# Exams: bits of my books

First are two extracts from ‘Making Learning Happen: 3rd edition’ due for publication by Sage in May 2014. The first part gives an overview, linking traditional exams to the seven factors underpinning successful learning, as outlined earlier in the book (and in previous editions). The Table which follows is new to the 3rd edition, and is the first section in a comparison of a number of assessment types.

**Traditional exams**

In particular, let’s take the example of time-constrained, unseen written examinations. In other words, candidates don’t know the questions until they see them in the exam room. They work against the clock, on their own, with pen and paper. Assessment systems in the UK are quite dominated by this kind of assessment, usually at the end point of increments of learning. The assessment can therefore be described as summative.

As an assessment process, exams can be *reliable or fair –* if there is a well-constructed marking scheme, each candidate can be reasonably confident that the marking will be fair and consistent.

The main problem with many traditional exams is that they don’t rate highly on *validity.* In other words, too often they measure what the candidate can *write* about what they have learned, in the relatively artificial conditions of solemn silence, against the clock. Where, however, exams are based on problem-solving, case study analysis, and so on, validity can be much higher.

Exams can be improved in terms of *transparency* where candidates have been involved in applying assessment criteria to their own or other people’s exam answers, and have found out all they need to know about how the examiner’s mind works.

One of the major advantages of exams is that we are reasonably certain (with due precautions) that the work of the learner is being marked – in other words, that side of *whodunit?* is assured. The the extent to which the assessed performance relates to the normal conditions in which the learning is intended to be applied – is less assured, and in some traditional exams the conditions under which achievement are measured are quite alien.

## 1 Traditional exams and wanting to learn

For many exam candidates, the ‘want’ to learn is damaged by the mere thought of looming exams. Many learners, if given the choice, go for learning modules that are continuously assessed rather than assessed by examination because of their fear – and even dread – of exams. Few assessment processes induce such high emotions. This is not the case for everyone, however. Some candidates love exams – and are very good at preparing for them and doing them. Not surprisingly, the cue-seekers mentioned earlier in this chapter are among those who are good at traditional exams. Their cue-seeking approach is thus rewarded by this pervasive assessment format.

## 2 Traditional exams and needing to learn

This is where the intended learning outcomes should come into their own. Ideally, if learners have systematically prepared to demonstrate their achievement of these outcomes, and practised doing so sufficiently, they should automatically remain able to demonstrate the same achievements under time-constrained, written exam conditions. However, there is often a gulf between the intended learning outcomes as published and what is *actually* measured by traditional exams. Due attention to achieving constructive alignment can overcome this problem. But there is another side to needing to learn. Candidates who prepare successfully for exams by mastering the intended learning outcomes so that they can demonstrate their achievement in answering likely exam questions often concentrate very firmly on what they perceive they need to learn, and don’t invest time or energy in things they decide can’t (or won’t) come up in the exams. We are therefore favouring strategic learners by the use of exams (and, of course, cue-seeking strategic learners do best).

## 3 Traditional exams and learning by doing

There is plenty of learning by doing *before* traditional exams. But not much further learning by doing happens *during* traditional exams. It can, however, be claimed that a looming exam is as good a way as any of causing learners to get their heads down and do some learning. We could argue, however, that preparing for an oral exam (viva) would have just as much effect on learning by doing.

## 4 Traditional exams and learning through feedback

This is where traditional exams do really badly. As far as feedback is concerned, they are mostly lost opportunities. By the time the scripts are marked, learners have often moved on to the next part of their learning and are no longer really interested in which questions they answered well and why, or (more importantly) in where they lost marks. Many learners were *very* interested in these matters immediately *after* the exam, and spent some hours in post-mortem mode trying to work out how their answers measured up to what was being looked for by their examiners. All the feedback that most learners receive – after some time – is their score, or their grade, or simply whether they passed or failed. It is feedback of a sort, but hardly formative feedback. We can, of course, argue that exams are intended to be summative measures, but they still seem to represent lost feedback opportunities. Where feedback *is* provided very quickly after an exam (for example, in computer-marked multiple-choice exams, where a feedback printout can be handed to each candidate on leaving the exam room), feedback can, indeed, play a much more powerful role even in summative testing.

## 5 Traditional exams and making sense of what is being learned

This, too, links badly to traditional exams. As with learning by doing, a great deal of making sense of the subject matter occurs *before* an exam and, indeed, could be argued to be happening *because of* the exam. But few exam candidates report later that the moment when the light dawned was *during* the exam. More often, they report that they only found out that the light *had not* dawned during the exam. And then we need to ask whether traditional exams are measuring the extent to which learners have made sense of what they learned. Too often, exams seem to measure what learners can *reproduce* rather than what they can *do.* Many learners can tell us about the frequent occasions where surface learning was all that they needed to engage in to address the task of answering a particular exam question.

## 6 Traditional exams and learning through verbalising

In Chapter 2, we explored the significant benefits to learners resulting from them verbalising orally, explaining things to each other, coaching each other, and so on. Sadly, traditional exams only really get at verbalising in writing – and with a pen at that usually. One of my worries about traditional exams is that learners tend not to discuss things with each other, but to go into competitive mode, and hide their learning achievements from each other rather than celebrating such achievements. We can, of course, try to counter this tendency, and encourage learners to work together in their preparation for exams, quizzing each other, explaining things to each other, and so on, providing good rehearsal for doing similar things in writing on their own in the exam room.

## 7 Traditional exams and learning through assessing – making informed judgements

We could argue that in the context of traditional exams, most of this kind of learning occurs in the minds of examiners, not learners! Too often, *exactly* *how* the informed judgements are made by examiners is hidden from learners; examiners seem to fear the consequences of sharing with learners details of how marking schemes work in practice, possibly dreading future appeals by learners against ‘academic judgement’. We can indeed encourage learners to self-assess practice exam performance, and to peer-assess each other’s practice as they head towards exams, but the competitive ethos of exams militates against them doing either of these wholeheartedly. Moreover, there is fear involved – fear of finding out that the performance is not going to be up to the standards desired – and this can lead to self-fulfilling prophecy and lower attainment in exams.

Perhaps the main problem regarding learning through making informed judgements in the context of traditional exams is that such assessments are often the ‘mystery black box’ in nature, where learners do the best they can, hoping it will be found to be satisfactory or better. Other assessment formats tend to be more open to learners regarding exactly how they work (though this is not always the case).

## Traditional exams: summary

The picture painted above of the links between traditional exams and the factors underpinning successful learning is very bleak. It does not *have to be* so bleak, however. With care, for example, exams can be designed which are much better at measuring ‘making sense’ than suggested above. Problem-solving exams and case study exams are much better at *not* rewarding reproductive learning. But the concerns remain about the damage that can be inflicted on many candidates’ *want* to learn, the artificial way that exams can skew the *need* to learn and the fact that so much work may be done by examiners making sure that the exams have been fair and reliable, yet very little feedback usually reaches learners. In some ways, it seems that traditional exams are diametrically opposed to all of the central factors underpinning successful learning! Couple this to the problems of achieving validity, fairness and transparency, and it is surprising that in some assessment cultures (including much post-compulsory education provision in the UK) traditional exams continue to hold sway to the extent that they do.

First element of a new table comparing eleven assessment types

| **Type of assessment** | **Status** | **Advantages** | **Disadvantages** |
| --- | --- | --- | --- |
| **1 Traditional Exams**  Exams are often referred to as the ‘gold standard’ because of their widespread use in secondary and higher education. Still the most common kinds of exam, are handwritten, invigilated, against-the-clock, with questions not being known by candidates beforehand.  Exams remain prevalent on many post-compulsory education courses, sometimes where questions are set by external examiners, and sometimes by the staff who teach the learners. Timescales vary, but two hours and three hours are relatively common in universities, though much shorter exams are perfectly possible (and perhaps desirable).  Many traditional exams offer candidates a choice of questions (e.g. attempt any 5 out of 8 questions, each carrying equal marks), but increasingly there may be a compulsory section, then a section providing choices. | **Validity:** poor, limited to what comes out of pens.  **Fairness:** can be good, but poor when answers are essay-type, and different markers would award very different marks for the same essay.  **Whodunit?:** relatively safe (though stories of ingenious cheating are legion!).  **Real world:** written exams are not at all close to the workings of the real world; most people never do a written exam again after leaving university.  **Feedback to learners:** very limited indeed, usually just a score/grade or pass/fail, which can leave candidates having very little idea about what they did well or badly. | Can avoid plagiarism and cheating.  Give data which can be ranked and handled quantitatively.  Exams are relatively familiar to learners entering higher education, as they’ve already experienced them at school.  Exams are already ‘hard-wired’ into many university systems, so a case doesn’t have to be made for continuing to use them.  Written exams are much better for some subjects than others: for example they can work well for mathematical and quantitative subject matter, and tend to work really badly for ‘wordy’ or descriptive matter. | As has been argued already in this book, exams tend only to measure what comes out of pens, a poor proxy for what might be in heads.  Many otherwise capable learners never show their best efforts under exam conditions.  It can take a long time to mark a set of exam scripts (properly). (There are economies of scale for large numbers of candidates as examiners become familiar with the marking scheme being used).  Problems with speed of writing and legibility, and difficulties candidates face when using a se.  ‘Sudden death’: a bad day can mar a lifetime.  A snapshot of achievement, rather than a real measure of it.  One of the main skills measured tends to be time-management – dividing the available time sensibly between the questions being attempted.  Promotes surface-learning: filling heads with information to use ‘on the day’ and forget as quickly as possible thereafter.  Question-spotting by candidates can pay off substantially, meaning that at least some candidates pass without having learned the whole syllabus reasonably well.  Where candidates have a choice of questions, it is really hard to get all of the questions to be of equal difficulty – leading to at least some candidates ending up with an easier exam than others overall.  Long-answer written exams contribute to the continuing trend for post-compulsory education to remain ‘elitist’ – i.e. to favour those who are good at such exams. |

**Extracts from ‘The Lecturer’s Toolkit: 3rd edition (London, Routledge), in the process of being revised for a 4th edition expected in 2015.**

***Concerns about traditional exams***

Much has been written about the weaknesses of traditional examinations – in particular time-constrained unseen written exams. In many subject disciplines, this assessment format seems to be at odds with the most important factors underpinning successful learning. Moreover, there is abundant evidence that even in discipline areas where the subject matter is well defined, and answers to questions are either correct or incorrect, assessors still struggle sometimes to make exams valid, reliable, authentic or transparent to students. In disciplines where the subject matter is more discursive, and flexibility exists in how particular questions can be answered well, it can be even harder to achieve demonstrable reliability in assessment, even when validity is well achieved.

Overall in higher education at present, with staff time under more pressure than ever before, there is evidence of a drift back to reliance on exams, which can be argued to be one of the more time-efficient and cost-effective methods of assessment, where it is fairly easy to achieve fairness and reliability, and with the added bonus that plagiarism or cheating cause less headaches to markers than in many other forms of assessment.

Some of the principal concerns that can be expressed about unseen written exams are summarised below. Half of the fourteen concerns below are conflicts between this kind of exam and the seven factors underpinning successful learning which we considered in Chapter 1. It could be said that unseen written exams are diametrically opposed to these factors working successfully. Several of the remaining concerns relate to our own procedures for marking students’ exam scripts.

1. **Exams don’t do much to increase students’ ‘want’ to learn.** Students often make choices in modular schemes strategically, so that they avoid this kind of assessment if they can. This can lead them to choose subjects in which they are less interested than those which they fear to select because they will be subjected to exams.
2. **Exams are not often a good way of alerting students to what they really need to learn.** Admittedly, students will often only get down to serious learning when an impending exam causes them to revise actively, but the fact that in unseen exams the actual assessment agenda has to be guessed at rather than worked towards systematically means that the resultant learning can be unfocused, and the assessment judgement becomes too dependent upon the success of the agenda-guessing.
3. **Exams are not ideal occasions for learning by doing.** Though students may do a lot of learning *before* formal unseen written examinations, their actual experiences of learning *in* such situations is very limited. In other words, a note could be placed on the door of the exam room stating ‘exam cancelled; you’ve already done all the learning that this exam could have caused’! The learning payoff during an assessment element should be considered more. It is therefore worth our while revisiting our testing processes to search for forms of assessment which are in themselves better learning experiences.
4. **The amount of feedback that students receive about exams is far from optimal.** Most systems require marked exam scripts to be regarded as secret documents, not to be shown to students on any account! It is worth asking what reasons underlie this philosophy? You might have noticed that among the student demands in their charter in 2010 was a request for feedback on exams. It is useful to reconsider the value that students can derive from seeing their marked examinations papers, where it should be possible to be able to demonstrate to students that the examination marking has indeed been reliable, fair, and valid. Moreover, the natural process of learning from mistakes should always be accommodated, even when the assessment judgements have already been taken down to be used in evidence against the candidates.
5. **Exams tend not to do much to help students make sense of what they have learned.** While there may be a significant amount of making sense of concepts and theories during the time leading up to exams, the assessment experience itself does little to help students to gain any further deepening of their grasp of these. One of the consequences of modularising the curriculum can be that some subject matter is introduced too close to an impending exam for the content to be satisfactorily digested.
6. **Written exams don’t do much for learning through *verbalizing.*** However, practising putting learning into spoken words pays enormous dividends even when the eventual testing is written. We need therefore to encourage students not to study for exams in solitary silence, but to work for at least some of the time explaining answers to likely answers to each other (or anyone else who will listen).
7. **Written exams don’t encourage students to learn by *assessing.*** The tendency is for students to do what they can and hope for the best. However, the more we can encourage students to practise making judgements on their answers to typical exam questions, the better they can make similar judgements while they are answering questions in exams. The more they know about how marks are gained and lost, the better they can structure their efforts towards doing written exams successfully.
8. **We mark exam scripts in a rush.** Most staff who mark exams agree that the task usually has to be completed in haste, in preparation for timetabled exam boards. The situation has been worsened by modularization and semesterisation developments in most institutions, which give tighter turn-round intervals between examinations and progression to the next element of study. While our marking may still be fair and reliable, it can be shocking to students who have spent a great deal of time preparing for unseen written exams to find out that their scripts are marked so quickly.
9. **Unseen written exams can lead to us placing too much emphasis on unimportant factors in candidates’ answers.** For example, factors such as quality of handwriting, or neatness of overall presentation of scripts can influence examiners, consciously or subconsciously. Many students nowadays are much more comfortable composing essays or reports using a keyboard, and adjusting their writing on-screen, cutting and pasting to bring their writing to a logical or coherent whole; this is well nigh impossible to do well with pen and paper, against the clock, in a threateningly silent environment. Moreover, there’s a factor we don’t see but that students know all too well – speed of handwriting.
10. **We’re often tired and bored when we mark exam scripts.** Because of the speed of marking, and the pressure to do the task well, we may not be functioning at our best while undertaking the task.
11. **We’re not good at marking objectively.** There is abundant data on the problems both of inter-assessor reliability and intra-assessor reliability, particularly with the more qualitative or discursive kinds of exam answer.
12. **Unseen written exams tend to favour candidates who happen to be skilled at doing exams!** We’ve created an examinocracy! If we look at exactly what skills are measured by unseen written exams, the most important of these from students’ point of view turns out unsurprisingly to be the techniques needed to do unseen written exams, and the same students can get rewarded time after time! This skill may have little to do with the competences we need to help students to develop to become professionals in the subject disciplines they are learning.
13. **Unseen written exams force students into surface learning, and into rapidly clearing their minds of previous knowledge when preparing for the next exam.** Students are encouraged to clear their brains of the knowledge they have stored for each exam in turn. This of course is quite contrary to our real intentions to help students to achieve deep learning.
14. **There are many important qualities which are not tested well by traditional exams.** For example, unseen written exams are limited or useless for measuring teamwork, leadership, and even creativity and lateral thinking, all of which have their parts to play in heading towards graduateness.

Despite these concerns, there is a lot we can do to make exams work better or in different ways, for example open book, open notes, time-unconstrained exams, in-tray exams, OSCEs and so on. Some discussion is given later in this chapter, and further developed by Race(2010, 2014).