

From: Biggs & Tang (2011)

Table 11.3 Example of criteria (rubrics) for grading a declarative ILO

Scale score	Marginal	Adequate			Good			Excellent		
	D 1.0	C- 1.7	C 2.0	C+ 2.3	B- 2.7	B 3.0	B+ 3.3	A- 3.7	A 4.0	A+ 4.3
ILO										
Explain	Able to identify and briefly write about limited points Very little evidence of using these points to provide reasoning to why they are interrelated	Able to identify a number of relevant points with some details Use these points to provide a fair reasoning or causality No evidence of a comprehensive overview of reasoning or causality			Able to identify a full range of relevant points with details Supported by relevant literature Points are organized to provide a comprehensive and cohesive reasoning or causality			As in 'good' but provides views on possible alternative causes and/or results depending on changes of conditions Able to link current reasoning to situations in real-life professional contexts		

could be used as formative feedback to the student or summatively. If the latter, as this is only one ILO out of five or so for a given course, we will need to state a final grade for that course and to calculate a student's overall grade for the ILO. The 'scale score' is actually taken from one university's conversion from grade to GPA-type scale: notice that again as in Table 11.1 (p. 240), the gap between grades is greater than the gap within grades in terms of scale score. When the final result has been calculated, we can convert to GPA score using Table 11.2 (p. 241), as before. It would in fact be most meaningful if on the student's transcript all the assessments of all the course ILOs were retained rather than overall GPA. This would give a potential employer, or the dean of graduate school, for instance, a detailed account of the student's strengths and weaknesses.

In practice, students at present want to know 'How did I do on that midterm assignment?' rather than 'How did I do on the "explain" ILO?' To some extent, then, it will be necessary to assess both the task itself to give student feedback, as well as the ILOs it may address. In time, however, when students and the public generally become used to outcomes-based teaching and learning it may well be that a profile of grades on the ILOs will become perfectly meaningful to all.

Assessing in large classes

Many teachers see no alternative to the final exam and the MCQ when assessing large classes. Using varied assessment tasks for higher level ILOs, especially those addressing functioning knowledge, is seen by many teachers as impractical in large classes.

However, it need not be thus. Of course, assessing the projects, assignments and portfolios of 400 students between the end of semester and submission of grades to the faculty board of examiners may be logistically and humanly impossible. But there are alternatives.

Cloze tests

These were originally designed to assess reading comprehension. Every seventh (or so) word in a passage is deleted and the student has to fill in the space with the correct word or a synonym. A text is chosen that can only be understood if the topic under discussion is understood, rather like the gobblet (p. 244). The omitted words are essential for making sense of the passage.

Concept maps

We have seen concept maps as a teaching/learning activity (pp. 141–2). They can also be used for assessment. They are particularly useful for giving an overview of the course. They need not take a long time to prepare and the

teacher can tell at a glance if a student has an impoverished knowledge structure relating to the topic or a rich one.

Goblets

Goblets are significant chunks of content with which the student should be familiar and to which the student has to respond (Brown and Knight 1994). They could be a paragraph from a novel or of a standard text, a brief passage of music, a Venn diagram, an archaeological artefact, a photograph (a goblet, an engine part) and so on. The student's task is to identify the goblet, explain its context, say why it is important, what it reminds them of or whatever else you would like them to comment on.

Goblets should access a bigger picture, unlike short answers that are sufficient unto themselves. That big picture is the target, not the goblet itself. Brown and Knight point out that three goblets can be completed in the time it takes one essay exam question, so that to an extent you can assess both coverage and depth. They could assess either declarative or functioning knowledge. Box 11.6 gives two examples of goblets for assessing ILOs for declarative knowledge.

Box 11.6 Two examples of goblets

For assessing the ILOs 'identify' and 'explain' in Law

A piece of legislation is provided. Students are asked to identify the context of the legislation, explain its importance, and its possible impact on a current legal crisis.

For assessing the ILOs 'analyse' and 'recommend' in Occupational Health and Safety

Students are given a photograph of a section of a construction site in function.

Students are asked to analyse the situation, identify any potential health and safety hazards and recommend measures to improve the situation.

Group assessment

Group assessment is appealing in large classes. With four students per assessment task, you get to assess almost a quarter the number you would otherwise. But there are problems, particularly of plagiarism and its equivalent, freeloading. It is necessary to be very careful about who does what in the project, which is where peer-assessment helps, and that each student obtains an overview of the whole task, not just of their particular contribution, for

example by writing a reflective report on how well each thinks they have achieved the ILOs (pp. 254–5).

Letter-to-a-friend

This is written by the student to a friend, imaginary or real, who is supposedly thinking of enrolling in the course in the following year (Trigwell and Prosser 1990). These letters are about a page in length and are written and assessed in a few minutes. The student should reflect on the course and report on it as it affects them. Letters tend to be either multistructural or relational, occasionally extended abstract. Multistructural letters are simply lists of course content, a rehash of the course outline. Good responses provide integrated accounts of how the topics fit together and form a useful whole (relational), while the best describe a change in personal perspective as a result of studying the course (extended abstract). Letter-to-a-friend also provides a useful source of feedback to the teacher on aspects of the course. Like the concept map, letters supplement more fine-grained tasks with an overview of the course.

Minute paper

The minute paper appeared as a TLA for large class teaching and as a learning activity and as feedback for the teacher (p. 143). It can just as easily be used summarively for grading purposes, but if so, the students should be told first as their strategies will be different. An obvious advantage is that the three-minute essay can be answered and assessed in, er, three minutes.

Peer- and self-assessment

Peer- and self-assessment can slash the teacher's assessment load quite drastically, even when conventional assessments such as exam or assignment are used (p. 229). An additional benefit is that self- and peer-assessment are particularly well suited for assessing functioning knowledge and values ILOs such as teamwork and cooperation, because such assessments are what are required in real life.

Let us recap the advantages:

- 1 Self- and peer-assessment give the students first-hand, active involvement with the criteria for good learning.
- 2 Students learn how to select good evidence.
- 3 Judging whether a performance or product meets given criteria is vital for effective professional action.

It is important that these educational justifications are made clear to the students, not only because the rationale for all teaching and assessing decisions should be transparent, but because it is necessary to get the students on side. A common belief is that assessment is the *teacher's* responsibility and

some students resent being required to do the teacher's dirty work (Brew 1999). Peer-assessment can also be stressful to some students (Pope 2001). It should be noted too that good students under-assess themselves, compared to how their peers would rate them, while poor students over-assess themselves (Lejk and Wyll 2001b).

How well do self- and peer-assessments agree with teacher assessments? Falchikov and Boud (1989), reviewing 57 studies, found that agreement was greatest with advanced students, least in introductory courses; and in convergent content subjects, such as science, medicine and engineering, rather than in arts and social science. Good agreement requires explicit criteria of assessment and discussion and training in using them (Fox 1989).

As an operational rule of thumb, Boud (1986) suggests that if self- and/or peer-assessments agree within a specified range, whether expressed as a qualitative grade or as a number of marks, the higher grade is best awarded (collusion can be mitigated by spot checking). He estimates this procedure can cut the teacher's load by at least one-third. Gibbs (1999) cut marking time for the teacher by 18 hours a week by using peer-assessment, while summative marks increased by 20% simply because peer-assessment is itself a powerful TLA.

Random assessment

One way of ensuring that students are motivated to put effort into a series of ATs is to use random assessment. In Gibbs (1999), 25 reports through the year were required, but as each was worth only a trivial 1%, the quality was poor. When the requirements were changed, so that students still submitted 25 reports as a condition for sitting the final exam, but only four reports selected at random were marked, two benefits resulted. The students worked consistently throughout the term and submitted 25 good reports, and the teacher's marking load was one-sixth of what it had previously been.

Short-answer examinations

These are answered in note form. This format is useful for getting at factual material, such as interpreting diagrams, charts and tables, but is limited in addressing main ideas and themes. The examiner is usually after something quite specific, and in practice operates more like the objective format than the essay (Biggs 1973; Scouler 1996). However, it has advantages over the standard MCQ in that it is less susceptible to test-taking strategies: the answer can't be worked out by elimination, it requires active recall rather than just recognition and it is easier to construct but not as easy to score.

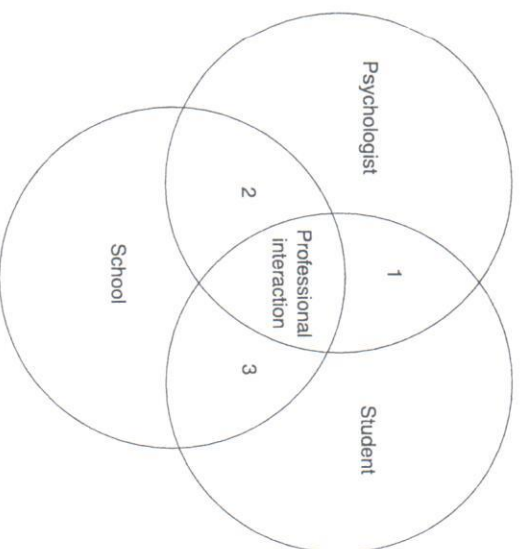
Venn diagrams

A simple form of concept map, where the boundary of a concept is expressed in a circle or ellipse, and interrelations between concepts expressed by the intersection or overlap of the circles (see Box 11.7). Venn diagrams, like

concept maps, are very economical ways of expressing relationships. They can be used for teaching purposes, in conveying relationships to students, and for assessment purposes, so that students may convey their ways of seeing relationships between concepts. Getting students to draw and briefly explain their own Venns, or to interpret those presented, can be done quickly, where the target of understanding is relationships between ideas.

Box 11.7 represents an item for an educational psychologist course II.O relating to professional interaction. There are three domains: psychologist, student and school, with each of which the psychologist has to interact at various times. For the student to be able to explain examples of the interactions (1) through (3) would indicate a high level of understanding of the

Box 11.7 An example of a Venn diagram



Write a brief sentence explaining an interaction that would occur in the sites in relation to professional interactions.

- 1 _____
- 2 _____
- 3 _____

This item is easily adapted to other areas by using different labels in each circle.

psychologist's role. This item could be adapted to virtually any situation: just label the circles differently. Task 11.2 asks you to think about precisely that.

For those who need to assess declarative ILOs for a large class, Task 11.2 asks you to design a gobbet, a concept map or a Venn diagram as an assessment task.

Task 11.2 Design a gobbet, concept map or Venn diagram for assessing declarative ILOs for large classes

Consider some the ILOs for declarative knowledge in your course, design a gobbet, a concept map or a Venn diagram as an assessment task.

For the gobbet

ILOs to be assessed:

1 The content information you are going to give to your students.

2 Questions or tasks your students need to answer or complete in relation to the content information.

For the concept map or Venn diagram

ILOs to be assessed:

1 The concepts or information to be given to the students (if appropriate).

2 Questions or tasks your students need to answer or complete in relation to the concept map or Venn diagram.

3 Presentation of the concept map or Venn diagram.

Educational technology (ET) can help in rapid assessment, as well as in assessment generally (see Chapter 12).

Before we end this chapter, Task 11.3 is an exercise on designing assessment tasks for declarative knowledge of your course.

Task 11.3 Design an assessment task or tasks for one of your course declarative ILOs

Select one ILO relating to declarative knowledge of your course and design assessment task(s) that will appropriately assess this ILO. To help you check the alignment between the task(s) and the ILO, identify what the students are required to do in order to complete the assessment task(s). The task requirements should be aligned to the ILO.

Course ILO: _____

Number of students in the course: _____

Assessment task	Student activities to complete the task
1	
2	

Now double-check whether the student activities are aligned to the verb(s) nominated in the respective course ILO.

After designing the task(s), you will need to write the grading criteria for either the ILO or for each of the tasks.

Summary and conclusions

Designing assessment tasks

In designing assessment tasks (ATs) there are several things to bear in mind. Clear assessment criteria or rubrics need to be established for each task, or for the ILO(s) each AT is meant to address. It is useful to think of ATs as a *source of evidence* of a student's achievement of any ILO. You can have one source of evidence or several, just as one task may provide evidence on more than one ILO. However, the ATs have to be *manageable*, both by students in terms of both time and resources in performing them and by staff in assessing students' performances.

Assessing and grading using extended prose

Declarative knowledge is typically assessed by writing answers to set questions or in essay-type or objective formats. The format chosen matters because it can affect the way students react to the task. Writing is either in the timed and invigilated 'exam' or unrestricted, as in the typical essay assignment. The stress typically felt in the 'exam' situation produces its own distortion in the quality of work done, especially by the Roberts of this world. A particular problem with assessing extended writing is the lack of reliability between assessors. Several suggestions are made to improve this.

Objective formats of assessment

The lack of reliability of assessing essays, plus the time they take to assess, has led many teachers to use objective formats, particularly the MCQ. The major problem with the MCQ, however, is that it is not at all suited to addressing high-level outcomes and that it is prone to encouraging strategic rather than knowledge-driven responding. An exception is the ordered-outcome format, which encourages students to target higher rather than lower level items.

Assessing and grading the task

Once the AT is aligned to the ILO(s) it is meant to address, the question becomes how, operationally, is the student's performance assessed? Is it assessed against the task or against the ILO(s) the task is meant to address? Teachers and students are used to task assessment and that is what many teachers will continue to do. There are three ways of task assessment: quantitatively, as has been the case traditionally; by assessing and grading the task analytically, addressing the task components using rubrics for each component; or by assessing the task as a whole and grading qualitatively.

Assessing and grading the ILO

The most logical, and operationally the simplest, way of assessing is by using the evidence gained from the various tasks directly to assess the ILO itself, by using rubrics designed for each ILO. The main objection to this method is simply that teachers and students are not yet used to it.

Assessing in large classes

Assessing large classes need not mean exams and MCQs. Cloze tests, concept maps, gobbers, group assessment, letter-to-a-friend, minute paper, peer-and

self-assessment, random assessment, and Venn diagrams and are all ways in which rapid assessments of higher order ILOs, involving little time in either sitting or assessing, can be made.

*Further reading**Computer scoring of essays*

Rudner, L. and Gagne, P. (2001) An overview of three approaches to scoring written essays by computer, *Practical Assessment, Research and Evaluation*, 7: 26. <http://parconline.net/getvn.asp?v=7&n=26> (accessed 2 February 2011).
 Shermin, M. and Burstein, J. (eds) (2003) *Automated Essay scoring: A Cross-disciplinary Perspective*. Hillsdale, NJ: Erlbaum.

Much of the background and enrichment material for this chapter is the same as for the next. Please refer to the Further Reading section in Chapter 12.