learning & development guide
information and education services division
6
This guide follows CALT Guides 1-3

Assessing by Portfolio

John Biggs, Visiting Fellow

Abstract

Common approaches to assessment focus on the congruence between students' answers to set questions and what the teacher has predetermined as optimal or correct. Not only do such formats typically elicit lower level learning outcomes than we want students to learn, they miss outcomes we may not have foreseen but which are nonetheless important and relevant. In assessing by portfolio (AP), the students select examples of their learning as evidence that the quality of learning nominated in the objectives has actually taken place. The items can be specified by the teacher, or chosen by the student, or both.

How is assessing by portfolio different?

When I stand in front of a class, I don't see stupid or unteachable learners, but boxes of treasures waiting for us to open.

(An in-service teacher education student, University of Hong Kong)

I received this gem in an assignment on assessing learning. It makes a very important point both about the nature of learning, and of assessment.

In the traditional assessment format, the teacher is usually the one who decides what questions to ask, what the evidence for learning is to be, and who judges the extent to which these criteria have been met. All too often this means requiring the student to retell in some form or another what the teachers has already told the student. It is a false dialogue: the questioner knows the answers to the questions that the questioner has put, and the respondent knows that the questioner knows! Such a format of assessment must miss the more important aims of tertiary courses. It is also very boring, for students to do and for teachers to evaluate.

In terms of the treasure metaphor, something like this occurs:

Teacher: How many diamonds have you got?

Student: I don't have any diamonds.

Teacher: Then you fail!

Student: But you didn't ask me about my jade.

Learners can construct treasure of all kinds, not just in diamonds. If teachers only ask questions for which they have a limited range of acceptable answers, they will miss the jade: the treasure they didn't know existed because they didn't ask.

In assessing by portfolio (AP), students must be absolutely clear as to what is required in the teaching objectives. They then decide what "treasures" to put on show by selecting items for the portfolio that they think provides evidence for the required learnings. Since students see the assessment as the "actual" curriculum (Ramsden, 1992: 187), what the portfolio does is embed our wanted curriculum in the assessment: our curriculum and the students' curriculum are now the same.

AP is like a job application (Moss, 1994). As the candidate has to convince a committee of the strengths on offer, so the student has to convince the teacher of what has been learned, how effectively, and its relevance to the unit or course in question. A key requirement of AP is a justification for the items selected: how each item contributes to what objective, and what is the overall picture of learning given by the portfolio. AP allows for a broad range of important learning:

1. *conventional learning outcomes*. These include high levels of understanding of concepts, theories, who-said-what, algorithms, the usual kinds of declarative knowledge taught in tertiary courses.
2. *unforeseen outcomes*. Any rich teaching/learning activity contains more opportunities for learning than the teacher might have bargained for. Particularly at the higher levels of understanding students can show a generativity in their thinking that is missed by conventional assessment probes.
3. *reflective or metacognitive awareness of their own learning*. The justification for the items selected require the students to self-evaluate in a very important sense. They have to see their own learning as experts would see it, explaining why this particular item meets the criterion for "highly distinguished" performance. (Which is not to say that students self-*grade* their work; that is an entirely separate issue you may or may not see as appropriate).

The parallel between AP and everyday professional decision-making extends beyond job applications. Being able to interpret external specifications that fall within one's range of expertise, and to judge whether one's work meets those specifications, are high order competencies that universities should be fostering; likewise the flexibility of AP in providing the opportunity to how one's work may exceed existing specifications and show how the job can be done in ways the client may not have foreseen.

Implementing AP

Following are the necessary steps in implementing AP:

1. ***Make it quite clear in the teaching objectives what the evidence for good learning may be.*** The objectives should be available to students at the beginning of the semester. In a graded course (High Distinction, D, Cr, P, etc.), as opposed to Pass/Fail, this statement would also contain what is required for HD, for D, and so on down the line. In Pass/Fail it is only necessary to specify what is required for P. Detailed examples follow.
2. ***State the requirements for the portfolio.*** These need to be made very clear, as otherwise some students are likely to dump barrowloads on you. Again, it's like real-life. If someone's job application is the length of a novel, that person has problems, the most serious being lack of the very judgment you want to assess.
   - *number of items*. In a semester long unit, four items is about the limit.
   - *approximate size of each item*. The total portfolio should not be much longer than a project or assignment you would normally set. I stipulate no more than 1,500 words for any one item, but that of course depends on the nature of the item. Some items, such as concept maps or other diagrams, require less than a page.
   - *a list of sample items,* but emphasise that it is better to show some creativity by going outside that list as long as the items are relevant. Items should not be repetitive, making the same point in different ways. For item suggestions, and what they are supposed to be assessing, see below.
   - *any compulsory items?* In my courses (in teacher education) I usually prescribe a journal, leaving the other items to student choice, but obviously the subject area will determine what sorts of items are most suitable.
   - *source of items*. It could be possible in some course designs for a portfolio to be built up from previous assessment tasks in that course.

For instance, in some problem-based courses students will be continually providing inputs, often on a pass/fail basis, over a year or two years. The final evaluation could then comprise - in toto or in part - samples of the best work students think they have done to date.

- *what are the items supposed to be getting at?* Obviously, your teaching objectives, but are they best addressed as a package, or as a list of separate items? The answer to this question determines how the portfolio is to be graded. A major virtue of the portfolio is that it is most usefully seen as a package. The student is in effect saying: "This is what I got out of your class. I have learned these things, and as a result my thinking is changed, as you can see. I am now able to..." If their package can show "how my thinking has changed" they have learned well indeed.

3. ***Decide how the portfolio is to be graded.*** The following questions need to be addressed:

   (a) *the assessment of individual items.* Pass/fail? "Marks" (but see below)? Qualitatively graded?

   (b) *the assessment of the portfolio as a whole* (the "package"). How are the assessments of the individual items to be combined?

   (c) *are there any other assessments for the unit apart from the portfolio?* If not, the (b) assessment above would stand as the final grade for the unit. However, you may decide to use the portfolio to assess open-ended learning outcomes, and a conventional assessment perhaps to establish "coverage" of basics. In that event, you will need to decide how to combine the two results.

These matters are elaborated below, with examples from my own use of AP.

Setting clear objectives

The objectives for the unit in question must be written in such a way students are quite clear as to what they are required demonstrate. When we teach, it is nearly always for "understanding". However, understanding takes many forms. In most tertiary objectives, it is appropriate to conceive of understanding as *performative* (Perkins & Blythe, 1993); that is, when students "really" understand something, they *behave differently* in that content area. Much learning at tertiary level, particularly in professional courses, is about getting students to behave differently in the sense of making theoretically informed decisions; it is not so much about declaring who-said-what and who-did-that, particularly in higher years. Much traditional assessment is however of such declarative knowledge: we ask questions, they write answers.

Teachers need to specify in what ways their students should perform differently if they have a deep understanding of what has been taught. For example, the essence of professional judgment is being able to make decisions that go beyond the textbook: new cases, novel applications, ill-defined problems, working from first principles. The key is not the specific tasks, but the *level of cognitive activities* required to make the decisions in the appropriate contexts. Expressing objectives in the form of *verbs* is helpful; they make it clear what the student is supposed to do (Biggs, 1996; and see below). The student can then choose productions that put these activities to use in the relevant content area.

Objectives and the grading scheme

Let us assume we are using the High Distinction, Distinction, Credit, Pass categories. There are four ways to go, with variations within each:

1. *You decide what qualities of learning students would need to demonstrate in their work to warrant a Pass, what beyond that warrants a Credit, then a Distinction, then a High Distinction; the portfolio is then judged as a package against these criteria.* In the following example, verbs (italicised) were used to designate levels of understanding: the levels, and their relation to grading categories, is structured in line with the SOLO Taxonomy as outlined in Biggs (1992). This allows us to state quite clearly the qualities of understanding required in handling the course content (in this case educational psychology) at each grading level. In this example I didn't use HD, D, etc but the letter-grades A, B, C, D. The principles are the same. The highest level of understanding that a portfolio consistently displayed (the rationale is crucial in deciding this) was graded according to that level, as follows:

   A: Metacognitive understanding, students able to use the taught content in order to *reflect* on

their own teaching, *evaluate* their decisions made in the classroom in terms of theory, and thereby *improve* their decision-making and practice; that is, to *formulate* a personal theory of teaching that demonstrably drives decision-making and practice, and to *generate* new approaches to teaching on the basis of taught principles and content.

B: Students can *apply* course content, and *recognise* good and poor applications of principles. They "understand" in that course content is used as a theory of teaching that drives action.

C: Students understand declaratively, in that they can demonstrate that they can *discuss* the important topics content meaningfully, and in context. However, they don't transfer or apply knowledge easily. (A strong cue is a "depersonalised" way of discussing content, however meaningfully; it has not been taken on board, it does not drive action).

D: Basic understandings, evidence of some effort in the acquisition of terminology; higher level understanding offset by some misunderstandings.

F: Worse than D.

2. *You decide the categories, as above, but more stress is placed on individual items.* This is easier to grade, but there is less emphasis on the shape of the whole. You could stipulate that for a Pass, they need to submit at a satisfactory level for all items, but the rationale does not take in the course as a whole. For a Credit, they need to show "excellence" in one other item (whatever "excellent" may mean in your subject at the year level in question: mastery of difficult concepts, originality,... that needs to be spelled out) and to have a rationale that shows how the items "hang together". A Distinction might then require "excellent" in two (or three) items and a rationale that shows some insight into the subject structure; High Distinction excellent in all four and a clear, original or personal, depending on the subject, picture of their learning in the rationale. In this case, they submit the best they can and you judge what criteria they have met. Here the rationale holds the items together.

3. *You decide what tasks need to be passed for what grades.* For instance, a P is obtained if students submit a journal, and it meets the specifications you have laid down (length, content, evidence of reflection, etc). That is the baseline. After that, they can if they wish submit more items. Thus for Credit, journal plus (say) a couple of book reviews addressing important topics in the unit but not set texts. Distinction is journal, book reviews, and an essay on, say, "What are the most important points I have gained from this Unit". High Distinction could be all the preceding plus a highly original piece of cognate work.

4. *Each item is allocated so many marks,* and the final grade depends on the sum: HD 85+, D...etc. This so-called "analytic" method of grading has real problems in this context, as discussed in the next section.

Important Note. So far, the grading is entirely criterion-referenced. **If the students do what is required, at the level required, that's it. If 80% meet the requirement for HD, then congratulations, they and you have done a great job!** (as long as you are *sure* your criteria are realistic...). The naming of the categories carries some baggage in this respect. To be "highly distinguished" is to be rare; the category "HD" therefore carries the suggestion that there shouldn't be many of them, in a way that "A" does not. Educationally, of course, the more students that genuinely achieve the higher level objectives, the better.

Holistic or analytic methods of assessing APs

While it might seem that AP is ideal for analytic marking, as in (4) above, this is against the spirit of AP and of referencing assessment to qualitative criteria. The backwash is negative, telling students to select mark-productive items, rather than to see how topics in the unit complement each other to create an overall learning experience. Thus, when a student disputes a grade that has been marked quantitatively, the discussion becomes an unseemly quibble over a mark here, a mark there. When a student disputes a grade that has been assessed qualitatively, the discussion becomes a seminar on the nature of learning; the justification for the item selection, and the "package" they make, is the topic of the debate. The ultimate focus of AP is thus not directly on individual items, although some will impress mightily in themselves, but ideally on the *package* they make.

In making holistic assessments, the details are not ignored; the question is whether, like the bricks of a building, or the characters in a novel, the specifics are tuned to create an overall structure or impact.

This is a matter of *judgment*, a responsibility that teachers can avoid under the shelter of quantitative assessment, by allocating points and letting the numbers make the big decisions (Moss, 1992). However, if one is competent to teach a topic, one should also be competent enough to assess it properly, by judging whether the whole performance matches the objectives.

Either teacher or student can make the summative judgments of the AP; the student's judgment could be an item, and then it is a separate issue as to whether that judgment is used in the grading or not.

Reporting in percentages

To grade holistically is not say that you can't report in percentages, if required for administrative/clerical reasons. Assessing and reporting are two different exercises. The normal procedure is to assess quantitatively, and report in categories on the basis of the accumulated marks. It is far preferable to reverse the procedure:

1. assess qualitatively into the grading categories, P, Cr, etc.
2. fine grade *within* the category: a bare P would be 51%, an excellent P but lacking the essential qualities of Credit would be low sixties (if Credit begins at 65).

Two judgments are made. A macro-judgment: What sort of learning is displayed here? and a mini-judgment: Which are the best examples and which the worst in this category?

This way you can satisfy the requirements of administrators, the need to sort out students particularly at the upper ends, to feed students their marks, but most important to send them the message: **Shoot for quality** - a category shift means a big bonus in "marks".

Other Issues

Examples of item types (obtained in a teacher education course)

Items may address:

(a) the "declarative" knowledge taught, that is the text-book content, which may be understood at various levels of insight, or

(b) functioning knowledge, or knowledge put to work in solving real life problems, or improved workplace functioning.

Either may be specific to a topic, or address the whole unit. The letter-to-a-friend, developed by Trigwell and Prosser (1990) to evaluate first year Science, for example, is a simple and powerful way of telling whether students see the unit as a bunch of topics, or as a total experience with an impact. Concept maps may address single topics, or the whole course. All items are evaluated in terms of depth of understanding displayed.

Examples of items relating to *understanding of declarative knowledge:*

> book and article reviews, self-set essays, concept maps, set ten objective test items with justifications for the correct response.

Examples relating to *functioning knowledge:*

> *applications:* lesson plans, actual lessons with student evaluations (the latter often led to reflection), test items based on SOLO and other content dealt with, critical incidents of good/bad teaching/learning experiences.
>
> *personal reflection and self-evaluation:* student/colleague evaluation of one's own teaching, diary entries, critical incidents, letter-to-a-friend, VT of a group discussion, later written from the perspective of each.

The possibilities are many, and of course depend on the content area. The item form is not as critical as the way it is interpreted and justified.

Workload

Student workload can be a problem, if a self-inflicted one. In my own experience, students need clear guidance. Some need convincing that often only a page or so may suffice for some items (as in the concept map, the letter-to-a-friend). Other items, such as the journal or some self-selected items, may be very extensive. Again, advice to restrict journals to critical events is necessary. Nevertheless, some portfolios are bulky, others slim. Strict word limits could be imposed, but this may limit the creativeness of some of the students. As always, the resolution of these issues depends on the teaching objectives.

As for *teacher* workload, times for assessing individual APs vary enormously. A concept map, or letter-to-a-friend, can be assessed in 30 seconds. Average for a four-item portfolio would be about half-an-hour for the whole. The time spent in assessing portfolios might be objectively longer than in assessing more conventional formats, but it *seems*

shorter, and more productively spent. Often a student addresses the teacher, asking for advice or to comment on a point arising; the portfolio is a personal activity, and needs responding to accordingly. This makes it far more interesting to assess than a standard essay/assignment.

Student reaction

I have used portfolios at all levels from first year to postgraduate. The reaction at first, particularly among less mature students, is usually negative. One group of first years, at the beginning of the first semester, were appalled at this early stage of their university life to have to face this new monster of AP. They wanted an objective test. So we did a deal. A very simple MCQ, a disciplinary device, really, to make sure they could at least recognize the more important topics addressed. By the end of the semester, the majority much preferred the portfolio.

One large evening class comprised part-time teachers. This was a plus, as every working day was a potential gold-mine of portfolio items. Even so, first reactions were negative, as the following excerpts from a letter-to-a-friend indicate:

> How about the assessment? Aiyaa! ANXIETY! ANXIETY! ANXIETY! I was so puzzled and worried about it when I received the handout on the first meeting.

At the end of the unit, reactions had changed:

> Now I have changed my perception of assessment and I have practised with it. It really works!
>
> (The same student as above)

> I do not see the portfolio as an assignment to be handed in, it's rather a powerful learning tool for the learner himself.

> I found lots of fun (in making my portfolio)... it led me to think about many questions that I never think of...

> What (we are expected) to prepare for the portfolio undoubtedly provide me a chance to reflect on my daily teaching. This would never happen if this module proceeds in the same way as the other modules. I would not be so alert about my own teaching and eager to make changes and improvements.

> Instead of bombing us with lengthy lectures and lecture notes, we have to reflect on our own learning experiences and to respond critically... I feel quite excited as this course is gradually leading me to do something positive to my teaching career and to experience real growth

> All (the teacher) said was "show me the evidence of your learning that has taken place" and we have to ponder, reflect and project the theories we have learnt in our teaching... How brilliant! If it had only been an exam or an essay, we would probably have just repeated his ideas to him and continued to teach the same way as we always do!

Courses and content most suitable for portfolio assessment

AP is possibly best suited for professional courses, and for expressive courses such as Fine Arts where portfolios have been used for years, but it can be used for virtually any course content.

It only requires that students attempt to demonstrate what they have learned in terms of the objectives.

One can keep a portfolio of maths problems, reflections on one's growing understanding of physics, biology, whatever. The essential requirement is that the objectives are stated clearly enough for students to understand what they are supposed to be learning. And isn't that all the time?

Further Reading

Biggs, J.B. (1992). A qualitative approach to grading students. *HERDSA* News, 14 (3), 3-6.

Biggs, J.B. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.

Biggs, J.B. & Collis, K.F. (1982). Evaluating the quality of learning: *The SOLO Taxonomy*. New York: Academic Press.

Moss, P.A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research, 62*, 229-258.

Moss, P.A. (1994). Can there be validity without reliability? *Educational Researcher, 23*(2), 5-12.

Perkins, D., & Blythe, T. (1993). Understanding up front: A performance approach to testing for understanding. Paper presented to Annual Meeting, American Educational Research Association, Atlanta, April.

Ramsden, P. (1992). *Teaching and learning in higher education*. London: Routledge.

Trigwell, K., & Prosser, M. (1990). Using student learning outcome measures in the evaluation of teaching. *Research and Development in Higher Education, 13*, 390-397.

learning & development guide

information and education services division

Other CALT Guides in this series:

1. *Getting the Most out of Assessment*
2. *Evaluating Teaching*
3. *Encouraging Students' Motivation to Learn*
4. *How Well Do We Help Students to Learn?*
5. *Demythologising the Teaching of International Students*

Assessing By Portfolio

Please write below any comments and suggestions you would like us to consider or include in the next edition of this publication.

Please send a copy of your comments and suggestions to:

Isabel Sendlak, Box 1, The Hunter Building.