Joyful online assessment: beyond high-stakes testing

Jon Dron <jond@athabascau.ca>, Athabasca University, Canada

Abstract: One of the biggest upheavals in education systems that has been caused by the Covid-19 pandemic has been the need to rethink traditional summative (terminal) assessment processes, especially high stakes tests exemplified by the proctored, unseen, written exam (PUWE). This paper presents a framework of constraints for designing alternative assessments, and describes a summative assessment design approach to fit those constraints that centres around eportfolios, combined with methods to increase trustworthiness while making a positive and even joyful contribution to learning. Thirteen distinct benefits are described, and outstanding issues are discussed.

Introduction

The Covid-19 pandemic of 2020 has, amongst many other disruptions, brought changes in assessment practices - especially summative, for-credentials assessment - for many teachers and institutions. One of the most notable casualties has been the traditional proctored, unseen, written exam (henceforth PUWE, pronounced 'poo-ie'), the archetypal form of high-stakes test (HST). In this paper I will argue that this is a very good thing. The manifold weaknesses of the PUWE and other HSTs have been known for many decades, but they persist because most of the alternatives are perceived to be less desirable due to some combination of reliability, discrimination, precision, credibility, difficulty, efficiency, susceptibility to cheating, or cost. In this paper I will present an alternative online assessment method that has been field-tested in a purely distance context for over a decade that addresses all of these concerns, while making a positive contribution to motivation and learning.

The PUWE has many failings, of which the following are among the more obvious:

- They are inefficient. if we assume that the purpose of teaching is to help students to learn, PUWES make a little, no, or even a negative contribution to learning, except in the ability to take further PUWEs (Kohn, 1999, p. 201). Furthermore, they are normally offered in *addition* to assignments and coursework that still need to be marked, and that usually provide better, more nuanced evidence of meeting the same learning outcomes.
- Used as extrinsic motivators HSTs cause focus on achieving the credential, not learning, and thus they persistently and reliably diminish intrinsic motivation in the subject (Kohn, 2011; Ryan & Deci, 2017; Ryan & Weinstein, 2009) as well as being active causes of cheating (Kohn, 2007) which is often the easiest way to gain the credential.
- PUWEs are almost always inauthentic measures of competence, providing conditions and activities that hardly ever reflect any in the real world, and thus they are very poor ways to judge actual capability.
- The shortest path in setting exam questions tends to be to demand students to remember facts, repeat procedures, or mechanically apply rules, so most focus on a shallow, naively objectivist concept of knowledge as a fixed and repeatable set of behaviours.
- Even when efforts are made to examine deeper, more creative, more expansive knowledge or skills, students are unlikely to reach great depth when forced to do so under pressure. HSTs discourage risk-taking behaviours in many students (Atkinson, 1957).
- PUWEs are unfair on many students who cope poorly with the artificial, high-pressure, inauthentic situation in which they are placed, and overly advantage those who thrive in such contexts. They cause boredom, frustration, and drop-outs (Mora, 2011)
- HSTs encourage teachers to teach to the test, rather than teaching to instill love of or curiosity in what is being learned and the joy of learning it.
- PUWEs are weak at the one thing that recommends them, the prevention of cheating. The majority of students in most countries cheat in exams (e.g. (Christensen & McCabe, 2006; Ma, McCabe, & Liu, 2013)). For each technology we invent to prevent cheating (be it a method or a tangible tool), someone will always develop counter-technologies. There is a whole industry devoted to the production of exam cheating aids. This is an unwinnable war for all concerned.

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

Online PUWES, whether through dedicated proctoring software/hardware or simpler video conferencing tools, bring new and, often, insurmountable difficulties of their own. Some are technical. Simple webmeeting tools are notoriously fragile at the best of times but the greater the sophistication and complexity of the defences against cheating that are built into the tools such as online proctoring systems, the less reliable they become. To make matters much worse, such systems generally rely on students' own computers and networks at home, few of which are professionally maintained, and skill in using software with which students may be unfamiliar. It is hard to imagine many things worse for a student already facing one of the highest stakes few hours of their lives already, to have to try to cope with an unfamiliar and hostile technology that might not even work, that may penalize them for departures from a machine's perceptions of 'acceptable' normal behaviour, or that has a high likelihood of failure during an exam session. Add to that massive concerns about invasions of privacy, together with a lot of effective counter-technologies designed to subvert it, from hidden displays and earpieces to others in the room, then throw in difficulties and inequalities in ability to use digital tools, and it becomes an entangled, unreliable, invasive, intrusive mess.

Of course, like any technology, PUWEs can be designed to be less harmful. For instance, Huntrods and Dron (2017) write of an exam that essentially has only two questions, relating to a project-based course: 'what did you do?' and 'what did you think of the course?' The fact that the questions are known to the students in advance and can be safely published on the web is of no help to cheats, and no one who has not taken the course could provide an authentic account of the process. Unlike most PUWEs, the reflective tasks are useful learning exercises in themselves, encouraging close examination of how and what students learned, as well as encouraging them to make new connections. They are engaging, personal, and meaningful activities. However, the pressure to perform an HST in an unrealistic setting with tight time constraints remains problematic and exceptions like this are all too rare.

Alternative methods

There are at least millions of alternative approaches, hundreds of which are commonly used, from oral exams to coursework projects to authentic evaluation of workplace learning to portfolios. In this paper I will be presenting one of these but, first, it is important to frame what assessment should and should not do. From a broad perspective, assessment serves three main purposes within an educational system. The first, and most important, is as an intrinsic and essential part of the learning process, as what is often described as 'assessment *for* learning' (Stiggins, 2002). Assessment, whether by self or other, provides formative feedback that helps learners to learn more, that identifies strengths and weaknesses in both the process and the outcomes of learning, and that allows learners to progress more effectively, efficiently, and proficiently on their learning journeys. The second, closely related to the first and, in an educational system, nearly as important, is to provide teachers with feedback on the effectiveness of both their teaching and the progress of the student, providing them with information needed to adapt to student needs and to improve their practice. The third, often described as 'assessment *of* learning' which – in the interests of efficiency, compassion, and common sense - should only be in *addition* to the first and second, is to summatively judge whether learning has been effective for the benefit of others, such as employers or other institutions, usually through the award of credentials such as credits, badges, or certificates for successful completion. These summative judgements also present an accountable face to the world for the institution, its professors, and its courses.

With that in mind, I present an evolving set of design constraints that I have developed to serve as both a checklist and generative tool for any assessment practice, that have been informed by and inform my own assessment practice. These are a mix of pragmatic, ethical, and research-informed constraints, respectively labelled (P), (E), and (R) in the list below to indicate their provenance. Their purpose is to act as reminders of both principles and values underpinning those principles on the basis of which any assessment can be evaluated.

Pedagogical constraints

To be of educational value, assessment practice should:

- (P) support and promote learning (there should be no assessment that is not also a valuable learning experience);
- (P, E) respect and accommodate diverse learning approaches and needs (this implies diverse learning paths, and accommodation of different kinds of evidence of learning);

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

- (P) encourage habits of learning that make future learning more successful (learners should learn to learn in ways that are useful, as well as learning the topic or skill at hand. This implicitly excludes revision-for-the-test and other extrinsically motivated methods that depend on compliance with teacher instructions);
- (P) Provide timely feedback that should be supportive of future improvement (grades are rarely if ever formative feedback);
- (E) be compassionate (feedback should consider the learner, their distinctive needs, and what is needed for them to improve, and students should know from the feedback that they are supported and cared for);
- (P) provide information to teachers about the effects of their teaching in order to facilitate better teaching (as Hattie (2013) puts it, to make learning and teaching visible).

Motivational constraints

To support motivational needs, assessment practice should:

- (R) based on the principles of self-determination theory (Ryan & Deci, 2017):
    - provide learners with autonomy (including control of both the learning process and the way it is judged);
    - be challenging but not impossibly so (learning, and proof of it, should not be too easy but nor should it be too hard);
    - enhance, support and promote relatedness to others (learning should be with others, or for others, in ways that matter to the student);
    - limit and, where possible, eliminate extrinsic motivation (assessment should not directly reward or punish, except insofar at the activity itself brings pleasure, satisfaction, or fulfilment);
- (E) never be the purpose of learning, only a consequence and enabler of it (this is difficult to achieve when credentials are awarded because, as Goodhart Goodhart (2006) tells us, when a measure becomes a target then it ceases to be a good measure. Attempts should be made to reduce the ill effects by treating credentials as awards rather than rewards);
- (P, E) be low stakes (students should be able to keep trying until they succeed - there is no such thing as failure, only success that has not yet been achieved);
- (P,E) bring joy (bearing in mind that the assessment may be difficult, challenging, frustrating at times, and not necessarily pleasurable: this is far more about satisfaction, achievement, and love of the activity, rather than things that make us smile or feel content).

Housekeeping constraints

Credential-awarding assessment practices should:

- (P) represent authentic activities and contexts for learning (with some exceptions when assessment is purely or primarily formative);
- (P) reliably demonstrate competence (including capability to achieve in other contexts, as well as in the context of use);
- (P) be resistant to cheating (breaking of rules to achieve a better result);
- (P) be resistant to gaming (use of rules to achieve a better result);
- (E) be fair, unbiased, and non-discriminatory to any student (includes issues of accessibility);
- (P) be shareable before the event without compromising their integrity;
- (E) be judged by more than one person (or, at least, a sample of work should be considered by others to ensure equitable and reliable marking);
- (P) be as inexpensive, and resource-friendly as possible, for both teachers and students (they should not take too long to mark, should not require specialized equipment or services, should not demand extra learning from students);

Clearly, the vast majority of traditional PUWEs and other HSTs fail on almost every count to conform to the pedagogical and motivational constraints, and are only partially successful in meeting the housekeeping constraints. We need a better way.

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

Bullet-proofing e-portfolios

Portfolios, a popular alternative to other forms of assessment, may equally be seen as both process and product. As a process, they can provide a structure for systematic reflection and development that can be highly beneficial to learning (Hertels, 2004). As a product (sometimes known as a showcase), whether for summative assessment or personal marketing, they provide a reliable, nuanced, and comprehensive picture of a learner's competence. However, without great care, these two roles or paradigms may conflict with one another, especially when there are high stakes involved and students' own agency and ownership are reduced (Barrett & Carney, 2005).

The example method described here is a variation on the popular eportfolio approach to learning and assessment. Eportfolios warrant their own name because of how they affect the process. Specifically, they provide opportunities to share reflections and other elements of the process with other people, including other students, and enable use to be made of social and professional networks for support and engagement along the way. Special purpose eportfolio tools generally provide blogs, social networking, file sharing, and other tools that support this process but, as Barrett (2005) observes, almost any tool that supports creating and sharing content will work, and there are motivational and practical benefits to students in having personal ownership and choice of the tools and systems on which they reside.

Background and context

At Athabasca University, Canada's open and distance university, teaching and learning occurs in two main forms. The first, mainly limited to graduate teaching, is the traditional paced course: weekly activities are performed by all, usually supported by asynchronous discussion forums, and assignments are submitted on a schedule, in courses created and supported by a single instructor. The second, used in most undergraduate teaching, is self-paced. Self-paced courses are developed by teams of specialists, such as editors, visual designers, learning designers, as well as teachers, with a process that may take from months to years, that typically results in high-quality courses that then persist, largely unchanged, for years until the next revision cycle. Each course has a coordinator (often, though far from always one of the authors) who manages the overall course process and may also tutor some students, but additional tutors are added as course numbers increase. It is thus scalable from one to thousands of students, because most of the teaching work is done up front before students ever arrive. Tutors and coordinators mostly play a background support role, marking work, dealing with problems, and monitoring progress, only very rarely engaging in didactic practices. In a couple of faculties, a call centre emphasizes this reactive service role still further.

Students may enrol at any month of the year. They receive a study pack, nowadays mostly or fully online, including all the content, textbooks, readings, and media required for the course, and are assigned a tutor. They then have six months (extendible to up to a year, for a price) to complete the work. Most courses end in a PUWE. Before Covid 19, though many had already migrated to online proctoring systems, many were still in-person, making use of a large network of physical exam centres scattered around the world. Most courses also include multiple assignments, marked by tutors. Normally these assignments are summative high stakes assessments, counting towards the final credential. Social interaction is typically limited. Students may, if they wish, contact tutors by telephone, email, or web form if they need help but, otherwise, assignment feedback often represents the main human interaction on a course. Though mostly online for many years, this general format has remained largely unchanged for several decades.

Social learning in a self-paced model

The self-paced model presents many distinctive challenges, perhaps most of all for social learning. There are many benefits in learning with others. Self-determination theory posits that relatedness (the need for social value in what we do) is one of the three essential cornerstones of intrinsic motivation (Ryan & Deci, 2017) so, though this may be achieved through other means (for example, within a student's local circle of friends, relatives, and workmates), it is a very good idea to build support for connecting with and relating to other students, as well as with instructors. Other social benefits include seeing a diversity of viewpoints, of watching different approaches to learning, of developing ideas through discussion and debate, and of feeling supported and valued in a community. In a self-paced model, the chances of any two students doing the same things at the same time are, in all but the highest enrolment courses, slim to negligible. This rules out most forms of collaboration. Furthermore, the carefully designed pedagogies of self-paced courses, that by their nature must anticipate most if not all learning needs in advance, tend to rely heavily on a structured approach to delivery of content which, as Moore's (1997) theory of Transactional Distance implies, is inversely related to dialogue in a learning transaction (Saba & Shearer, 1994).

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

There are, however, some opportunities for peer support and cooperation. Pedagogical designs that are intended for individual study seldom if ever work well by simply appending collaborative social elements like discussion forums: the social aspect has to be designed into them from the start. This does not stop many teachers from trying. Some courses mandate contributions and give grades for participation though, in the light of the foregoing discussion on the risks of extrinsic motivation, not to mention the problems of assessing work in progress (before all the outcomes have been attained), this is far from ideal, and rarely realizes the most significant benefits of learning with others, especially as dialogue is rarely sustained or personal. Sometimes, whether mandated or not, more advanced students will help those at earlier stages in a course through discussion forums, however, which can help to reduce the loneliness as well as to share knowledge. Such interactions persist for the lifetime of a course, after many students have left, making them a potentially useful resource for later students as well as those receiving or giving help. Unfortunately, the hierarchical organization of discussion forums, difficulties searching, and problems identifying useful threads, make these less useful than they might be, and they are lost when, typically every 3 or 4 years, the course shifts to a new revision (and thus moves to a new instance on the LMS, which is an unfortunate necessity in order to avoid confusion for those still working on the previous revision).

Partly to address such needs, ten years ago we developed the social learning commons Athabasca Landing, an Elgg-based social media environment described in detail in (Dron & Anderson, 2014). The Landing is, in brief, a smorgasbord of social software tools (blogs, wikis, groups, discussions, microblogs, file sharing, bookmark sharing, etc) that can be used by anyone - faculty or student - to create, share, and connect with others within the university. Discretionary access control on every post allows members to share with as few or as many people as they wish, right up to fully public. Accounts, posts, and discussions on the Landing are persistent, unbound to the schedules of courses, and normally available in perpetuity, even when students are no longer studying at the university, thus providing a level of personal ownership not found in most systems. Every post can be the basis of discussion, and social connection is embedded throughout.

The method

The general approach that I am about to describe has been used in three self-paced and three instructor-paced masters level courses at Athabasca University for over ten years now, covering diverse areas from beginner-level programming to management practices, game design, ethics and social issues in computing. The courses involve very diverse activities, learning needs, levels, and pedagogical approaches, but share a common underlying methodology. I will first describe how this works, followed by a discussion of the strengths and flaws in the approach.

- All of the courses make use of portfolios of evidence as the sole formally graded assessment activity, and all involve use of a shared reflective diary - normally either a blog or wiki on the Landing – that is shared with and can be commented on by others, that serves as a repository for any learning outputs, and that forms the foundations of the portfolio. Submission of compiled portfolios and/or links to them on the Landing occurs through a learning management system (LMS) site which also contains the static course materials that describe the processes and goals, and that provides a little didactic content.
- Most courses require students to invent or use a real scenario, on which most or all of the rest of the work is based. For example, web development students incrementally build a single website to meet the needs of a scenario and personas identified at the start, while database management students structure diverse activities around a detailed business scenario, normally related to or drawn from their own work or personal interests..
- Structure is provided in the form of units of learning activity of varying length, each of which centres around solutions to specific problems and/or requires creative inquiry. Because the process demands the learning of what, to many, are new skills, the first unit (usually numbered ‘0’) is always designed to introduce the course process and to help explain how learning is meant to happen on the course, and which involves activities that elicit a lot of rich information about the students, their interests, their values, and their aspirations, which helps tutors to get to know them well.
- Each course provides a detailed set of learning outcomes but, rather than simply describing intended outcomes in ways that would only make sense to someone who has achieved them, these are used to structure the assessment process. Instead of giving marks for individual assignments, grades are given for each of those learning outcomes. The evidence for each outcome may come from one or more parts of the course, including projects, essays, problem-solving exercises, reflections, help given to others, discussion

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

contributions, shared resources, comments on other students' work, and so on, compiled in the portfolio. The evidence is not necessarily limited to work done within the course framework, though it should normally have occurred within the timeframe of the course itself.

- In recent iterations, though not yet implemented, students may create their own course-relevant outcomes, which can account for up to a maximum of 15% of the course mark overall. Those not wishing to do so, or who only use part of the allocation, can instead use that 'spare' 15% to provide greater weighting to other outcomes that matter to them, up to a maximum of 5% per outcome. To limit gaming of the system, decisions about the allocation must be made before the final submission, usually within the first few units or weeks of the course, in consultation with the instructor.
- In all the courses, as part of the portfolio-as-product, students are required to explicitly map work in the portfolio to the intended outcomes, usually using a provided template, and to give their own opinion about the grade it deserves. The evidence provided in the mapping document normally consists of annotated links to the relevant work on the Landing, though may sometimes include snippets or further reflections to explain or provide evidence of competence.
- No grades are given for any work during the courses themselves, though various different approaches to feedback are used in each course to ensure that students can improve their work, including tutor feedback, peer feedback, and self-evaluation. This is greatly facilitated by the embedded dialogue in all parts of the Landing.

- There is minimal instructor-authored conventional didactic content in any of the courses, and then mainly at an introductory level to provide scope and context. There are no mandated textbooks, though some courses provide open educational resources as optional support or background reading. Most didactic content, used to assist in solving the problems, is in part provided by either open educational resources, or links to web-accessible content (where necessary taking advantage of the university's extensive electronic library subscriptions but mostly pointing to free resources). Mainly, though, content is either provided directly in the work of former students, or in the form of annotated social bookmarks in the Landing, many of which are provided by students themselves. Sharing and annotation of useful bookmarks is another kind of evidence that students may use in their portfolios to demonstrate competence, so the course content evolves and improves over time, as well as staying current. These are read/write courses.
- Each course includes an FAQ on the Landing, where common problems can be addressed, poor descriptions can be clarified, additional content can be provided, and the instructor can maintain an active presence. This supplements and extends the static content in the Moodle LMS and, again, allows the course to evolve over time.
- In all the courses, though some work can optionally be made more private when students are uncomfortable with sharing personal reflections, almost all work is shared with (at least) others in the group, students are encouraged to look at and comment on what others have done, and may (if they wish) used such dialogue as evidence of competence for any relevant learning outcomes. Students thus learn from and with one another, including past students.

Thirteen benefits

This approach to course design and assessment brings at least thirteen distinct advantages over most alternatives:

1. It gives students far greater autonomy in, as well as ownership of, the assessment process and its products. They can choose work for their portfolios that represents their best efforts, can take quite different, personally relevant paths through the learning process, can link the work to past experience and future interests or needs. By mapping the evidence to outcomes, they control a lot of what the instructor will see. Being able to choose some outcomes or adjust the weightings gives greater ownership of the assessment process itself.
2. The presence of others is tangible and motivating, as well as providing multiple perspectives, feedback, and support, emotional and academic. Students on the Landing also benefit from the visibility of other students, professors, professional staff and researchers who also use the Landing. Sometimes, former students contribute. The course has permeable boundaries. Those not wishing to engage still share, and still benefit from the conversations of those who do, but are not put into situations that may be uncomfortable for them (except where such situations are unavoidably implied by the course learning outcomes, such as the ability to debate in a course on ethics).

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

3. The mapping process itself is a useful learning activity, demanding reflection on the complete course and the connection of multiple activities within it, as well as recollection to reinforce learning. This also gives students a chance to double-check that they have met the intended outcomes and, if any are lacking, to make good, so quality of submitted portfolios tends to be high.
4. From a teacher's perspective, the mapping usually makes marking of what might otherwise be a lengthy, unique, and complex submission at least as easy as exam marking. On the whole, students at all levels tend to fairly accurately assess themselves, in part because of the feedback provided throughout, so it is usually only necessary to give detailed feedback when there are disagreements. It is useful that, for the most part, learning outcomes can only really be understood well by those who have met them so, when they have not been met, it is often obvious without having to look too hard because inappropriate mappings will be provided.
5. Another benefit for teachers is that it is extremely accountable to third parties and transparent to students in the event of a dispute. It is guaranteed to provide solid proof that the course outcomes have been met, with absolutely clear and unequivocal evidence to support it. It also makes it very easy indeed to demonstrate perfect alignment between teaching, assessment, and outcomes.
6. It makes cheating extremely unlikely. Part of the reason for this is the many eyes that act as a disincentive. Part is the pride that students take in presenting themselves, not just their work, and the ownership they have of the process and products. Part is the fact that (in most cases) it is work that they have chosen to do themselves, with relatively little teacher coercion or control, so it is interesting, engaging, and relevant to them: they love doing it, so cheating is far less likely. Part is the fact that it cannot – for the most part - be plagiarized and, though contract cheating remains a risk, it cannot be done without doing all the work required for the course, so the costs of hiring someone to cheat are much higher than for other forms of assessment, including PUWEs. Additional mechanisms are provided in all the courses. Some are purely psychological: for instance, all include a pair of eyes (usually mine) where cheating might occur (especially in final portfolio submission) that research suggests can greatly diminish the chances of dishonest behaviour by around 35% (Dear, Dutton, & Fox, 2019). Some make use of difficult-to-fake products, like dated files, log files, and so on, or require a prior component to have been created for the next to work. The paced courses include portfolio presentations to the class, that virtually eliminate the possibility of contract cheating, though this mechanism is not available in self-paced courses in which students do not meet in real-time.
7. It reveals learning and competence that most traditional forms of assessment would miss. By allowing any evidence to be used, work that would otherwise be invisible to teachers can be revealed. The option in recent iterations to add outcomes that are of greater interest or relevance to them can highlight outcomes beyond those are intentionally taught.
8. It supports far greater diversity in learners than one-size-fits-all methods like PUWEs. Apart from the flexibility in process and content that the portfolio process provides, one of the biggest benefits is that those who do not wish to or cannot engage with others will not necessarily fail as a direct result because they can use other evidence of meeting outcomes apart from engagement with peers. There are strong extrinsic as well as intrinsic motives to engage, however, because interactions can be used as evidence of (for example) problem-solving or competence.
9. For any individual component, the stakes tend to be low because limitations can be overcome through other means. The portfolio itself is very high stakes, but there are plentiful opportunities for students to perfect it, to take their time (especially in self-paced courses), and to get feedback along the way to ensure it is as good as it can be. Some students, not unusually, do rush everything at the end but most have plenty of feedback and opportunities to get it right.
10. Risk-taking and innovation is encouraged, because no single unit of the portfolio is critical to success, there is support along the way to warn against dangerously high risks, and multiple kinds of evidence can be offered for the same outcome, including those that might not normally be submitted for fear of losing marks. As long as students provide evidence of meeting the outcomes, the approach supports any kind of learning process.
11. The students' rich descriptions of the learning process and path through course materials reveals vast amounts of information about the course design - far beyond simple proxies for competence revealed in conventional assignments - that the teacher can use to make improvements to it. Course evaluation is not a separate chore to be completed at the end, but an integral and continuous part of the process. Often, students volunteer recommendations for improvements and positive and negative appraisals of the process in their reflections.

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

12. Students provide an excellent source of materials and resources for augmenting, clarifying, and adapting the static course materials, so that courses can remain current and topical for much longer than most of their ilk, and they can evolve to better address learner needs as and when they are revealed. Often, students share good practices that then evolve to become the norm. For example, in one relatively high-enrolment course, students have incrementally adapted frameworks for reflection provided in the course materials to create a better, more reliable, easier-to-accomplish template that is now used by the majority in preference to my own. These truly read/write courses evolve and improve while they are running, not just when explicitly revised.
13. Perhaps best of all, for most students, it can be meaningful, fun, and deeply engaging. Portfolios of this nature can more easily bring joy to both learners and teachers than most conventional forms of assessment, especially as use of better conventional forms does not prevent their inclusion in the portfolio themselves. This is especially valuable in overtly creative disciplines like poetry composition, arts, and design, but there are no disciplines that are so dry that no joy can be found in them. Students often talk of their pride and pleasure in what they produce, and willingly share their portfolios with wider audiences, and one of the biggest complaints is that it takes too much time because they do not wish to stop.

Five costs

Of course, no process is perfect for every student, all the time. The following are the main issues that I have found.

Unlearning how to learn

The course process, and especially the process of reflecting, mapping, and assembly of portfolios, is an extra learning burden for most students. It can sometimes cause confusion or fear because it tends to be unlike what most are used to or expecting. Despite carefully designed 'unit 0' elements that are intended to address this and to 'unteach' students the habits they have acquired in other courses and institutions, there is typically a need to provide constant and repeated reminders, support, and instruction on how to learn on the course, including pragmatic technical concerns like where and how to post reflections as well as more general difficulties with understanding the course process itself, reasons for lack of grades, explanations of outcomes, and so on. In instructor-paced courses this is easy: constant interactions and regular instructor blogs can focus on difficulties and issues to help keep students on track. In self-paced courses it is more difficult, and may require personal intervention and help by the tutor. It can be especially difficult to know whether a student is taking a break because they are busy in the workplace, or whether they have dropped out or submitted work somewhere invisible or unknown to the tutor.

The need for continual instructor engagement

The fact that course content consists largely of problems to be solved demands a greater degree of support than one that provides content in pre-digested forms for the students, or that wraps itself around a textbook. Tutors must constantly monitor how students are doing, as well as providing reactive support, to ensure that they are not going down fruitless or harmful paths, and are not overwhelmed by problems they cannot solve. Other students can play an extremely important role in this and, especially given the fact that they may use this as evidence of competence, many do. Shared learning diaries are extremely useful as ways to let teachers know a lot about how students are progressing and thus to reduce the chances of cheating, but work is needed to stay informed (automated notifications can help with this), and, until a post is made, leaves teachers completely in the dark about what students are doing and what help they need.

The process relies heavily on dedicated, passionate, caring tutors who fully understand the process and can put the time and effort needed into supporting students along the way. Unfortunately, not all tutors are equal in these respects. Much care is needed to ensure that they are fully engaged and on board with the methods. For courses with high enrolments, it can be difficult to maintain tutors' boundaries. I encourage them to respond to any student in need but there are cost and contract implications: each tutor has their own students and it is too easy for less able or willing or time-rich tutors to freeload on those who respond sooner.

Dunning-Kruger problems and grade-gaming

Though the majority of students think carefully about the mapping process and provide grades that closely match those of their formal markers, some students are overly modest in their self-grading, while a similar number try to

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

game the system (or suffer from an extreme Dunning-Kruger problem) by giving themselves solid As for every outcome. This can greatly lengthen the time taken to mark the work, so it is really important to provide clear and unambiguous criteria for grading, plentiful examples and explanations of good and less-good practice, and reminders where possible, as well as to offer help when needed.

The Dunning-Kruger problem can cause issues, especially on self-paced courses in which students choose for themselves when to submit their work, and who may therefore believe their work to be finished prematurely. It is not too unusual for students to believe their work to be excellent when it is very much not. There are several solutions to this. The first is social: when students share work with one another then it is normally fairly clear which provide good examples to follow/standards to meet, and which do not. I have noticed that the problem most commonly occurs among students who have been over-hasty, rushing to meet deadlines, who may not have taken the time to look closely at the work of others. However, weaknesses may not always be obvious to other students - there are risks of the blind leading the blind - and the problems can occur among even the most dedicated and careful of students. Secondly, therefore, it is important that students receive useful formative feedback that helps them to make improvements well before they choose to submit it for grading.

Cheating

Though cheating is very much less likely to occur, it is not impossible. When most work revolves around and must relate closely to a student-specified scenario that cannot be the same as any other it is extremely difficult and normally impossible to plagiarize more than a minor fraction of the work presented in the portfolio without also failing to meet the outcomes. However, once in a while students try. For example, on two occasions we have seen cases where one student has copied their reflections from those of another, despite the extremely high probability that the same person who marked the first would mark their copy, and the fact that the original author might easily have discovered the plagiarism themselves. There also remains the possibility that someone else might impersonate the student, despite the very high time commitment involved for the one doing the work. All the paced courses include an ungraded but mandatory live presentation of the portfolio to the other students, primarily because it is a wonderful learning opportunity for all, but it also has the side effect of virtually eliminating this possibility. However, it remains at least a potential problem in self-paced courses where such class meetings are infeasible.

Instrumentality

The fact that cheating still occurs speaks to the significance of the overall system in dictating student behaviour. While courses or even programs might be devised to be relatively cheat-proof, engaging, and personally relevant, as long as the overall system makes extrinsic motivation such an important structural feature, some students will still feel pressure and incentive to take the shorter paths. A not uncommon complaint, especially from visiting students taking a single course to fill in gaps in their home institutions or for equivalent credit when they have failed a course, is that they would much prefer to be told what to do, step by step, and to perform assignments and tests with clear, fixed and measurable outcomes. This is one reason that the ‘unit 0’ is so important and, for the most part, I have tried to provide templates, default readings, and simpler instructions to cater for such needs, but it remains a persistent complaint and an ongoing challenge. It is a source of pride to me when students complain that they are not being taught – that they have to do it for themselves – as long as they are learning and enjoying the learning process, but many students have learned to learn by being told what to do and how to do it, and it is hard to cater for those who are resistant to doing this themselves, whether through lack of confidence or instrumental focus.

Conclusions

The method presented here has been working for more than 10 years and applied in multiple contexts for many different subjects using many different pedagogical methods. It is one of, perhaps, millions of possible solution to the problem of designing assessments that are both joyful supports for learning and reliable summative judgment tools, and that address each of the self-set constraints listed. As a summative process this approach is reliable, discriminatory, accountable, authentic, and relatively non-threatening. As a learning process it is challenging, engaging, personal, meaningful, social, flexible, and, above all, fun. For the teacher, it makes marking interesting and seldom very difficult, though the need for close monitoring of and engagement with the learning process itself can be exhausting, time-consuming, and challenging, at least when compared with the traditional hands-off self-

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781

paced model. For those who love to teach, however, it is fulfilling in ways that the industrial model of the past cannot even approximate. It allows instructors and tutors to express caring, support, and engagement, to be visible parts of the learning process in ways that at least match those of their classroom-bound peers. This approach is natively online. It relies upon and exists within digital tools and systems, that are substantiated only in bits and bytes in machines. But, perhaps more than most in-person approaches (at least as processes and procedures) it is also profoundly and innately human.

References

Abrami, P., & Barrett, H. (2005). Directions for research and development on electronic portfolios. *Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie*, *31*(3).

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological review*, *64*(6p1), 359.

Barrett, H., & Carney, J. (2005). Conflicting paradigms and competing purposes in electronic portfolio development. *TaskStream web site*.

Christensen, J. M., & McCabe, D. L. (2006). Academic misconduct within Higher Education in Canada. *The Canadian Journal of Higher Education*, *36*(2), 1-21. Retrieved from http://athena.uwindsor.ca/units/csja/csja.nsf/0/28a0c81be449007c85257287006e6401/$FILE/Christensen-Hughes.Article%20%231.pdf

Dear, K., Dutton, K., & Fox, E. (2019). Do 'watching eyes' influence antisocial behavior? A systematic review & meta-analysis. *Evolution and Human Behavior*, *40*(3), 269-280. doi:10.1016/j.evolhumbehav.2019.01.006

Dron, J., & Anderson, T. (2014). *Teaching crowds: Learning & Social Media*. Athabasca: AU Press. Retrieved from http://teachingcrowds.ca

Goodhart, C. A. E. (2006). The ECB and the conduct of monetary policy: Goodhart's law and lessons from the euro area. *JCMS: Journal of Common Market Studies*, *44*(4), 757-778.

Hattie, J. (2013). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Taylor & Francis.

Hertels, E. (2004). Experiential learning portfolios in professional programs: A Canadian perspective. *J. Zubizarreta (2004). The Learning Portfolio. Bolton, MA: Anker Publishing*.

Huntrods, R., & Dron, J. (2017). *Engagement with Robots: Building a Social, Self-paced, Online Robotics Course*.

Kohn, A. (1999). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes* (Kindle ed.). Mariner Books.

Kohn, A. (2007). Who's Cheating Whom? *Phi Delta Kappan*, *89*(2), 89-87.

Kohn, A. (2011). The case against grades. *Educational Leadership*, *69*(3), 28-33.

Ma, Y., McCabe, D. L., & Liu, R. (2013). Students' Academic Cheating in Chinese Universities: Prevalence, Influencing Factors, and Proposed Action. *Journal of Academic Ethics*, *11*(3), 169-184. doi:10.1007/s10805-013-9186-7

Moore, M. G. (1997). Theory of Transactional Distance. In D. Keegan (Ed.), *Theoretical Principles of Distance Education* (pp. 22-38). Routledge.

Mora, R. (2011). " School Is So Boring": High-Stakes Testing and Boredom at an Urban Middle School. *Penn GSE Perspectives on Urban Education*, *9*(1), n1.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications.

Ryan, R. M., & Weinstein, N. (2009). Undermining quality teaching and learning:A self-determination theory perspective on high-stakes testing. *Theory and Research in Education*, *7*(2), 224-233. doi:10.1177/1477878509104327

Saba, F., & Shearer, R. L. (1994). Verifying key theoretical concepts in a dynamic model of distance education. *The American Journal of Distance Education*, *8*(1), 36-59.

Stiggins, R. J. (2002). Assessment Crisis: The Absence of Assessment for Learning. *Phi Delta Kappan*, *83*(10), 758-765. doi:10.1177/003172170208301010

This is a preprint. Final version: Dron, J. (2020). Joyful Online Assessment: Beyond High-stakes Testing. *Innovate Learning Summit 2020 2020*, 58-67. Retrieved from https://www.learntechlib.org/p/218781