examples of best practice

- Best practice in centres includes professional development sessions focusing on standardisation and using benchmark assessments, which are held regularly throughout the year.

- Some centres conduct standardisation meetings on writing assessments, where every teacher is given three papers before the meeting, with task instructions, assessment criteria and three benchmark samples, one with marks and comments as well as two that participants mark before the meeting.

- It is common practice that prior to students undertaking a seminar presentation, the teachers who will be assessing their performance take part in a standardisation meeting in which they view three videos of benchmark presentations and review the marking of these. The same teachers then conduct a similar session with their students.

- In one centre, as part of their induction, teachers new to a level are encouraged to access a range of benchmark writing tasks via the college online learning system before the course commences.

2. Develop moderation policy and procedures

Clearly-documented policies and procedures are central to effective moderation. These should be established by academic management staff as part of course planning and reviewed periodically.

A sound moderation policy or procedures includes:

a.   how staff and students are informed of the policy or procedures;

b.   which assessments/macro skills are moderated;

c.   a selection of samples of assessment;

d.   which members of staff are involved;

e.   the role of cross marking and double-blind marking;

f.   the timing of moderation;

g.   moderation as part of continuous improvement.

Examples of best practice

How staff and students are informed of the policy or procedures

- Policies and procedures relating to moderation are clearly set out and available for both staff (e.g. via meetings, induction, email, LMS portal) and students (e.g. in the student/course handbook, first day of class, on LMS).

Accessibility of Documentation

- Best practice is to have documents readily available and online access makes documents easily accessible to all stakeholders and ensures the process is transparent. One institution has its moderation policies available on its website.

Choosing which assessments/macro skills are moderated

- Moderation is most important in the marking of the productive skills of writing and speaking, which involve higher degrees of subjectivity. However, while listening and reading tests are generally

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

designed to be marked with reference to a list of acceptable responses, there may still be room for subjectivity in marking, depending on the type of items used. For example, if short answers are part of the task, there may be a variation in student responses. Tasks should be designed to minimise this variation. If practicable, teachers should mark listening and reading tests while sitting together, so that consensus can be reached in cases of ambiguity or uncertainty.

- With speaking assessments, colleges can provide a recorded sample of student performance for standardisation and moderation. Where this is not possible, speaking may be moderated by having two assessors present at the time of the assessment.

- For each macro skill in which assessments are to be moderated it is best practice to moderate all assessment tasks which contribute to students' final results. This may include **diagnostic assessments** completed soon after students start the course, as well as **formative assessments** done during the course, and will invariably include all **summative assessments** undertaken at the end of a course or unit of study. The advantage of also moderating formative assessments, such as practice tests, is that the process can play the role of standardisation in preparation for the corresponding summative assessments.

Selection of samples of assessment for moderation

- It is not possible to moderate all student work, so a common approach is to choose a range of samples at random, based on level of achievement: clear pass; pass/borderline; and fail. Many colleges have a policy of double or triple marking all failed assessments or assessments where students score significantly higher than expected, based on their previous performance.

Which members of staff are involved

- Moderation processes typically involve collaboration between a number of staff, especially teachers and senior teachers/academic managers. Collaboration is much more straightforward when staff meet face-to-face, though it is possible for moderation to take place with staff working online or individually.

- All teachers in one college are involved in marking the assessment under the guidance of the Course Coordinator. The Course Coordinator facilitates the session and, while respecting and involving all, has the final say if consensus is not achieved.

The role of cross marking and double-blind marking

- In the process of moderation, cross marking (marking assessments of students from another class) is best practice. In this way, teachers are not marking their own students' work and should therefore avoid bias.

- It is standard practice in many centres to cross-mark all major assessments during the initial assessment as well as during moderation meetings.

- Best practice is when failed papers are double-blind marked, so that each assessor is unaware of the grades or comments made by the other marker.

- One centre ensures that all failed papers are double or triple-blind marked and the coordinator decides on the final mark.

Timing of moderation

- Ideally, moderation happens as soon as practicable after the assessment has been conducted. Some colleges plan standardisation sessions in non-teaching weeks, so that the moderation process is expedited on marking day, while other colleges prefer to combine standardisation and moderation on the day of marking, so that the standard is set and fresh in the minds of all markers.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Moderation as part of continuous improvement

- Moderation can highlight issues that can contribute to improvement of the college academic program. These may include:
  - ➢ significant divergence between assessor ratings, which may indicate the necessity of professional development for staff;
  - ➢ potential problems with the validity of the assessment task;
  - ➢ suggested improvements to teaching methodology and/or course content that prepares students for the assessment;
  - ➢ ambiguities or omissions in the assessment criteria;
  - ➢ possible instances of suspected student plagiarism or collusion; and
  - ➢ observations on the distribution of marks and whether adjustments need to be made to the task or criteria.

If any of these arise during moderation and standardisation, the person responsible documents them, identifies clear action points and ensures that they are followed up.

Examples of best practice

- One college maintains a spreadsheet tallying each assessment and keeps records of marks given by teachers. Significant patterns of high or low marks are identified and the coordinator addresses any potential concerns with rater-reliability in a constructive manner with the teacher concerned.

- Through external moderation, another college found that most marks for a particular assessment task were in the middle of the range (with relatively few that were high or low) and therefore decided to review its moderation procedures. This in turn led to changes being made to the assessment criteria.

3. Consider external moderation

People from outside the institution such as specialist consultants, experts in the field or academic staff from other ELICOS institutions can bring a valuable perspective to assessment processes; however, for many colleges operating at a single location this can be problematic. While those outside the centre may be able to comment on the validity of specific assessment tasks and criteria, there may be limited benefit in having them take part in the moderation of students' work unless they have a thorough grounding in the wider context in which the assessment is situated. For example, if external moderators do not have a sound awareness of the course content and have not taken part in standardisation, their inter-rater reliability is likely to be compromised.

Institutions with more than one location, on the other hand, need to engage in cross-campus moderation to ensure consistency across the wider organisation (which might be considered, more properly, as internal moderation). If face-to-face moderation meetings are difficult to arrange, tele/video-conferencing and the sharing of documents electronically may be an effective alternative.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

External validation

Definition

Sound policies and processes for validation and moderation of assessments and programs can provide an organisation with confidence in its internal quality assurance mechanisms. However, receiving institutions may need assurance from *outside* the organisation that the outcomes of its DEPs really are equipping students for success in their further study. This is called *external validation*.

External validation generally comprises three main features: benchmarking; the input of experts; and student feedback.

Benchmarking

Benchmarking, in this context, is where institutions work together to examine their policies, practices and student outcomes in a structured way so that they can identify where they can make improvements (Australian Government Tertiary Education Quality and Standards Agency, TEQSA Guidance Note Benchmarking. Retrieved from: www.teqsa.gov.au/sites/default/files/BenchmarkingGNFinal_0.pdf.

In this Guidance Note, TEQSA identifies five main aspects of benchmarking:

- organisational benchmarking, where internal comparisons are made (internal validation);
- course benchmarking of courses and student results;
- process benchmarking of practices and processes;
- outcomes benchmarking, where student outcomes may be compared with those from other organisations; and
- best practice benchmarking, where an institution compares its processes and outcomes with those of an organisation that is seen as 'at the forefront' in terms of programs and outcomes.

Under TEQSA Provider Registration Standard 5.6, providers need to demonstrate that they have some kind of process to compare their performance with that of others. This is called *outcomes benchmarking* and it is an often utilised form of benchmarking in external validation at ELICOS centres.

ASQA states that validation should include using a 'statistically valid sample' (ASQA Fact Sheet – Conducting validation, 2015. Retrieved from: www.asqa.gov.au/media-and-publications/conducting-validation1.html. It defines a statistically valid sample as:

1.   large enough that the validation outcomes of the sample can be applied to the entire set of judgements, and

2.   taken randomly from the set of assessment judgements being considered.

Input of experts

In order to externally validate programs or assessments, it may be useful to utilise those who may have extended experience in either program development or assessments. For the purpose of this document, this input will be considered as 'input of experts'.

Student feedback

Student feedback on courses and assessment can be a form of validation of programs as well as providing a valuable form of input into improving programs and assessment.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Best practice in external validation

1.   Form collaborative partnerships

ELICOS providers stand as gateways to further/higher education, and therefore collaboration between the ELICOS provider and the receiving institution is often key to ensuring benchmarking practices lead to future student success. The direct entry provider and the receiving institution work closely together in a collaborative manner to decide on the most appropriate methods of providing external validation to their programs

2.   Define financial responsibilities

In a collaborative relationship, financial responsibilities for the DEP provider and the receiving institution(s) are clearly defined, and resources required for external validation are factored in to an institution's budget as an essential cost.

3.   Manage time

External validation occurs at regular intervals and the timing is carefully managed so as not to increase operational burden.

Examples of best practice

Benchmarking

The most common forms of outcomes benchmarking carried out by English Australia member colleges are exit-testing of students using an industry-recognised proficiency test; tracer studies of DEP graduates carried out by receiving institutions; and comparison analyses undertaken with other DEP providers. Benchmarking can provide a valuable external reference point in validating ELICOS programs and assessment.

It is imperative that the frequency, reporting and funding of benchmarking is agreed to by receiving institutions and ELICOS providers and that these agreements are documented. Examples of best practice in the various types of benchmarking and the frequency, reporting and funding of these can be found in the table on page 19.

Input by experts

•   Some colleges utilise external markers of major assessments as a means of external validation of assessments. The external markers are either teachers with considerable experience who may be considered experts, but not working as teachers at the associated college, or experts at 'head office' who are responsible for the validity of courses and assessments.

•   Some colleges have an advisor from the associated receiving institution who sits on its Academic Board. The advisor may oversee and/or approve any major changes to the curriculum and will usually have specific postgraduate qualifications in assessment and course development. They may also check that the course objectives are mapped and connected to the assessments.

•   As well as working collaboratively with on-campus providers, experienced academic staff at receiving institutions are actively involved with external off-campus ELICOS providers, whether in Australia or overseas, to ensure programs and assessments meet the same rigour and consistency as on-campus providers. These experienced staff provide collaborative support and guidance to these providers, including professional development resources and/or advice where appropriate.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Validation and Moderation

Student feedback

- A number of colleges survey ELICOS students towards the end of their first year of their further studies course in order to validate and evaluate the content and outcomes of their direct entry course. Feedback is taken and suggestions actioned in order to improve the course and better meet the needs of the students.

- One college considers that students who have completed a year of their degree are in a good position to be able to give valuable feedback on their ELICOS course and how it prepared them for their studies. This information validates successful practices in program construction, delivery and assessments and also highlights areas for improvement. The college considers all input then formulates an action plan for improvements in specific, identified areas.

The cycle of external validation and continuous improvement

External validation of courses and assessments provide quality assurance in student outcomes and also provide valuable feedback for improvements. Best practice in external validation and evaluation of courses and assessments includes, as an essential feature, the cycle of continuous improvement. This process then feeds back into improvements in internal validation and moderation processes.

Best practice in validation and moderation of courses and assessments is one that is achieved collaboratively by continual improvements achieved by seeking and acting on feedback and input from relevant stakeholders.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Moderation and Validation

External validation: Outcomes benchmarking. Examples of best practice:

Method	Frequency	Reporting	Funding
Proficiency assessment A sample group of ELICOS students who are on pathways for undergraduate and postgraduate courses undertake an industry-accepted proficiency test, such as IELTS or TOEFL, which is accepted by the receiving institution for entry to their courses. In most centres, two groups of students, one group in the last weeks of their course before they are due to enter Bachelor Degree courses, and one group about to enter post-graduate study, undertake the external test. Students are usually selected on a voluntary basis. This sample benchmarking allows the receiving institution and the ELICOS centre to have confidence that students on direct entry pathways to the receiving institution are demonstrating a comparable level of language level proficiency as students who enter the university via an industry-accepted proficiency test.	Annually or biennially	Included in the college's Academic Board report.	ELICOS institution or receiving institution, depending on the Memorandum of Understanding.
Tracer studies Analysis is typically undertaken by the receiving institution (RI) of anonymised data of students who have come to them on an ELICOS pathway. The tracer studies commonly look at the following factors, compared with non-ELICOS pathway students: • retention rate • weighted average mark (WAM) Depending on the capabilities of the RI's database, this analysis can examine: • students undertaking undergraduate studies • students undertaking postgraduate studies • students undertaking degrees in different faculties and undertaking different courses.	Annually Comparisons are usually made over the past four years.	Included in the college's Academic Board report.	Receiving institution
Comparisons with other colleges A number of different colleges, especially where there is a relationship between the different colleges, undertake benchmarking of course results and pass rates. This provides an external means of validating student outcomes.	Every course-annually, as agreed between the colleges involved.	Stakeholders at all colleges and RIs involved.	Any costs are commonly borne by the colleges involved.

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Moderation and Validation

Sample Assessment Validation Checklist 1

Course	
Assessment Task	
Validation Team	

Description	Agreed (Y/N)
There is clear alignment between the unit learning outcomes and the assessment task/s.	
The assessment measures student achievement of the intended learning outcome/s at an appropriate level.	
The assessment is as objective and as fair as possible, taking into account learning styles and diverse student cohorts	
The assessment documents are free from ambiguities, grammatical errors and spelling mistakes.	
The timeframes set for all assessments are reasonable.	
The weighting is appropriate to the requirements of the tasks or questions.	
Marking criteria, rubrics and guides are clear.	
Examination papers only:	
All relevant materials have been provided within the papers.	
Alternate, equivalent versions of exams are available for special/alternative examinations.	
The presentation and layout of the examination paper is in the approved university format.	
Recommended actions	

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Moderation and Validation

Acknowledgement and signatures:

Validation Team	Signed	Date

Coordinator/Director of Studies acknowledgement of receipt and agreement to report and/or action items raised:		
Signed:		Date:	

Guide to Best Practice in 'Direct Entry' Programs in ELICOS: Supplement in Moderation and Validation

Sample Assessment Validation Checklist 2

Assessment tool:	Course:
Outcomes Assessed:	Validation Team:
Version:	Date:

	Y/N	Comments
1. The assessment tool:		
a. addresses the learning outcomes being assessed.		
b. is at the appropriate level for the students.		
c. has clear instructions.		
d. is appropriately formatted and well designed in terms of layout, with sufficient space for student and assessor responses.		
e. is easy and practical to use.		
f. does not contain any gender, cultural or other bias.		
g. is labelled with date and version number.		
2. Does the assessment tool include:		
a. instructions for the assessor?		
b. instructions to the student?		
c. assessment criteria?		
d. guidelines for administration?		
e. guidelines for recording and reporting?		
Modifications Required:		

Appendix B: Validation: example simple text analysis

EAP4 W10 ASSESSMENT TASK 7 READING TEST
US Immigration, Mosaic/Melting Pot, Cycle of Success

Student ID	Text 1 Qns 1-15	Text 2 Qns 16-30	Text 3 Qns 31-45	Total	%
1	4	10	7	21	47
2	8	9	5	22	49
3	6	11	5	22	49
4	9	10	5	24	53
5	9	9	8	26	58
6	7	12	9	28	62
7	8	10	10	28	62
8	8	9	12	29	64
9	8	13	8	29	64
10	13	11	12	36	80

Text 1 - US Immigration (Qns 1–15): Summary Headings (1–5), Vocabulary (6–15)

Text 2 - Mosaic/Melting Pot (Qns 16–30): TRUE, FALSE / NOT GIVEN (16–25), Multiple Choice (26–30)

Text 3 - Cycle of Success (Qns 31–45): Multiple Choice (31–35), Paragraph (36–40), Ordering a Conclusion (41–45)

	CODE	
	X	Incorrect response
		Bottom third
		Top third

Guide to Best Practice in 'Direct Entry' Programs in ELICOS:
Supplement in Moderation and Validation

The previous table illustrates a simple way that Excel can be used to analyse text responses. Once a test has been trialled, the following steps are recommended:

1. Order student results from lowest to highest.
2. Identify the top third and the bottom third. In this example, the top third is shaded blue and the bottom third is shaded yellow.
3. Analyse text responses according to rate of correct/incorrect responses.
4. Assumptions for text anlaysis are as follows:

   - If 70% or more are getting an item incorrect, the item is too difficult.
   - If 70% or more are getting an item correct, the item is too easy.
   - Items with a correct response rate between 30% - 70% are generally useful.
   - Most students who get a high final score should be getting any given item correct more than those whose score is low.

5. **In the above example, many of the responses need to be reviewed as they are either too easy or too difficult.**

The above is an example only. A larger sample should be used to anlayse text responses.

- It is highly recommended that a number of assessment methods be used to gather evidence about a student's performance.
- For more complex anlaysis, consult the following resources:

- *Association of English Language Teachers in Europe (ALTE)*: http://www.alte.org/
- *Materials for the Guidance of Test Item Writers:* http://www.alte.org/attachments/files/item_writer_guidelines.pdf

Guide to Best Practice in Direct Entry Programs in ELICOS:
Validation and moderation

Appendix C: Sample Assessment Tool Database used at one centre to monitor assessment task usage

COURSE: EAP5 ASSESSMENT TASK DATABASE

	READING	LISTENING	WRITING	SPEAKING	TRIAL TESTS
Intake 9 (2016)					
W5	EAP5-W5 T6 Reading Test (A) QUESTION BOOKLET v1.0 NC MH WE 30112016	**EAP5-W5 T5 listening Test (A) AUDIO & PPT v1.0 WE 02122016**	EAP 5 W5 Writing Test Workplace Motivation Synthesis	**EAP5-W5 T3 Seminar TEACHER FEEDBACK v1.0 MH 04082016**	
W10	EAP5-W10 T6 Reading Test (B) QUESTION BOOKLET With answers v1.0 NC MH WE 30112016	**EAP5-W10 T5 Listening Test (A) AUDIO PPT v1.0 WE 02122016**	**EAP5-W10 T4 TIMED WRITING TEAMWORK**	**EAP5-W10 T3 Seminar Presentation FEEDBACK & RESULT v1.0 MH 09092016**	
Intake 1					
W5					
W10					
Intake 2					
W5					
W10					
Intake 3					
W5					
W10					
Intake 4					
W5					
W10					

Glossary

Assessment	A process to determine a student's achievement of identified learning outcomes. It may include a range of written and oral methods and practice and demonstration. Tertiary Education Quality and Standards Agency, Glossary of Terms, 2012. Retrieved from: http://www.teqsa.gov.au/media-publications/glossary
Authenticity	The evidence presented in an assessment is the learner's own work. *Authenticity* in assessment also refers to 'the degree of correspondence of the characteristics of a given language test task to the features of a language task' (Bachman and Palmer: 1996, cited in Brown & Abeywickrama: 2010: 36). In other words, an assessment task is 'authentic' to the extent that it mirrors 'real world' language use.
Assessment Tools	Incorporate both the assessment instrument and the rubric or instructions for gathering and interpreting student performance. It also includes scoring guides, rating scales and descriptors/criteria or frameworks used to evaluate student performance. Assessment tools include: • Oral presentations • Interviews • Tutorial discussions • Research Essays • Examination or timed essays • Reflective Journals • Project work (group/individuals) • Student Portfolios • Tests of Achievement (Listening, reading, language skills).
Assessment validation	The process of ensuring that assessment tasks are well designed to address the learning outcomes being addressed. The goal of assessment validation is to ensure that the assessment tools provide valid, reliable, sufficient and fair evidence. Each assessment tool must be validated to ensure it meets the principles of assessment (face validity, reliability, fairness and flexibility) and rules of evidence (content validity, sufficient, authentic and current).
Content validity	A principle of assessment which ensures the assessment covers a representative sample of the content/skills that should be assessed. For example, a summative assessment task should contain items that cover the entire course not just a portion of it. Assessment tasks should aim for high content validity.

Guide to Best Practice in Direct Entry Programs in ELICOS: Validation and moderation

Core principles of assessment	Fundamental principles which guide validation in assessment and include: face validity, reliability, fairness and flexibility.
Currency	The assessment evidence demonstrates a learner's current level of competency, skills and knowledge.
Diagnostic assessment	This 'is designed to diagnose specified aspects of a language' (Brown & Abeywickrama: 2010: 348) and is often used to gauge students' abilities in terms of the four macroskills in order to ensure that they are placed in an appropriate class for their level.
Face validity	'Face validity refers to the degree to which a test looks right, and appears to measure the knowledge or abilities it claims to measure, based on the subjective judgment of the examinees who take it, the administrative personnel who decide on its use, and other psychometrically unsophisticated observers' (Mousavei, 2009, cited in Brown & Abeywickrama: 2010: 35).
Fairness	A principle of assessment that means the individual learner's needs are considered in the assessment process. Where appropriate, reasonable adjustments are made to assessments to take into account individual learner's needs. All learners are informed about the assessment process at the beginning of a course. (Australia Skills and Quality Authority, Users Guide to the Standards for RTOs, 2015. Retrieved from: https://www.asqa.gov.au/standards/about-standards-rtos-2015/standard-one/clauses-1.8-1.12
Flexibility	A principle of assessment that affords flexibility to the learner by drawing upon a range of assessment methods that are appropriate to the context, which may include reports, essays, oral presentations and tests. (Australia Skills and Quality Authority, Users Guide to the Standards for RTOs, 2015. Retrieved from https://www.asqa.gov.au/standards/about-standards-rtos-2015/standard-one/clauses-1.8-1.12
Formative assessment	Information collected (generally via a range of formal and informal methods) during a course to determine students' progress towards course outcomes and learning goals. Its purpose is to provide feedback in order to adapt or change teaching contexts and approaches or to adapt or change student learning and study strategies. (*National standards for ELICOS providers and courses (ELICOS Standards), 2011,* Glossary.)
Intra-rater reliability	The 'condition in which the same scorer yields consistent scores across all tests.' (Brown & Abeywickrama: 2010: 352.)
Inter-rater reliability	The 'condition in which two or more scorers yield consistent scores for the same test.' (Brown & Abeywickrama: 2010: 352.)

Item Analysis	An analysis of each item or question on a test to determine the proportion of students selecting each answer.
Moderation	The process of ensuring that marks or grades are awarded appropriately and consistently. It is a collaborative process with the primary aim of achieving consistency in the application of assessment criteria.
Reliability	Evidence considered in the marking of an assessment is consistently interpreted and assessment results are comparable irrespective of the assessor conducting the assessment.
Rules of evidence	Rules which relate to the way evidence is collected in assessments. The 'rules' state that this evidence must be: valid, current, sufficient and authentic.
Summative Assessment	Often carried out during or at the end of a course of study to determine and specify students' achievement of course outcomes and learning goals. (*National standards for ELICOS providers and courses (ELICOS Standards), 2011,* Glossary.)
Standardisation	Standardisation differs from moderation in that the samples of student work under scrutiny are benchmarked assessments taken from previous courses. The purpose of reviewing these assessments is to assist in preparing assessors to mark to the appropriate standard before they are expected to mark current or future student work. Standardisation and moderation are ideally undertaken as complementary processes.
Sufficiency	The assessment must allow for enough evidence to be collected to make an accurate judgment about a learner's skill level.
Validity	The assessor is assured that the learner has the skills, knowledge and attributes as described in the module or unit of competency and associated assessment requirements.

References

Arkoudis, S., Baik, C. & Richardson, K. (2012). *English Language Standards in Higher Education: From entry to exit*. Melbourne, VIC: ACER Press

Australian Government, Australian Skills Quality Authority, *Fact sheet—Conducting validation.* Retrieved from:

http://www.asqa.gov.au/media-and-publications/conducting-validation1.html

Australian Government, Australian Skills Quality Authority, Users' guide to the standards for RTOs 2015. Retrieved from:

https://www.asqa.gov.au/standards

Australian Government, Department of Education and Training, National Standards for ELICOS providers and courses. Retrieved from:

https://internationaleducation.gov.au/Regulatory-Information/Education-Services-for-Overseas-Students-ESOS-Legislative-Framework/ELICOSnationalstandards/Pages/Default.aspx

Australian Government Department of Education and Training (2016). *Research Snapshot: Study Pathways of International Students in Australia, 2014-2015.* Retrieved from:

https://internationaleducation.gov.au/research/Research-Snapshots/Documents/Study%20pathways%202016%20snapshot.pdf

Brown, H., & Abeywickrama, P. (2010). *Language Assessment, Principles and Classroom Practices (2nd ed.).* White Plains, NY: Pearson Education

Elder, C. & O'Loughlin, K. (2003). 'Investigating the relationship between intensive EAP training and bandscore gains on IELTS'. *IELTS Research Reports 4*. Canberra: IELTS Australia

English Australia (2017). *Guide to Best Practice in Direct Entry Programs*. Sydney: English Australia

Tertiary Education Quality and Standards Agency. *TEQSA Guidance Note – Benchmarking*. Retrieved from:

http://www.teqsa.gov.au/sites/default/files/BenchmarkingGNFinal_0.pdf

Tertiary Education Quality and Standards Agency, Glossary of Terms, 2012. Retrieved from:

http://www.teqsa.gov.au/media-publications/glossary