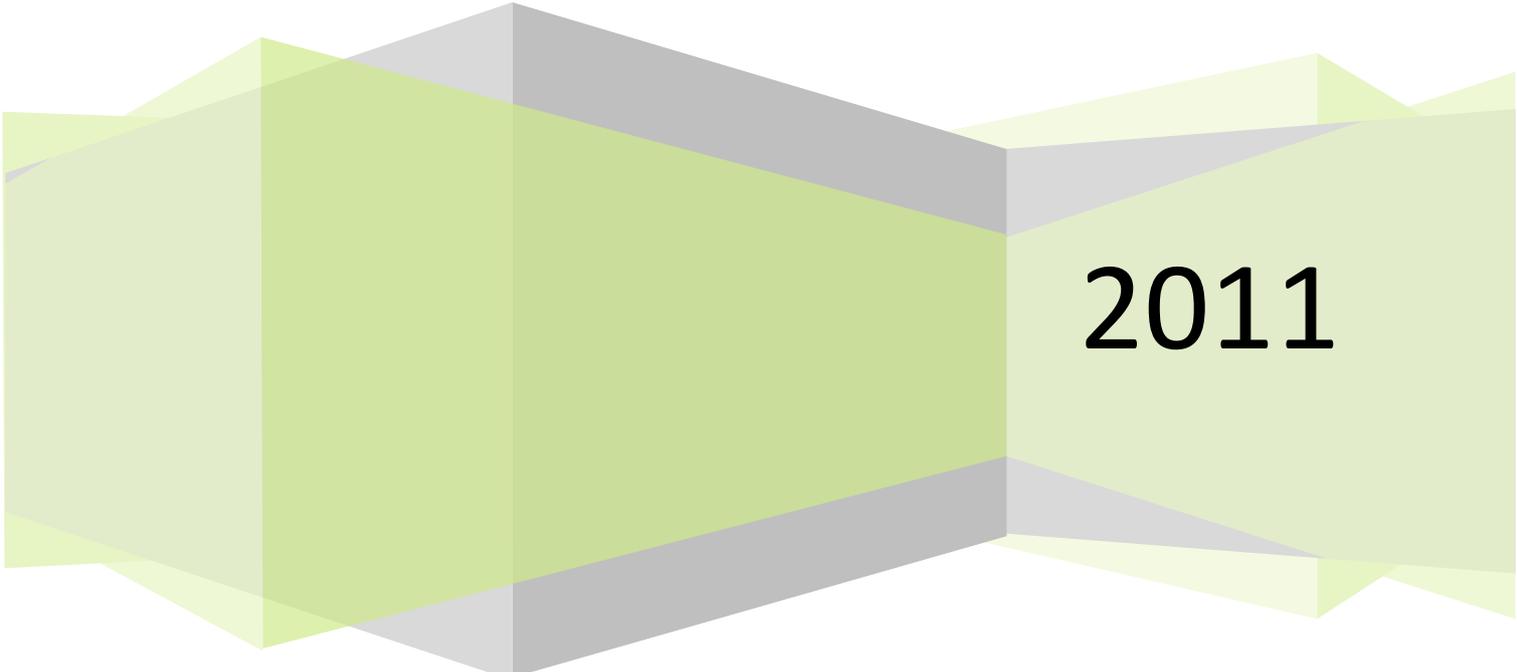


*Quality assurance framework for
Student assessment*

**Guidelines for the design and implementation
of effective student assessment**

Lai Jiang



2011

Introduction

BACKGROUND, RATIONALE AND PURPOSE

Founded in 2008, the educational network-LINQED is a project within the 3rd Framework Agreement (FA3) between ITM Antwerp and the Belgian Directorate General Development Cooperation (DGD). Quality in education is the central theme of LINQED. Its overall goal is to strengthen training capacities in the field of international health at all member institutions. Developing a reference framework for educational quality assurance was considered a necessary approach to achieve this goal. On the one hand, there was a consensus that common frameworks for (internal and external) quality (and quality assurance) would hardly take into account all specificities (e.g. resources, culture and organisation) of network partners. On the other hand, the scientific basis of pedagogy and the endeavour to establish high quality postgraduate training for the provision of health care is universal. Notwithstanding variations, there is an increasing degree of equivalence of structure, process and product of postgraduate medical/public health education worldwide. Hence, providing a set of “guiding quality criteria” for key processes in education was considered to be necessary especially in view of (international) collaboration. These guiding criteria mainly clarify what is meant by the quality of specific aspects (e.g. student assessment, curriculum design or didactic approach).

With the aforementioned rationale, this document aims at formulating guiding criteria and standards as a basis for Internal quality assurance (IQA) of one of the key processes in education—student assessment. **There is no intention to foster uniformity of educational practice.** It is not LINQED's intention to prescribe how higher education providers must implement the precepts set out in the next section, the aim of which is to assure good assessment practice. Instead, **it is expected that this document will stimulate the sharing among partners of concrete experiences with various IQA methods and tools that deploy this framework.** In the near future, exchanges of IQA tools and practices among network partners will take place.

It is important to note that the identified criteria resulted from a synthesis of many important quality assurance issues that were discussed in the third annual LINQED workshop. By analysing specific assessments, participants came to a consensus on several criteria which are of vital importance to student quality assurance:

- Ensure ***cognitive complexity*** of assessment tasks—assessments must reflect higher cognitive skills and assess students’ thinking processes;
- Ensure assessments are ***meaningful***—assessments must benefit students as well as teachers;
- Ensure the chosen assessment methods are ***fit for*** their ***purposes***—assessment goals and methods should be compatible with the overall educational goals;
- Ensure implemented assessments are ***fair***— assessments should provide opportunities to all students to demonstrate their abilities in various ways;

- Ensure implemented assessments are ***transparent***—scoring criteria and standards must be clearly communicated from the start and understood by all stakeholders (i.e., learners, assessors and external controlling agencies);
- Ensure positive ***educational consequences*** of assessments—assessments must have a positive effect on learning and instruction;
- Ensure assessments stimulate ***self-regulated learning***—assessments should provide opportunities for students to practice in self-/peer-assessment; students should always be encouraged to use assessment feedback to improve learning;
- Ensure assessments are set up and carried out in a ***comparable*** way— various assessment methods and different assessment tasks may be used to assess the same competence. In this case, criteria and assessment conditions should always be consistent to make sure that scoring is consistent for all students.

Unsurprisingly, what was highly valued by LINQED participants is also advocated in several evidence-based research on student assessment (e.g., [Baartman, 2008](#); [Black & Wiliam, 1998, 2003](#); [Huba & Freed, 2000](#); [Suskie, 2004](#)). Four criteria, which were not discussed in the workshop but emphasised in other assessment guidelines and research, have been added to the current framework. They are:

- ***Authenticity***—the content of assessment tasks, the way they are assessed and the context in which they are assessed must appropriately reflect the competency needed to solve the problems that students will encounter later in their work (AAHE, 1991; Baartman, 2008).
- ***Appropriate decisions*** — in order to yield reasonably accurate and truthful conclusions, multiple sources and types of relevant information about persons or programs should be used when making educational decisions/judgements (Baartman, 2008; Bresciani, 2003; NCME, 1995; Suskie, 2004).
- ***Costs- efficiency***— when developing assessments, one must keep a balance between the time and resources needed and their benefits (Baartman, 2008; Huba & Freed, 2000; Suskie, 2004).
- ***Acceptability***— the collection, interpretation, and use of student learning evidence is a collective endeavour, and is not viewed as the sole responsibility of a single office or position. The principles, criteria and standards specified for assessments should be accepted and approved by all stakeholders to ensure they will be used (Baartman, 2008¹; Council of Regional Accrediting Commissions, 2004).

Furthermore, this framework intentionally takes due account of available resources that may give rise to the implementation of good practice and strategies on student assessment. A set of practical guides can be found in the appendix and/or links (e.g., [the European Association for Quality Assurance in Higher Education \(ENQA\)](#) ; [World Federation for Medical Education \(WFME\)](#) ; [The Learning and Teaching Support Network \(LTSN\)](#); [AAC&U](#); [Quality Assurance Agency for Higher Education \(QAA\)](#)).

¹ The framework of quality for Competence Assessment Programmes (CAPs), Baartman, 2008 is presented in [Appendix A](#).

This framework is directed at the LINQED network. However, we expect this document to achieve a wider circulation among those with an interest in quality assurance on student assessment. These readers will hopefully find it useful and inspirational.

NATURE AND PURPOSES OF ASSESSMENT

“If you want to change student learning then change the methods of assessment”.

— Brown et al., 1997, p. 9

“Assessment is the process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences; the process culminates when assessment results are used to improve subsequent learning”.

— Huba & Freed, 2000, p. 8

“Assessment is a planned continuous process to gather and synthesize information relevant to the purposes of (a) discovering and documenting students’ strengths and weaknesses, (b) planning and enhancing instruction, or (c) evaluating progress and making decisions about students”.

— Cizek, 1997, p.10

The three quotations above provide three basic guidelines adopted for the current framework:

- ***Student assessment is a crucial aspect of a successful teaching and learning system.***

Brown et al.’s words pinpoint why we should put considerable emphasis on the quality of student assessment—the nature of the assessment strongly influences what students actually do to study. Good assessment drives students to achieve and demonstrate the intended learning outcomes, i.e., gaining competence at the end of the training programme to perform professional tasks at an appropriate level.

The assessment of competence consequently calls for an **assessment programme** approach (Bartman, 2008) as it is agreed that no single assessment method is sufficient to assess competence, because it is such a complex whole of knowledge, skills, and attitudes (e.g., Van der Vleuten & Schuwirth, 2005). A programme of different assessment methods could enable teachers and assessors to better capture the complexity of competence and as such generate a more valid picture of the student’s development.

- ***Using assessment as a continuous process of guiding and supporting learners’ progress throughout (and beyond) education should be our primary focus.***

A considerable body of evidence has shown that assessments can be strongly associated with gains in learning if they are used formatively. An effective way to integrate formative assessment² into practice is to provide feedback and engage learners to generate

² The terms ‘formative’ and ‘summative’ evaluation of student learning —what we today would tend to call assessment— were first introduced by Benjamin Bloom in 1969 to refer to two kinds of roles assessment can play. The first role is the traditional role that assessments play in judging and classifying students. Summative assessment is a judgement (usually a mark) which encapsulates all the evidence up to a given point. For instance, the final examination for obtaining the diploma/certificate is a summative assessment.

appropriate learning activities for future work while using assessment. **Feedback should play a pivotal role in helping learners to develop their self-regulative skills (e.g., plan, monitoring, evaluation and adaptation).**

- **Assessment can either be used as a process towards improvement (i.e., assessment for learning), or a process towards accountability (i.e., assessment of learning). An assessment process can, and often should, involve more than one of these assessment purposes.**

Assessment is usually construed as formative (*diagnostic and guide*) or summative (*evaluate and certify*). An assessment process can, and often does, involve more than one of these assessment purposes. For example, an assessment component submitted during a module may provide formative feedback designed to help students improve their performance in subsequent assessments. An end-of-module or end-of-programme examination or other assessment normally results in a summative judgement being made about the level the student has attained, but any feedback on it may also have an intended formative purpose that can help students in assessment later in their programme, or on another programme (QAA: Assessment of students, 2006).

- **In addition to its benefits for students (i.e., to guide and evaluate their development and to certify them), assessment should also be used to enhance the quality of instruction.**

Assessment should help teachers make better educational decisions. Teachers should use the results of the assessment to determine whether these particular outcomes are essential components in the sequence of learning. If problems are identified, remedial work should be provided to give students a second chance to acquire a particular skill or concept.

The term “assessment” used in this document adopted the concept of *Competence Assessment Programme* (CAP) which was defined by Baartman (2008) as a combination of both traditional and new assessment methods that can be used for both formative and summative assessment purposes and both for learners and instructors.

The concept of CAP does not set the boundary on how the actual combination of assessment methods should be because the contents of a CAP depends on the competences under assessment and the breadth of the educational programme (i.e., a specific course, a semester, a school year, etc.) (Baartman, 2008, p. 31).

In the next section, The LINQED Quality Assurance framework on student assessment is presented. Precepts and explanations are given to answer the question “**what are the criteria we should take to guarantee the quality of CAPs?**”

Quite in contrast, the purpose of **formative assessments** is to make sure that the evidence about student achievement is produced, interpreted, and used by teachers, students, or their peers to acknowledge the existence of a ‘gap’ between the actual level of the work being assessed and the required standard. In addition, formative assessment also provides an indication of how the work can be improved to reach the required standard. A crucial feature of formative assessment is that individual students/teachers use the obtained information and that it **benefits their learning/teaching**.

The LINQED Quality Assurance framework on student assessment— Criteria & Standards

The LINQED QA framework on student assessment recommends the following set of global standards. The standards are structured according to 12 key criteria. STANDARDS are specified for each criterion using two levels of attainment:

- **Basic standard.** This means that the standard must be met by every institution and fulfilment demonstrated during evaluation.

Basic standards are expressed by a “**must**”.

- **Standard for quality development.** This means that the standard is in accordance with international consensus about the best practice for public health schools and basic medical education. Fulfilment of - or initiatives to fulfil - some or all of such standards is left to LINQED member institutes to document.

Fulfilment of these standards will vary with the stage of development of the assessment system, resources and educational policy in member institutes. Even the most advanced institution might not comply with all standards. Standards for quality development are expressed by a “**should**”.

Annotations are used to clarify, amplify or exemplify expressions in the standards. The explanations provide a rationale, and in some cases examples, to support the precepts. Where examples are provided, their purpose is to illustrate concepts, and sometimes to refer to what might be considered good practice, depending on the context and subject, and the students being assessed. These examples are not intended to form a checklist and were chosen to exemplify the concepts being explained.

PRECEPTS AND EXPLANATIONS³

It is very important to note that not all methods included in a CAP must meet all criteria listed below, but that the assessment programme as a whole must.

During the LINQED workshop, there were intensive discussions on how to properly design four assessment methods (i.e., multiple choice questions, short answer questions, essay and professional behaviour assessment). One of the group exercises asked participants to reflect on their own example assessments: "How can the content of the current test be reformulated using a different assessment method (e.g., can one redesign multiple choice questions into an essay type of assessment?)? What are the advantages and disadvantages of using these two methods for the given content?" By doing this mental exercise, we realised that different assessment methods have their own strengths and weaknesses. The selection of the assessment method largely depends on what needs to be assessed. Diversity of assessment methods is needed to test a wide range of intended learning outcomes (knowledge, skills and attitude). These ideas which reflect the FITNESS FOR PURPOSE criterion were underlined in the action plans of several institutions (for an example see the action plan of [UMSS-Bolivia](#))

1. Criteria: Fitness for Purpose

Basic standard:

- Assessment content **must** be clearly compatible with educational objectives⁴ and with what has been taught.

Quality development:

- The assessment methods **should** be appropriate for assessing predefined practice requirements in terms of knowledge, skills and attitudes.
- A complementary set of assessment methods **should** be applied to fulfil these requirements.

Annotations:

Assessment must be designed to measure the achievement of the intended learning outcomes (i.e., the competences needed in the future work place) and other programme/course objectives. If a mismatch between the messages of the instruction and the assessment is perceived by students, it is unlikely that assessment will have a positive impact on student learning (Segers, Dierick & Dochy, 2001).

To test a wide range of intended learning outcomes, diversity of assessment methods between and within different subjects is to be expected and welcomed. The definition of methods used for assessment may include consideration of the balance between formative and summative assessment, the number of examinations and other tests, the balance between different types of examinations, the use of normative and criterion referenced judgements, and the use of portfolio and special types of examinations, e.g. objective structured clinical examinations (OSCE).

³ The content of this section is mainly based on the documents: [Assessing the assessment- Development and use of quality criteria for competence Assessment Programmes](#), Baartman, 2008; [Code of practice for the assurance of academic quality and standards in higher education](#): Assessment of students, 2006; [Assessment and classroom learning](#), [Develop the theory of formative assessment](#), Black & Wiliam, 1998, 2009).

⁴ Depending on the phase of education, the focus of training objectives may vary. However, an overall objective must reflect the competences needed in the future workplace.

2. Criteria: Authenticity

Basic standard:

- The content of an assessment task **must** represent real-life problems of the assessed subject domain.
- Assessment tasks **must** cover all knowledge, skills and attitudes of the required competence for students' future careers.

Quality Development:

- Knowledge, skills and attitudes **should** be assessed in an integrated way, as they are used as an integrated whole in a job situation;
- The assessment tasks **should** be as realistic as possible and as close as possible to the ones that can be encountered in an occupational area;

Annotations:

“Assessment tasks **should** be realistic” (i.e., “tasks are authentic”) is important for the construct validity and for the impact of assessment on student learning or competency development (the latter is also called consequential validity, which will be discussed in more detail in the next criteria: educational consequence). Construct validity of an assessment is related to whether an assessment measures what it is supposed to measure. With respect to competency assessment this means that tasks must appropriately reflect the competency that needs to be assessed and that the content of an assessment involves authentic tasks that represent real-life problems of the knowledge domain assessed. However, as Gulikers (2006) pointed out, this does not mean that every assessment task should be very complex (even though most authentic problems are complex, involving multi-disciplinary, ill-structure, and having multiple possible solutions). Real-life problems can also be simple and well-structured with one correct answer and requiring only one discipline.

The following criterion *educational consequence* (p.4) is to ensure the consequential validity of assessment (describes the intended and unintended effects of assessment on learning and/or teaching).

Although this criterion was not explicitly discussed, participants from Nepal had recognized its importance. Their institutional action plan is related to this criterion (see the action plan of [BPKIHS-Nepal](#)).

3. Criteria: Cognitive Complexity

In the LINQED workshop, Eduardo Suarez Barrientos (UMSS, Bolivia) illustrated this criterion by using two types of assessment as examples (i.e., multiple choice questions and short answer question). In [the framework](#) he and his colleague proposed, very detailed rules were suggested for developing assessment items that assess higher-order skills and ensure a proper ratio of the items that cover essential and important learning content. In line with Eduardo's proposal, many institutions had included this criterion in their action plan (e.g., [UGM, Indonesia](#), [BPKIHS-Nepal](#), [UMSS, Bolivia](#))

Basic standard:

- The assessment programme **must** comprise methods that measure students' different cognitive levels (e.g., [Bloom's Taxonomy](#)).

Quality development:

- An assessment task, depending on the phase of education, **should** elicit the thinking processes⁵ used by practitioners to solve (complex) problems in their occupational field.

Annotations:

Assessment should not only focus on the product, but also on the thinking processes: how and why did students act and make choices during their work on a task. Thus, assessment should be designed to prompt students to provide a rationale for their decisions when performing an assessment task. For instance, the thinking process can be explicitly assessed during the presentations and the criterion-based oral exam by asking students how they tackled a problem.

Note. Judgments regarding the cognitive complexity of an assessment need to start with analyses of the tasks but also need to take into account student familiarity with the problems and how students attempt to solve them. It is incorrect to assume that open-ended scientific problems will necessarily require the use of more complex cognitive processes by students (Linn, Baker & Dunbar, 1991). In early 1987 the National Academy of Education warned of the dangers if the cognitive processes are not revealed during assessment:

“It is all too easy to think of higher-order skills as involving only difficult subject matter as, for example, learning calculus. Yet one can memorize the formulas for derivatives just as easily as those for computing areas of various geometric shapes, while remaining equally confused about the overall goals of both activities.” (Linn, Baker, & Dunbar, 1991, p. 54)

⁵ The examples provided by University of Queensland may give an idea on how we design assessments that require students to explicit their reasoning (For details see [Appendix C](#)).

4. Criteria: Educational Consequences

Basic standard:

In the LINQED workshop, Ganes Retno Rahayu (UGM, Indonesia) highlighted this criteria while introducing [Crooks et. al.s' assessment design cycle](#). When reflecting on an assessment programme as a whole (i.e., thinking all assessment methods are used), this question should be asked "Are positive consequences achieved? Is there any serious negative impact?"

- Assessments **must** stimulate a deep study approach aiming at understanding and application.
- Assessments **must** stimulate the development of professional competencies.
- Assessments **must** emphasise formative in-training methods and *constructive feedback* (i.e., focusing on the task⁶, given regularly and on time⁷, and being specific to the task⁸).

Quality development:

The following practices **should** be encouraged (QAA: Assessment of students, 2006):

- The use of peer assessed activities during formal teaching sessions where students, either in pairs or groups, comment constructively on one another's work.
- The use of self-reflective or other types of student self-assessment counts.
- Involving, for example, employers, patients or clients in providing part of the feedback to students on their performance.
- Setting assessment tasks such as extended assignments that involve students researching a topic and producing work based on their research.
- Opportunities should be provided to students to immediately use the feedback they receive to close the gap between current and desired performance, especially in the case of planned assignments.

Annotations:

"Students' learning is guided by the assessment to come and the objectives being assessed become the students' learning objectives" (Elton, 2002). For instance, if assessment assesses memory learning, then students tend to engage in memory learning, which may be unintended (LTSN: Assessment in Universities: a critical review of research, 2002). Hence the educational consequence of assessment (positive vs. negative) relies on its nature.

⁶ Literature on assessment indicated that task-involving evaluation is more effective than ego-involving evaluation, to the extent that even the giving of praise can have a negative effect on low-achievers.

⁷ Timing is important: students benefit from feedback on their work at a time when they will be able to use it and are most likely to take notice of it, for example, during a module rather than at the end.

⁸ Feedback should give assessment criteria, so that students are very clear on what was and will be expected of them. Specific comments on errors and suggestions for strategies provide diagnostic information for students to readjust learning direction and attention distribution. The suggestions for strategies also aim at helping students to adopt better (efficient/proper) learning strategies.

Feedback of assessment can ensure a positive effect of assessment on student learning by informing the student where and how their learning and performance can be improved. Therefore, it is important to design a 'feedback loop' into assessment tasks so that students can apply formative feedback (from instructors, peers or others) to improve their performance in the next assessment.

Note: one should be aware that the effectiveness of feedback depends on several detailed features of its quality (see footnotes 4, 5 & 6), and not on its mere existence or absence.

Peer-evaluation is an especially useful assessment method because it enables students to understand assessment criteria and deepens their learning in several ways, including:

- learning from the way others have approached an assessment task (structure, content, analysis) and
- learning through assessing someone else's work, which encourages them to evaluate and benchmark their own performance and improve it.

In addition, it is often the case that students, who are only interested in the mark or grade, frequently ignore feedback on summative assignments. The following are some specific strategies suggested to counter this tendency, helping students to use external feedback to regulate and close the performance gap (Nicol & Macfarlane-Dick, 2006):

- provide feedback on work in progress and increase opportunities for resubmission;
- introduce two-stage assignments where feedback on stage one helps to improve stage two;
- students can be supplied with written feedback on summative assignments and the mark or grade withheld until after a period of reflection;
- specifically provide some 'action points' together with the normal feedback to students or involve students in groups to identify their own action points in class after they read the feedback on their assignments.

*One of the key concepts that became crystal clear for all LINQED workshop participants is that **assessment must be used for learning.***

“How can assessment support and enable learning?” [The session](#) organised by Lai Jiang (ITM) on formative assessment opened the door for considering the roles assessment may play for promoting learning besides its summative function. Several suggestions given by participants (for examples see LINQED workshop outcomes) were formulated concerning the formative use of different assessment methods.

5. Criteria: Fitness for Self-assessment

Basic standard:

- Assessment **must** stimulate self-regulated learning.

Quality development:

- Various strategies **should** be used to make sure that students understand the learning goals, understand the assessment criteria, and have the opportunity to reflect on their work.

Annotations:

An ultimate goal of education is to prepare students to be self-regulated⁹ for learning throughout their lives. One effective way to develop self-regulation in students is to provide them with opportunities to practise regulating aspects of their own learning and to reflect on that practice. Self-assessment tasks are an effective way of achieving this, as are activities that encourage reflection on learning progress. To promote effective self-assessment, one must clarify and share learning intentions and criteria with students so that they are able and willing to apply them to their work, and in making judgements about how their work relates to the standards (Boud, 1986).

Various ways exist to create opportunities for students to effectively reflect on their study. For instance (QAA: Assessment of students, 2006):

- integrating external and internal feedback: self-assessment with integrated tutor feedback¹⁰ can help students identify and correct more errors (the ones that students were not aware of)
- enabling students to experience a range of assessment methods that take individual learning needs into account and, where appropriate, encouraging them to reflect on and synthesise learning from different parts of their programme;
- where oral examinations take place, ensuring that opportunities are available for a student to practise and receive constructive feedback, and that the practice and feedback are timed to enable students to refine their work and, if necessary, to further develop the personal skills needed to present their arguments effectively;
- including students in the evaluation of assessment practices.

⁹ Self-regulated learning is an active constructive process whereby learners set goals for their learning and monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features of the environment. (Pintrich & Zusho, 2002, p. 64).

¹⁰ More information can be found in [appendix D](#) on how to provide good feedback (what & when).

6. Criteria: Fairness

It is important that consistent approaches are used to process marks for judgement /rewarding certificates across all subjects. These approaches should support the key principles of fairness to all students and maintaining academic standards. With regard to this criterion, few institutions have drafted action plans (for an example see the action plan of [ITM, Belgium](#)).

Basic standard:

- Clear and published criteria for marking **must** be developed.
- Students **must** be assessed using published criteria, regulations and procedures which are applied consistently.

Quality development:

- where possible, **should** not rely on the judgements of single examiners;
- require and enable students to demonstrate their capabilities and achievements within each module or programme by offering a diversity of assessment;
- **should** take into account all the possible consequences of examination regulations;
- **should** have clear regulations covering student absence, illness and other mitigating circumstances;
- **should** ensure that assessments are conducted securely in accordance with the institution's stated procedures;
- **should** be subjected to administrative verification checks to ensure the accuracy of the procedures.

Annotations:

“Some unfairness may occur through slips, errors and misunderstandings of participants who genuinely believe they are acting honestly: examples include differences in the understanding of the criteria for assessment, prejudice on the part of assessors, mistakes in requests for evidence and unrecognised personal difficulties (illness, lack of access to facilities, conflicting demands for the candidate's time and so on). The thorough briefing of administrators, assessors and candidates on the regulations is a crucial component of the assessment system” (Monk, 2005).

It is important that assessment is designed to recognise student achievement, including exceptional ability. Other than in pass/fail assessments, grading criteria can be used to differentiate between students' performance.

The *Quality Assurance Agency for Higher Education* provided practical guidelines on how to establish transparent and fair mechanisms for marking and moderating marks. Details can be found in [Appendix E](#).

Finally, it is crucial to note that fairness not only refers to the aforementioned administrative issues. Essentially, fairness also implies that **every student gets equal opportunities to demonstrate his/her learning**. From this perspective, Suskie (2000) gave the following suggestions:

1. Have clearly stated learning outcomes and share them with your students.
2. Match your assessment to what you teach and vice versa.
3. Use many different measures and many different kinds of measures.
4. Help students learn how to do the assessment task.
5. Engage and encourage your students.
6. Interpret assessment results appropriately.
7. Evaluate the outcomes of your assessments.

7. Criteria: Transparency

Assessment can help students identify problems and point them in the direction students need to work toward but only when they understand the criteria and standards of the assessment. One way to do this is to develop clear rubrics for assessment and communicate them to students. A good example of a rubric for thesis evaluation can be found in [Appendix E](#). In several institutions' action plans developing rubrics has been set as one of the target goals for quality assurance (for examples see action plans of [ISP-Ecuador](#) & [School of public health, MU, Uganda](#)).

Basic standard:

- Assessments **must** be carried out by people who understand the role of assessment in the progression of students towards the achievement of the knowledge and skills associated with the intended qualification.
- The communication on the purpose of the assessment with students **must** be thorough—making sure that they have a good understanding of the criteria (i.e., they know what they are expected to demonstrate), and hence have the opportunity to reflect on their work.

Quality development:

- Students **should** be clearly informed about the assessment strategy being used for their programme, what examinations or other assessment methods they will be subjected to, what will be expected of them, and the criteria that will be applied to the assessment of their performance.
- External controlling agencies **should** be able to get a clear picture of how a CAP is developed and carried out.

Annotations:

Clarifying and sharing learning intentions and criteria with learners is considered to be essential for encouraging student to use assessment for learning. This is because students can only achieve learning goals if they understand those goals, assume some ownership of them, and can assess progress (Sadler, 1989; Black & Wiliam, 1998). Strategies that have proved effective in clarifying criteria, standards and goals were given by Nicol & Macfarlane-Dick (2006). (For details see Appendix G).

8. Criteria: Meaningfulness

The action plans of several LINQED partner institutes explicitly state that the results of assessment for learning and teaching should be used (for examples see action plans of [BPKIHS-Nepal](#)).

Basic standard:

- For learners, the assessment **must** get students to deal with meaningful problems¹¹ that provide worthwhile learning experiences.
- For teachers and employers, the assessments **must** be meaningful in terms of the requirements of the future job.

Quality development:

- The evidence is evoked (i.e. assessment results). This **should** be interpreted in terms of learning needs and used to make adjustments to meet those learning needs better.
- Teachers **should** use the cumulative evidence generated from assessments to acknowledge students' levels of understanding and skill and adapt their teaching accordingly.
- The evaluation of assessment methods **should** include an evaluation of how they promote training and learning.

Annotations:

Various strategies exist for teachers to generate and collate quality information about students' learning. Here are some examples (Nicol & Macfarlane-Dick, 2006).

- Using variants of the one-minute paper—questions that are posed to students before a teaching session begins, and responded to at the end of the session (e.g., What was the most important argument in this lecture? What question remains uppermost in your mind now at the end of this teaching session?). These strategies can be adapted to any classroom situation or discipline. Moreover, they help to develop important meta-cognitive skills such as the ability to think holistically and to identify gaps in understanding.
- Having students request the feedback they would like when they make an assignment submission (e.g., on a pro forma with published criteria).
- Having students identify where they are having difficulties when they hand in assessed work.
- Asking students in groups to identify a question worth asking, based on prior study, which they would like to explore for a short time at the beginning of the next tutorial.

It is up to the individual institutions to determine the frequency and regularity with which the evaluation of assessment practice is conducted. This might appropriately take place as part of an annual monitoring process or be integrated with internal institutional periodic reviews.

9 & 10 Criteria: Reproducibility & Comparability

¹¹ The University of Queensland provided some concrete examples in the document "assessment for learning" which shows how to formulate meaningful assessment questions. You can find these examples in [Appendix C](#).

Basic standard:

- The final (high-stake) decisions about students **must** be based on multiple assessors, multiple occasions, multiple contexts, and multiple assessment methods.
- Institutions **must** have transparent and fair mechanisms¹² for marking and for moderating marks¹³.

Quality development:

- The conditions under which the assessment is carried out **should**, as much as possible, be the same for all learners and scoring should be consistent.
- The use of internal moderation **should** be encouraged to ensure the consistency of the scoring procedure.
- The reliability and validity of assessment methods **should** be documented and evaluated.
- A set of assessment tasks that are consistent with respect to key features of interest **should** be applied.
- The different stages of training **should** be recorded in a training logbook. An appeal mechanism concerning assessment results **should** be established and, when necessary, a second opinion, change of trainer/supervisor or supplementary training **should** be arranged.

Annotations:

The purpose of an assessment is not a performance in one specific situation observed by one assessor but should enable the assessor to draw more general conclusions about a learner's competences. Reproducibility includes the idea of multiple human judgment and the necessity of adequate sampling of tasks. Comparability of assessment emphasises the common key feature of interest across various assessment methods, the assessment tasks, criteria and working conditions which ensures consistent scoring.

Gandes Retno Rahayu (UGM, Indonesia) introduced the [Crooks et al' assessment design cycle](#), and drew our attention to these two criteria. The central questions related to reproducibility is "Is biased interpretation or explanation avoided?" "Are there sufficient tasks to be assessed?" As to comparability, we need to ask ourselves "Are parts of the target domain assessed or given appropriate Weight?" "Is biased interpretation or explanation avoided?" "Is there an appropriate standard?"

¹² The Quality Assurance Agency for Higher Education provides examples on how to establish such transparent and fair mechanisms (for details see [Appendix E](#)).

¹³ For an example on how to carry out moderation, see the document "[moderation](#)"

11 & 12 Criteria: Cost-Effectiveness & Acceptability

Basic standard:

- All stakeholders (students, teachers and employees) should approve of the assessment criteria and how the CAP is carried out.
- Investments in time and resources on assessment must be justified by the positive effects of competence assessment, such as improvements in learning and teaching.

Quality development:

- Development of assessment principles, methods and the number of examinations in accordance with changes in educational objectives, learning goals and methods.
- To be practical, especially for large-scale assessments, ways **should** be found to keep the costs of performance-based assessments at an acceptable level.

Annotations:

CAPs are generally more complex than traditional tests and may be more difficult to carry out. During the design process of a new assessment, enough attention must be paid to its feasibility. One way to ensure the feasibility is to involve all the stakeholders and take their wishes and possibilities into account.

The cost (time, resources) for carrying out CAP for both students and assessors should be justified by the benefits of the CAP. When assessors and students are asked to invest more time and energy into assessment work, it should be always made clear to them that these expenditures are worth the investment of resources. In this sense, these criteria are also related to [transparency criteria](#).

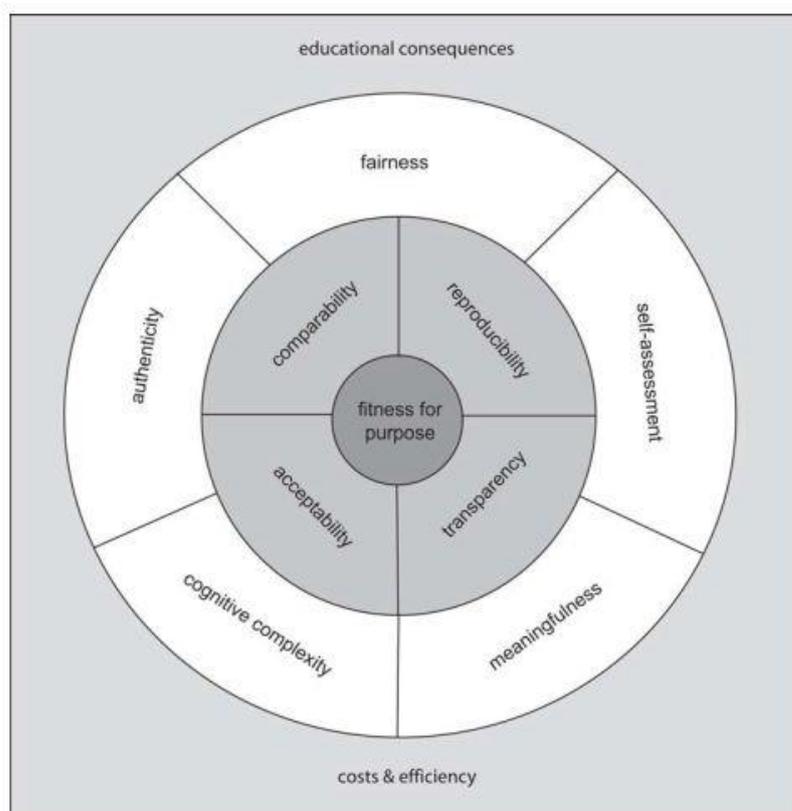
Framework of quality for Competence Assessment Programmes (CAPs)

The assessment of competencies is very complex, mainly due to the fact that a competency comprises a complex integration of knowledge, skills and attitudes (Van Merriënboer, Van der Klink, & Hendriks, 2002). Because assessing competencies is such a complex endeavour, it seems impossible to assess a competency using only one assessment method.

Therefore, a new trend of assessment method advocates the integration of different assessment methods into a Competence Assessment Programme (CAP), in which newer forms of assessment can be used in combination with more classical methods.

The framework of CAPs asserts that the well-known and widely used classical psychometric quality criteria of validity and reliability are not suitable for evaluating the quality of CAPs. Based on a literature study a framework of ten quality criteria for CAPs is proposed. The definitions of the criteria can be found below.

For a more elaborate description of the framework please read **Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & Van der Vleuten, C. P. M. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. Educational Research Review, 2, 114-129.**



Fitness for purpose emphasises the alignment among standards, curriculum, instruction, and assessment. The assessment goals and methods used should be compatible with the educational goals.

Appendix A

Authenticity relates to the degree of resemblance of a CAP to one's future professional life. A CAP should assess those competences needed in the future workplace.

The authors distinguish five dimensions that can influence authenticity: the assessment task, the physical context, the social context, the assessment result or form, and the assessment criteria.

Cognitive complexity resembles authenticity in the sense that it also relates to a person's future professional life, but it focuses more directly on the fact that assessment tasks should also reflect the presence of the cognitive skills needed (Hambleton, 1996; Linn et al., 1991). An assessment task, depending on the phase of education, should elicit the thinking processes used by practitioners to solve complex problems in their occupational field. In this respect, Hambleton remarks that the use of performance assessments is no guarantee that higher cognitive skills are indeed being measured. Therefore, this should always be thoroughly investigated.

Educational consequences are mentioned as a criterion for competence assessment by many authors (Dierick & Dochy, 2001; Linn et al., 1991; Messick, 1994; Schuwirth & Van der Vleuten, 2004) and pertains to the effects a CAP has on learning and instruction. A collection of evidence is needed about the intended and unintended, positive and negative effects of the assessment on how teachers and learners view the goals of education and adjust their learning activities accordingly. This criterion is also related to effects such as washback (Alderson & Wall, 1993; Prodromou, 1995).

Fitness for self-assessment means that CAPs should stimulate self-regulated learning. They should include specific methods to foster this learning such as practice in self-assessment and giving and receiving feedback.

Fairness specifies that a CAP should not show bias to certain groups of learners and reflect the knowledge, skills and attitudes of the competence at stake, excluding irrelevant variance (Hambleton, 1996; Linn et al., 1991). Possible causes of bias are improper adjustment to the educational level of the learners or tasks containing cultural aspects that not all learners are familiar with.

Transparency relates to whether a CAP is clear and understandable to all participants. Learners should know the scoring criteria, who the assessors are, and what the purpose of the assessment is. As a possible indication of the transparency of an assessment, Hambleton (1996) suggests to check whether learners can judge themselves and other learners as accurately as trained assessors.

Meaningfulness implies the fact that a CAP should have a significant value for both teachers and learners (Hambleton, 1996; Messick, 1994), to which the importance could be added in the mind of future employers. A possible way to increase meaningfulness is to involve learners in the (development of the) assessment process. McDowell (1995) underlined that

Appendix A

for learners to perceive an assessment as meaningful, they need to perceive a link between the assessment task and their personal interests. An assessment might also become more valuable to learners when they themselves can determine when they are ready to take the assessment and can thus gain most from it.

Reproducibility of decisions relates to the fact that the decisions made on the basis of the results of a CAP should not depend on the assessor or specific assessment circumstances. This does not mean that a CAP must be objective. In many new assessments, assessors subjectively judge the performance of learners. Therefore, multiple assessors, assessment tasks and situations should be combined.

Comparability addresses the fact that a CAP should be conducted in a consistent and responsible way. The conditions under which the assessment is carried out should be, as much as possible, comparable for all learners and scoring should occur in a consistent way, using the same criteria for all learners .

Costs and efficiency are especially important because CAPs are generally more complex than traditional tests and more difficult to carry out. This criterion relates to the time and resources needed to develop and carry out the CAP, compared to the benefits. Evidence needs to be found that the additional investments in time and resources are justified by the positive effects, such as improvements in learning and teaching.

Acceptability is described as “all stakeholders should approve of the assessment criteria and the way the CAP is carried out. They can have faith in the CAP’s quality.”

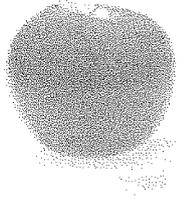
Appendix B

Elaboration of the six levels of thinking in Bloom's taxonomy ¹⁴						
1 Remembering <i>Can the student RECALL information?</i>	2 Understanding <i>Can the student EXPLAIN ideas or concepts?</i>		3 Applying <i>Can the student USE the new knowledge in another familiar situation?</i>	4 Analysing <i>Can the student DIFFERENTIATE between and RELATE constituent parts?</i>	5 Evaluating <i>Can the student JUSTIFY an opinion, decision or course of action?</i>	6 Creating <i>Can the student GENERATE new products, ideas or ways of viewing things?</i>
<p>Recognising Locating knowledge in memory that is consistent with presented material. <u>Synonyms</u></p> <ul style="list-style-type: none"> Identifying Finding Selecting Indicating <p>Recalling Retrieving relevant knowledge from long-term memory. <u>Synonyms</u></p> <ul style="list-style-type: none"> Retrieving Naming Reproducing Recounting 	<p>Interpreting Changing from one form of representation to another <u>Synonyms:</u></p> <ul style="list-style-type: none"> Paraphrasing Translating Representing Clarifying Converting Rewriting Restating Expressing <p>Exemplifying Finding a specific example or illustration of a concept or principle <u>Synonyms</u></p> <ul style="list-style-type: none"> Instantiating Illustrating... Representing Giving examples of Showing <p>Classifying Determining that something belongs to a category (e.g.,</p>	<p>Summarising Drawing a logical conclusion from presented information. <u>Synonyms</u></p> <ul style="list-style-type: none"> Abstracting Generalising Outlining Precising <p>Inferring Abstracting a general theme or major point <u>Synonyms</u></p> <ul style="list-style-type: none"> Extrapolating Interpolating Predicting Concluding Extending Generalising <p>Comparing Detecting correspondences between two ideas, objects, etc. <u>Synonyms</u></p> <ul style="list-style-type: none"> Contrasting Matching 	<p>Executing Applying knowledge (often procedural) to a routine task. <u>Synonyms</u></p> <ul style="list-style-type: none"> Carrying out Measuring Constructing Demonstrating Computing Calculating Manipulating Operating Preparing Producing Drawing up Practising <p>Implementing Applying knowledge (often procedural) to a non-routine task. <u>Synonyms</u></p> <ul style="list-style-type: none"> Using Estimating Predicting Solving 	<p>Differentiating Distinguishing relevant from irrelevant parts or important from unimportant parts of presented material. <u>Synonyms</u></p> <ul style="list-style-type: none"> Discriminating Selecting Focusing Distinguishing between Separating (Sub)dividing Examining Relating <p>Organising Determining how elements fit or function within a structure. <u>Synonyms</u></p> <ul style="list-style-type: none"> Outlining Structuring Integrating (Re)arranging Categorising Ordering Deriving 	<p>Checking Detecting inconsistencies or fallacies within a process or product. Determining whether a process or product has internal consistency. <u>Synonyms</u></p> <ul style="list-style-type: none"> Testing Detecting Monitoring Concluding Assessing Appraising Discriminating Determining <p>Critiquing Detecting the appropriateness of a procedure for a given task or problem. <u>Synonyms</u></p> <ul style="list-style-type: none"> Judging Questioning Justifying Defending Discussing 	<p>Generating Coming up with alternatives or hypotheses based on criteria <u>Synonyms</u></p> <ul style="list-style-type: none"> Hypothesising Proposing Developing Engendering Synthesising Providing options <p>Planning Devising a procedure for accomplishing a task. <u>Synonyms</u></p> <ul style="list-style-type: none"> Designing Formulating Combining Compiling Devising Revising Putting together Suggesting <p>Producing Inventing a product <u>Synonyms</u></p>

¹⁴ From "Revised Bloom's Taxonomy" retrieved 20 May, 2005 from <http://rite.ed.unt.edu/oz-teachernet/index.php?module=ContentExpress&func=display&ceid=29> and *Using Learning Outcomes to Design a Course and Assess Learning Outcomes*. http://www.hlst.heacademy.ac.uk/guide/current_practice/Learning.html and Moon, J. Linking Levels, Learning Outcomes and Assessment Criteria. Retrieved 30 May, 2007, from http://www.see-educoop.net/education_in/pdf/edinburgh-moon-oth-enl-t02.pdf *Assessment resource developed by Dr Clair Hughes (TEDI/The University of Queensland)*

Appendix B

	concept or principle). <u>Synonyms</u> <ul style="list-style-type: none"> • Categorising • Subsuming • Organising 	<ul style="list-style-type: none"> • Mapping <p>Explaining Constructing a cause-and-effect model of a system.</p> <u>Synonyms</u> <ul style="list-style-type: none"> • Elucidating • Constructing models 	<ul style="list-style-type: none"> • Changing • Discovering • Explaining how • Verifying • Finding 	<p>Attributing Determining the point of view, bias, values, or intent underlying presented material.</p> <u>Synonyms</u> <ul style="list-style-type: none"> • Deconstructing • Comparing • Contrasting • Diagnosing 	<ul style="list-style-type: none"> • Criticising • Arguing • Including • Rating • Ranking • Valuing 	<ul style="list-style-type: none"> • (Re)constructing • Composing • Modifying • Altering • Building • Enlarging
--	---	--	---	--	---	---



What kinds of assessment tasks are likely to promote learning for understanding?

The straightforward answer is “those kinds of assessment which require a deep understanding of the work”. Learning for understanding implies outcomes at the upper end of the SOLO taxonomy or of Bloom’s taxonomy. To succeed in assessment which promotes such outcomes, the student will need to have a deep understanding. If the students *expect* to be assessed in this way then they will tend to learn in the appropriate manner (Marton and Säljö, 1976b; Newble and Jaeger, 1983).

Understanding as the ability to apply theory

In any area where theory can be applied to specific instances, students might demonstrate some understanding by applying the theory in straightforward cases. Generally this will require that they have a basic understanding of the theory (what various theoretical terms mean in specific situations, for example). To test a deeper understanding of the theory, try

setting assessment exercises which require the use of subtle aspects of the theory. For example:

- Why is the sky blue on a sunny day? Explain by reference to appropriate physical theories. Your answer will be judged not only on the correctness of the reason, but also by the cogency of the theoretical justification.
- Is ‘existence’ a predicate? Please explain your answer with appropriate examples.

What is ‘difficult’ or ‘subtle’ will depend on how the issues have been explored in class or in learning materials. What requires reasoning in some contexts will require only recall in others. For example, “Why is the sky blue...?” may have been treated as an example in class. A more probing question might then be, “In Brisbane at the end of some sunny days there is a spectacular sunset, but not at the end of others. Why might this be so?”

Understanding as the ability to make learning personally meaningful

Another way in which students can demonstrate understanding of an area (and one which actively promotes 'deep', personally meaningful learning) is to ask students to write about the area in their own words, to give examples from their own, personal experience, or to relate theory or more 'academic' learning to their personal experience:

- Give three examples of the inappropriate use of stereotypes

in advertising. Justify your answer by reference to the theory of ...

- Interview three people of your acquaintance who are at least 50 years old. What are their feelings about the new 'deeming' policies?...
- Keep a record of **everything** you eat for three days. Analyse your diet using the model presented in lectures and the textbook. Given your life style, is your diet adequate for your needs? Please explain your answer.

...ask students to write about the area in their own words, to give examples from their own, personal experience...

Assessment and learning

Assessment directly contributes to learning both by clarifying what is desirable or required and by closing a feedback loop between students' learning efforts and their achievements. Telling students what is required will assist them to direct their learning efforts. Students will learn many things from a course, some of which you may not have intended (Biggs, 1996; Cobb, 1996; Terwel, 1999), but if they have clear learning goals they are more likely to achieve the intended outcomes. Similarly, learning is enhanced if students are helped to see what they have (and have not) learned — that is, if they receive feedback.

Feedback

Good feedback on assessed work tells the student four things:

- 1 what are the good or successful features of the assessed work;
- 2 what are the poor or less successful features of the assessed work;
- 3 how the student can improve in this piece of work; and
- 4 how the student might do better work in future.

Good feedback is also *timely*. Provided too soon it may stop the students themselves reflecting on their work; provided too late it may no longer be salient to the student and thus may be ignored.

When might feedback be 'too soon'?

'Too soon' might apply to the physical location of the feedback or it might apply to the time at which it is provided. For example, writing

overall feedback on an essay on the first page of the essay may encourage some students to read only that overall feedback and not to read more detailed comments written throughout the text of the essay. Giving a mark or grade is a crude kind of feedback, satisfying none of the criteria for good feedback, but feedback nonetheless. Giving a grade before feedback in the form of a comment may result in students looking only at the grade.

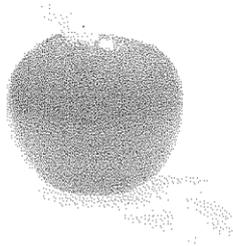
When might feedback be 'too late'?

'Too late' generally will apply to the time at which the feedback is supplied. Feedback is too late when it no longer is salient to the student. Thus, if the student feels that the task is complete and no future task will have much in common with this one, then the feedback is unlikely to have any effect. If the task was carried out so long ago that the student can no longer remember clearly and meaningfully what they did, then the feedback is not likely to have much effect.

Helping students use feedback

Think about how students might use feedback on their assessment exercises. Consider building into your plan for the course ways to help them do so. For example:

- Have some assessment which is purely for feedback. Such work might be set exercises from a text book (especially if the text contains worked solutions for problems), brief tests in class, moderate size tasks done by groups, brief presentations (for example, in tutorials). The feedback need not necessarily come from the teacher. It can be generated by the student him or herself (especially if clear



Give students the option (with or without penalty) of resubmitting a piece of assessable work or reattempting a test in the light of feedback.

marking criteria are supplied or negotiated), or by a peer of the student, or by a group of students. Carefully structured feedback sheets can be of great benefit here.

- More simply, have students submit a piece of assessable work in stages, allowing them to benefit from the feedback on each stage to inform their work on the balance (e.g. in a project, submit a plan; submit results; submit final report including results, as well as analysis and conclusions).
- Give students the option (with or without penalty) of resubmitting a piece of assessable work or reattempting a test in the light of feedback. This can be seen as *presubmitting* the assigned work or *preattempting* the test—in this case *all* students would have the option of accepting their first result or trying to improve on it.
- In the session where students have their work returned, ask the students to pair up. Each student in a pair should read the feedback on the other student’s work. The students should then suggest to each other ways that similar work might be improved in future.
- Each student, within a couple of days of receiving feedback, should write on an index card “As a result of the feedback on this assignment one thing I am going to do is...”. Then various things are possible, depending on the nature of the assignment. For example, if the feedback might lead to students actually doing some kind of work, students could exchange the cards and contact each other at an appropriate time to offer encouragement. If the process is aimed principally at encouraging reflection, then each student could be urged to look at an appropriate time (for example, when they are about to start work on the *next* assignment) at the card they themselves wrote.
- Ask students to submit the plan for a project for assessment and feedback. You might ask them to include, as part of their eventual final report, a section which explains how what they did was modified as a result of the feedback, and why.

Code for practice-Marking and grading (Sections)¹⁶

Institutions have transparent and fair mechanisms for marking and for moderating marks.

Publicising and using clear assessment criteria and, where appropriate, marking schemes, are key factors in assuring that marking is carried out fairly and consistently across all subjects. An important principle is that students and markers are aware of, and understand, the assessment criteria and/or schemes that will be used to mark each assessment task.

Precepts and explanations relating to external scrutiny and moderation of marking are included in Appendix F, Code of practice - External examining.

Internal moderation is important in assuring that examiners apply assessment criteria consistently, and that there is a shared understanding of the academic standards students are expected to achieve. Evidence of moderation is an important feature of internal procedures. Different methods of internal moderation are more or less appropriate for particular situations. In some circumstances, moderation may be limited to sampling a representative number of scripts from a cohort of students, perhaps with an emphasis on borderline cases. In other cases, moderation may involve double, or second, marking.

Some of the factors institutions may wish to take into account in developing policies and procedures on marking and moderation include:

- How to ensure that marking and grading at faculty, school and departmental level are appropriate and comparable. Institutional level guidance can suggest the circumstances in which it might be preferable either to give precise numerical marks or to use grades or bands of marks when assessing student work.
- The need for clear guidance about how borderline marks or grades are defined and treated.
- The circumstances in which anonymous marking is appropriate and when it is either not practical or inappropriate (for example in work-based assessment, or in the performing arts). Advice about where in the process anonymity ends is normally included in institutional guidance on this topic.
- When double or second marking should be used and what approach should be taken, for example, whether or not the assessment of students' second marker normally has access to the first marker's comments and/or marks and highlighting the importance of demonstrating that double or second marking has taken place.
- The methods to be used when assessments from larger groups are sampled by internal or external examiners.
- The processes governing and recording any internal moderation and verification of marks and the procedure to be followed when an internal or external moderator disagrees with the original marks.
- The usefulness of undertaking an analysis of marking and marking trends to facilitate comparisons and provide evidence on standards. Some institutions may find it appropriate to incorporate such analysis in annual monitoring processes.

Institutions publicise and implement clear rules and regulations for progressing from one stage of a programme to another and for qualifying for an award.

¹⁶The Quality Assurance Agency for Higher Education (September 2006). [Code of practice for the assurance of academic quality and standards in higher education](#). Section 6: Assessment of students (pp.16).

It is important that students, staff and examiners are aware of the ways in which assessment results will be used, including how they affect progression within a programme and their contribution to the overall programme outcome.

The results required to pass each stage and to progress to the next stage of a programme (where appropriate) need to be clearly stated and explained to students at the beginning of the programme. The purpose of this is to ensure that students understand the impact of individual marks on their ability to progress and ultimately to complete the programme.

In modular systems, it is important to make clear the effect that passing or failing an individual module will have on the student's eligibility to take other modules, as well as the overall implications for progression and completion. For each taught programme or group of programmes, institutions may wish to consider putting in place fair and easily understood procedures for combining individual marks and/or grades to come to a final programme mark. These procedures will need to be transparent and easily accessible to students, staff and examiners and be previously evaluated by the institution to assure their reliability and validity.

Consistent approaches to progression and combining marks for awards across an institution support the key principles of fairness to all students and maintaining academic standards. Flexibility at subject level may be appropriate to reflect different discipline needs and marking conventions, including those in practice-based subjects.

This might include allowing faculties, schools or departments to decide which assessment marks can contribute to a final degree mark. Such flexibility can often be accommodated within the overarching rules set by the institution, but where this is not possible, approval at institutional level of any variation helps to promote fairness. Consistency of treatment in the ways outlined above should enable an institution to recognise comparable levels of student achievement across disciplines in similar ways.

Guidance at institutional and programme levels which includes references to the following can support the implementation of this precept:

- The extent to which a student's overall success in a programme can include failure in part of the programme when permitted by institutional rules and regulations. In modular systems, guidance can helpfully distinguish between core and optional modules and include details about any modules that must be passed to meet PSRB requirements. It is important to ensure that students receiving an award have achieved or exceeded the learning outcomes for the programme.
- Defining which marks contribute to the decision about whether a student receives an award.
- On what basis re-takes or re-submissions can occur, making clear the number and timing permitted and the accompanying procedures; for example, re-sitting examinations; re-submitting a dissertation; repeating a work-based or other type of practical assessment; or repeating an oral examination.

the rules for deferring or not completing an assessment, together with any special assessment conditions or penalties that may apply, including any restriction on the marks, grades or levels of award that can be obtained on the basis of retaken or deferred assessments. It is helpful if such rules cover a wide range of circumstances, including any progression permitted or awards conferred because of a student's absence due to illness or other personal circumstances.

Appendix F

Rubric used in UCL London

CIHD MSc essay marking criteria (December 2010)

Distinction = 70%+

Pass = 50%+

	80%+	70-79%	60%-69%	50%-59%	40%-49%	0-39%
Clarity of argument Clarity and structure of argument	Exceptionally clear and persuasive argument. Includes clear introduction, logical progression, excellent signposting and conclusion which follows from the points made. In addition, shows a high level of sophistication in developing a complex argument from different strands.	Clearly and persuasively argued. Includes clear introduction, is well-signposted, shows logical progression, and has a conclusion which follows from the points made.	Well-argued, but some lack of coherence and clarity. Introduction present, mostly logical progression, conclusion which follows from some points.	Outlines of argument are present, but not developed in a compelling manner. Some problems with structure, insufficient introduction and conclusion, lack of signposting. Adequately written, but meaning sometimes let down by poor expression.	Little attempt to develop a coherent argument. Lack of structure: introduction or conclusion missing, little or no signposting, points are not made in a logical sequence. Meaning often unclear, with ideas expressed in a confusing way.	No attempt to develop a coherent argument. No structure: introduction or conclusion missing. Argument based on anecdotes and generalisations only. Writing highly unclearly and ideas very poorly expressed.
Use of evidence breadth of sources and evidence; sources and evidence used to illustrate argument; synthesis of knowledge; understanding of limitations of data / literature	Demonstrates a systematic approach to the literature, using up-to-date sources and evidence well beyond what is expected. Sources and evidence chosen are highly appropriate to the argument. Shows an exceptional ability to synthesise knowledge. Shows clear understanding of limitations of data / literature and develops alternative approaches.	Demonstrates a systematic approach to the literature, using a wide range of up-to-date sources and evidence. Sources and evidence chosen are highly appropriate to the argument. Demonstrates an excellent ability to synthesise knowledge. Shows clear understanding of limitations of data / literature and suggests alternative approaches.	Demonstrates use of an adequate range of sources and evidence. Sources and evidence are generally appropriate to the argument. Good attempt to synthesise different pieces of knowledge. Shows some understanding of limitations of data / literature and makes an adequate attempt to show the implications.	Use of sources and evidence is somewhat below what might be expected. Argument relies on only a few sources. Argument shows a more limited ability to synthesise relevant knowledge. Limitations of data / literature acknowledged, but little attempt to show implications of this.	Poor range of or inappropriate sources and evidence, and not up-to-date. Sources and evidence show only tangential relation to argument. Knowledge not well-synthesised. Little attempt to acknowledge. Limitations of data / literature and little or no awareness of the implications.	Extremely limited range of or inappropriate sources and evidence, with no relevance to argument. Little or no attempt at synthesising knowledge. Limitations of data ignored.

Appendix F

	80%+	70-79%	60%-69%	50%-59%	40%-49%	0-39%
Grasp of theory ability to describe and apply theory relevant to the argument.	Shows evidence of being able to advance existing theoretical frameworks.	Excellent grasp of key theories and their application.	Demonstrates adequate knowledge of relevant theories; minor deficiencies in description and application.	Demonstrates awareness of relevant theories, but not always described and applied accurately.	Relevant theories inaccurately described, and little attempt at application.	Little or no theoretical content.
Critical thinking and originality (a) critical / evaluative / analytic approach taken to sources. Balance. (b) originality and independence of thinking in tackling question.	Demonstrates outstanding ability to identify strengths and weaknesses of sources. Appropriately balanced when undertaking critical evaluation. Takes a highly original and independent approach to the discussion.	Shows excellent ability to identify strengths and weaknesses of sources. Appropriately balanced when undertaking critical evaluation. Shows flashes of originality and independence in a highly competent account.	Shows evidence of critical evaluation applied to sources in an appropriate and relevant way. Indications of some lack of balance in critical evaluation. Shows evidence of originality and independence though not necessarily well-tailored to answering the question.	Shows some limited critical ability; some errors in identifying strengths and weaknesses of sources. Strong lack of balance in critical evaluation. Little originality and independence demonstrated, or what is tangential to the question.	Little sign of a critical approach being taken. Key errors in identifying strengths and weaknesses of sources. Highly unbalanced in presentation of different sources. No originality or independence demonstrated.	No sign of a critical approach being taken. Highly unbalanced in presentation of different sources. No originality or independence demonstrated.
Presentation: language and referencing	Negligible grammatical and spelling deficiencies. Reference list comprehensive and accurately presented, according to CIHD guidelines.	Negligible grammatical and spelling deficiencies. Reference list comprehensive and accurately presented, according to CIHD guidelines.	Only minor grammatical and spelling deficiencies. Well-presented. Reference list comprehensive and accurately presented, according to CIHD guidelines.	Greater care required in spelling and grammar. Reference list not complete and/or has some inconsistencies in format that are not in line with CIHD guidelines.	Significant spelling and grammatical mistakes. Guidelines for referencing not adhered to, and several references missing.	Significant spelling and grammatical mistakes. Missing or totally inadequate reference list. No attempt to adhere to CIHD guidelines.

Notes:

1. 80%+. In general the assessment demonstrates attributes beyond the quality expected at this level of qualification.
2. This template is a guide for markers and gives an idea of the level of performance expected from students. Note that poor performance in one area may be compensated by better performance in another.
3. Where a task requires written or spoken English, the quality of the English forms part of the assessment. In the case of foreign speakers of English, discretion is exercised for occasional mistakes (e.g. non-native word order or incorrect use of articles), but all students are penalised for incomprehensible English.

Clarifying learning intentions and criteria

(i.e., what are learning goals and what is good performance)

HOW

One way of clarifying task requirements (goals/criteria/standards) is to provide students with written 'examples' of performance and statements that describe assessment criteria and/or the standards that define different levels of achievement. Examples are important. This is because many studies have shown that most criteria for academic tasks are complex, multidimensional (Sadler, 1989) and difficult to articulate; they are often 'tacit' and unarticulated in the mind of the teacher. One or few examples can make what is required explicit and define a valid standard against which students can compare their work.

Other strategies that have proved effective in clarifying criteria, standards and goals include:

- Providing better definitions of requirements using carefully constructed criteria sheets and performance-level definitions;
- Increasing discussion and reflection about criteria and standards in class (e.g. before an assignment);
- Involving students in assessment exercises where they mark or comment on other students' work in relation to defined criteria and standards;
- Workshops where students in collaboration with the teacher devise or negotiate their own assessment criteria for a piece of their work

References:

- American Association for Higher Education (1991). *Nine principles of good practice for assessing student learning*. Sterling, VA: Stylus.
- Baartman (2008). *Assessing the assessment' Development and use of quality criteria for Competence Assessment Programmes*. Doctoral dissertation.
- Black, P., & Wiliam, D. (1998). *Assessment and classroom learning*. *Assessment in Education*, 5(1), 7-71.
- Black, P., & Wiliam, D. (2003). "In Praise of Educational Research": formative assessment. *British Educational Research Journal*, 29(5), 623-37.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Boud, D. (1986). *Implementing student self-assessment*. Higher Education Research and Development Society of Australia.
- Bresciani, M. J. (2003). Expert-driven assessment: Making it meaningful. Educause Center for Applied Research (ECAR) Research Bulletin, 2003(21).
- Brown, G. J. B., & Pendlebury, M. (1997). *Assessing Student Learning in Higher Education*. London: Routledge.
- Cizek, G.J. (1997). Learning, achievement, and assessment: Constructs at a Crossroads. In G.D. Phye (Ed.), *Handbook of Classroom assessment: Learning, achievement, and Adjustment*. (pp. 1-32). San Diego: Academic Press.
- Council of Regional Accrediting Commissions (2004). *Regional accreditation and student learning: A guide for institutions and evaluators*. Atlanta: Southern Association of Colleges & Schools.
- Crooks, T. J., & Kane, M. T. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*, 3(3), 265.
- Crooks, T.J. (1988) The impact of classroom evaluation practices on students, *Review of Educational Research*, 58, 438-481.
- Elton, L. (2002). Good Assessment Practice. Retrieved in August 2011 at <http://www.materials.ac.uk/events/assessmentgp.pdf>
- Elton, L. & Johnston, B. (2002). Assessment in Universities: a critical review of research. Generic Centre of the learning and teaching support network (LTSN). <http://eprints.soton.ac.uk/59244/1/59244.pdf>
- Gulikers, J. T. M. (2006). *Authenticity is in the eye of beholder*. Doctoral dissertation.

- Huba, M. E., & Freed, J. E. (2000). Applying principles of good practice in learner-centered assessment. In *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning* (pp. 65-90). Needham Heights, MA: Allyn & Bacon.
- Isaacs, G. (2001). *Teaching & learning in Higher Education Series: Assessment for learning*. The University of Queensland.
- Linn, R.L., Baker, E., & Dunbar, S. B. (1991). [Complex, performance-based assessment: Expectations and validation criteria](#). *Educational Researcher*, 16, 1-21.
- Monk, J. (2005). Assessment of work-based reports: an analysis of assessment frameworks. In: *International Conference on Engineering Education and Research*, 1-5 March 2005, Tainan, Taiwan.
- National Council on Measurement in Education (1995). [Code of professional responsibilities in educational measurement](#). Washington, DC: Author.
- Nicol, D. J. & Macfarlane-Dick, D. (2006) 'Formative assessment and self-regulated learning: a model and seven principles of good feedback practice'. *Studies in Higher Education*, 31(2), 199-218.
- Pintrich, P. R. & Zusho, A. (2002) Student motivation and self-regulated learning in the college classroom, in: J. C. Smart & W.G. Tierney (Eds) *Higher Education: handbook of theory and research* (vol. XVII) (New York, Agathon Press).
- Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science* 18, 119-44.
- Segers, M., Dierick, S., & Dochy, F. (2001). Quality standards for new modes of assessment. An exploratory study of the consequential validity of the OverAll test. *European Journal of Psychology of Education*, 16(4), 569-586.
- Suskie, L. (2000). *Fair assessment practices: Giving students equitable opportunities to demonstrate learning*. AAHE Bulletin, 52(9), 7-9.
- Suskie, L. (2004). *What are good assessment practices? In Assessing Student Learning: A Common Sense Guide*. Bolton, MA: Anker.
- The Quality Assurance Agency for Higher Education (2006). [Code of practice for the assurance of academic quality and standards in higher education. Section 6: Assessment of students](#).