**Diverse assessment methods and approaches**

**A table adapted from one in Brown, S. and Race, P. ‘Using effective assessment to promote learning’ chapter 5 in *University teaching in focus: a Learning centred approach* edited by Lynne Hunt and Denise Chalmers Routledge 2012**

|  |  |  |  |
| --- | --- | --- | --- |
| **Method** | **Advantages** | **Disadvantages** | **Notes** |
| **Exams**  Traditional unseen, time-constrained exams, which largely use essay style questions. | Regarded as fair, avoiding problems of plagiarism and cheating, and amenable to yielding data which can be handled quantitatively. | Traditional exams only measure ‘what comes out of students’ pens’ which is an inadequate proxy for ‘what’s in their heads’. | For a detailed discussion of the pros and cons of traditional exams and alternative exam approaches, see The lecturer’s toolkit 3rd edition Chapter 2 in Race (2006). |
| **Open-book or open-note exams**  Students can take into the time-constrained exam specified or unspecified texts and notes, so that the questions focus not on memory and recall, but on interpretation and analysis | Can take away the focus from memorizing. | A great deal of time is used in marking exams, often against the clock. |  |
| **Take-away papers**  Students are given, say, a week in which to prepare an answer to a given topic, effectively as a short-term course work assignment | Allows a more normal approach to researching and preparing answers than in traditional exams. | Take away papers can disadvantage students with hectic home lives. |  |
| **Short-answer questions**  Students produce short responses to large numbers of questions, enabling high coverage of topics, with less reliance on elegance of sentence construction and argument. | Moves away from ‘speed of extended writing’ as a necessary skill, and allows a wider range of subject material to be tested. | Some argue (erroneously we believe) that this constitutes ‘dumbing down’ |  |
| **Essays** | Enables students to demonstrate the ability to construct an argument and to write fluently, coherently and at length. | Rarely used in mainstream occupations, so not an authentic means of assessment for most courses.  Readily plagiarisable.  Requires teachers to spend too much time marking.  Students may be graded on their essay technique, rather than their subject mastery. | One of the most heavily used methods of assessment on humanities courses, with least justification. |
| **Reports** | More authentic than essays, as students often enter careers where report-writing is a requirement. | Reports are often based on practical and field work and done collaboratively by students, but are usually required as individual write-ups, so the report-writing itself may end up being assessed, when perhaps the collaboration was more important. | This is the fall-back means of assessment for many courses. |
| **Multiple choice questions with feedback responses** | Excellent for quick testing of factual material. Feedback on correct or incorrect choices can be given instantly, allowing rapid formative feedback on learning. | While suitable for formative purposes, it is much harder to design high-quality multiple-choice questions for summative assessment.  Questions must be piloted extensively to determine facility values and discrimination indices to select which questions are suitable to include in summative tests. | Getting students to design multiple-choice questions *and* feedback responses is an excellent learning activity, and can lead to the development of substantial question banks for formative use. |
| **Sophisticated computer-based tests**  Using a wide range of question types including drop-down menu selection, drag and drop, clicking points on graphs, interpreting data from maps and diagrams etc. | Best suited to large cohorts and multiply presented courses, where it can be exceptionally efficient.  Helps new students build confidence as they self-test on new and familiar material. | Tends to require considerable work to design good computer-aided assessment. When it is bad, it leads only to efficient guesswork. | Needs expertise in question design, subject content, and technology to support it. |
| **Portfolios** | Allows learners to present wide-ranging evidence of achievement, and to show originality and creativity alongside mastery of subject knowledge. | Takes time to mark, and assessment reliability can be quite low as different assessors tend to look for different things when assessing wide-ranging evidence of achievement. | Portfolios can be maintained and show development over a considerable time scale, and can be useful evidence of achievement to show to prospective employers. |
| ***Viva voce* individual oral tests or interviews** | Allows probing questions to check for understanding.  Widely used for high-stakes assessment, such as at doctorate and masters levels.  Regarded as authentic, as many careers and professions may depend on face-to-face skills at answering questions and giving persuasive explanations. | Some candidates can be let-down by nerves in face-to-face ‘grillings’.  Evidence of achievement may be ephemeral, and it is hard to analyse retrospectively, unless recordings have been made.  Difficult to guarantee fairness between candidates, especially when variations in levels of probing occur.  With large cohorts, it can be difficult to timetable oral assessment in ways which prevent ‘leaking’ of questions to forthcoming candidates. | Gordon Joughin (2010) proposes a strong case for increased use of oral assessment as part of a balanced range of assessment methods in present-day contexts.  Much more extensively used in Northern Europe and Scandinavia than in the UkK |
| **Presentations** | Allows candidates to demonstrate oral communication skills alongside subject mastery.  Authentic, as oral presentation skills are often important in future employment.  Peer-assessment can make presentations a better learning experience for all.  Can include assessment of the ability to respond to questions from the audience. | Time-consuming.  May be hard to strike a balance between mastery of content and skills of presentation.  ‘Raising the bar’: expected standards can become higher over a series of presentations.  Unless recorded, presentations are ‘ephemeral’, making it hard to compare a series reliably. | ‘Impression’ marks can be associated with the quality of presentation slides or handout materials. |
| **Posters** | Allow candidates to integrate a range of evidence of achievement in an agreed visual format.  Can be compared with each other in an amenable way.  Lend themselves to peer assessment, and assessment by third parties, for example professionals in the field or employers. | The visual appearance and design of the poster can rate too strongly in the overall assessment, as opposed to the thinking behind the poster.  Some candidates may have better availability of resources (colour, photos and so on), compromising fairness of assessment. | Exhibitions of posters can be kept photographically or online, and used to train future candidates in the preparation of their own posters.  Many researchers in science and engineering subjects use posters in live contexts to demonstrate their research outcomes |
| **Projects** | Can be used for in-depth investigations that develop research skills. (Jenkins & Healey 2012)  Allow candidates to demonstrate originality and creativity. | Reliability of assessment can be compromised when projects are quite different in scope and range.  Tendency for presentation to influence assessment unduly. |  |
| **Case studies** | Increased authenticity because they are ‘real-world’ scenarios.  Allow candidates to demonstrate originality and creativity, especially where candidates construct case studies.  Demonstrate creative problem-solving especially where all candidates work with given case studies. | Written communication may dominate the assessment (unless an element of oral assessment is used as well).  It is impossible to have a range of case studies of exactly equal ‘difficulty’. | A bank of case-studies can be developed, to allow students the opportunity of rehearsal to develop the required skills and competences to do well.  Commercial case studies are also widely available |
| **Simulations** | Can measure a range of practical skills and competences beyond written and oral communication.  Highly authentic when relating to particular professions and contexts. | Take time to design, which is offset by advantages when high numbers of candidates are involved. | Heavily used in professional subjects like aviation and surgery to test practical competence in safe contexts, where computer-based simulations are extensively used. |
| **OSCEs (Objective Structured Clinical Exams)** | Allow candidates to be tested on high-level skills, in authentic contexts (e.g., interpreting x-rays, interviewing patients (actors), interpreting case notes, making diagnoses, deciding on prescriptions etc.).  Can be quite quick. | Designing OSCE scenarios can be time-consuming, but the design can be done by groups of students in prize competitions – the real prize being that some of the scenarios they submit are actually used in their own OSCEs. | OSCEs are highly regarded as authentic, valid and reliable assessment processes in medical and clinical education, but the process can readily be extended to contexts such as business, policing, law etc.. |
| **Reflective journals** | Deepens learning by reflection, and demonstrates analysis, creativity and originality. | Risk being narrative rather than reflective. | When reflective journaling works well, students continually develop their learning on the basis of reflection. |
| **Critical incident accounts** | Allows students the opportunity to choose particular incidents, analyse them in depth, and demonstrate creativity and problem-solving skills.  Can be word-constrained, thereby increasing ‘cut-to-the-chase’ writing or speaking.  Critical incidents are authentic, developing candidates’ skills for real-world problem-solving. | Incidents inevitably have differing levels of difficulty. | Accumulate a bank of critical incidents providing students rehearsal opportunities to develop skills at handling this assessment format.  Try presenting students with pre-determined incidents. Their performance can then be compared. |
| **Assessed seminars** | Ideally, seminars should be prepared and led by students (individually, or in pairs), and should involve peers and tutors as active audience participants.  A series of seminars can give each student the chance to prepare in-depth an element of the curriculum, and present it to the group, and be questioned on it by the group.  The assessment can be high in validity, relating to depth of knowledge, as well as communication skills and the ability to answer probing questions. | It is hard to allocate topics of equal difficulty to a large group of students.  Over a series of seminars, ‘drift’ occurs, as later presenters tend to bring in and use what they’ve learned from earlier seminars, thereby benefiting from earlier presenters.  Difficult to assess audience participation at anything more than a basic level.  Students who have already presented may switch off and failing to contribute or attend later seminars. | Students’ continued participation in a seminar series can be increased if assessment is triangulated by the inclusion in a summative exam of one or more questions relating to things addressed only in the seminar series. |
| **Annotated bibliographies** | A useful way to engage students with the relevant literature, rather than just collecting information on it.  Candidates can demonstrate their depth of study of the sources and the breadth of the source material they have reviewed.  Plagiarism is limited. Although students may choose the same sources, it would be easy to spot identical annotations. | The extent of the literature may mask the depth of thinking *about* the sources.  Avoid this by setting an exact number of sources to be included, and by asking students to include some elements of prioritisation of how well the respective sources measure up to two or three given criteria, alongside their own judgements about the sources. | Annotated bibliographies can be turned into a resource-bank, and issued to future students as a starting place for them to develop their own bibliographies. |
| **In-tray exercises**  For example, in an exam on ward management, students could find on their exam desks no questions, but a set of paperwork for them to study for a while, including logs of patients on the ward, doctors available, other facilities available on the ward etc. Then (say) 20 minutes into the exam, they each received a slip of paper, for example ‘Incident at the airport. The following three patients will arrive in 15 minutes. What will you do?’ The students write down their decisions, based on the information available to them. Other ‘incidents’ are given to them at different stages in the exam. | This kind of assessment is strong on authenticity, as it measures the skills the candidates will need in their careers.  Reliability of assessment is high, as normally there will be ‘best’ choices in terms of the decisions and actions required.  This kind of exam focuses on thinking, rather than merely writing.  Since all the students have the same in-tray exercises, the assessment is fair. | While exams based on in-tray exercises get away from measurement of ‘speed of writing’ towards ‘quality of thinking’, different students’ ‘speed of reading /absorbing’ the information provided can be a problem. | Because of the relevance of ‘in-tray’ exercises and problems, a bank of such resource materials can be useful in the day-to-day teaching and learning of the subjects involved, and particular exercises can be used as examples to add variety to whole-class contexts such as lectures. |
| **Artefacts** (e.g. sculptures, paintings, architectural designs, engineering models) | Assessment of artefacts such as these is high on validity and authenticity.  Artefacts are useful as evidence of achievement to show prospective employers. | Assessment can be compromised in terms of reliability where different judges have their own idea of what constitutes excellence. | Where the artefacts can be retained (or photographed) by the institution, they provide excellent indicators of the standards of evidence of achievement for new students to work towards |