Beyond Badges & Points: Gameful Assessment Systems for Engagement in Formal Education

Barry J. Fishman, University of Michigan, fishman@umich.edu
Sebastian Deterding, Hans Bredow Institute, s.deterding@hans-bredow-institut.de

Abstract: This symposium brings together a range of gameful assessment designs at different levels of formal education to explore how gameful design might lead to greater student engagement and improved learning outcomes. We use the term “gameful assessment” to describe assessment frameworks or approaches that employ game design principles to foster student motivation and learning. The symposium examines systems in both K-12 and higher education, and considers both the conceptual underpinnings of these systems and the design space of current tools developed to make it easier for instructors to implement gameful grading systems. Data related to the success (and struggles) of each system will be discussed.

Introduction

There is no question that well-designed video games are powerful learning environments (Gee, 2003; Squire, 2011) that motivate players and keep them engaged throughout ever-increasing challenges. Recently, the lessons to be learned from good video games have been extended beyond the literal design and use of games for learning to the use of game design principles to conceive of a different way to organize instruction, turning formal education itself into a game-like experience (Deterding, 2012). Instead of playing video games in K-12 or college classrooms, or learning through the engagement in the practices, affinity spaces, and larger ecology of video gaming (Salen, 2008; Squire, 2006), the idea is to use game design principles to create a gameful framework for learning and teaching. One common industry term for this approach is “gamification” (e.g., Kapp 2012); we prefer the phrase “gameful design” to avoid the charged connotations gamification, and to signify a wide focus on the full palette of game design to create gameful experiences for learning. These might include design methods and models, like play-centric design, or game design principles like “clear goals” or “supporting autonomy” – rather than a narrow focus on a small set of interface design patterns (points, badges, etc.) commonly associated with “gamification” (Deterding, Dixon, Khaled, et al., 2011). As many critics of the 2011-2012 “Badges for Lifelong Learning” DML Competition have noted, the ‘standard’ commercial points-and-badges implementation of gamification often replicates rather than transforms traditional grading systems, inadvertently replicating their shortcomings as well, such as a focus on performance metrics rather than learning and mastery (e.g. Reid, 2011) and an over-reliance on extrinsic rewards that can decrease deep and lasting engagement.

Indeed, a significant challenge in changing school practice is the inertia that comes from many directions, especially from students, who have grown used to the traditional approach to grading and other course structures. As Davidson (2012) frames the situation:

[Our students] were well taught and learned well the lesson implicit in our society that what matters is not the process or the learning but the end result, the grade.... where “success” has been reduced to a score on a test.... The message we’re giving our students today is all that really counts is the final score. (Davidson, 2012)

One criterion for the success of gameful approaches to educational design is that the problem is alleviated rather than exacerbated. Current work is happening at various levels of intensity and scale, with whole schools like Quest2Learn (Salen, Torres, Wolozin, et al., 2011) representing the upper end of the spectrum. Restructuring at the level of individual classrooms often represents the vanguard of experimentation, as instructors have substantial control over their local teaching choices. An early and prominent example of this experimentation is a 2009 university course on the theory and practice of game design taught as an MMOG (Sheldon, 2011). Many found this example to be intriguing, and set out to try variations on it themselves. These efforts often employ, either implicitly or as explicit design strategies, core motivational theories such as goal theory (Linnenbrink & Pintrich, 2000) or self-determination theory (Ryan & Deci, 2000).

As many within the education community experiment with these techniques, the time is right to gather together several different examples in order to evaluate whether, when, and how these designs
translate into more motivating and therefore more successful learning environments. In short: Can the design methods and principles that work to motivate players in games also motivate learners in typical classroom instruction – not just within discrete topics but at the level of an entire course, or an entire program or school? Specifically, this symposium focuses on the gameful design of assessment systems as part of larger learning designs. The use of game design to motivate learners goes beyond ‘just’ assessment (O’Mahoney et al., 2012), but since assessment systems provide a strong framing element for the broader learning goals of formal education, we believe they provide a good entry point and focus. The symposium includes examples from both K-12 and higher education. It also includes some emergent technological tools that are intended both to embody gameful course designs and to make them easier to enact. An explicit goal of this symposium is to bring key issues in course design to the surface, leading to an elaborated research and design agenda to inform future progress.

**Mapping the Design Space of Assessment Forms in Gameful Classrooms: Rationales, Patterns, Issues, Solutions**

Sebastian Deterding, Hans Bredow Institute, s.deterding@hans-bredow-institut.de

In recent years, a sizeable number of educators across settings and age groups have experimented with re-designing the whole classroom experience in the image of well-designed video games. At the 2008 GLS Educator’s Symposium, we conducted a one-day workshop convening 13 educators with experience crafting their own “gameful classroom” in order to facilitate the exchange of best practices and lessons learned, and to build a grounded understanding of current practices and problems. The workshop, together with follow-up surveys and interviews of further creators of gameful classrooms, revealed an astounding variety of designs, sometimes inspired by Lee Sheldon’s *The Multiplayer Classroom* (2011), sometimes independently bootstrapped solutions.

Several initiatives are currently on the way towards developing software platforms to support and scale such gameful design approaches across a larger number of classrooms, specifically gameful assessment frameworks (e.g., Fishman & Aguilar 2012; classrealm.com). Arguably, the design of such platforms should be informed by the experience of existing practitioners and designers, and start from an informed understanding of the total set of current solutions. To enable knowledge exchange between current practitioners and system designers, and to build a foundation for systematic research, this presentation offers a mapping of the design space of assessment forms in gameful classrooms, based on inductive coding (Schadewitz & Jachna, 2007) of interviews, taped moderated discussions, document and interface analyses of a total of 23 gameful classroom designs. The mapping is organized in three interlinked parts:

- **Rationales**: All surveyed educators have strong reasons for ‘daring’ to deviate from standard educational practices. These reasons inform the specific design solutions they devised, and the specific solutions are seldom understandable without knowing their rationale. We sketch the various rationales active in gameful classrooms, linking them to existing discourses and theories in game-based learning and learning theory.

- **Design Patterns**: Using (game) design patterns (Björk & Holopainen, 2005) as an analytic lens and organizing concept, we will describe the different forms of gameful assessment we found, and how they link to the underlying rationales.

- **Issues and Solutions**: We will outline the most common assessment-related challenges educators reported when implementing gameful classrooms, and solutions they found.

The final section of this presentation will explore patterns we found to be consistent and consistently reported as successful (e.g. emphasis on formative assessments, increased autonomy and reduced threat of failure through unlimited redoing and task options, turning ‘gaming the system’ into an explicit part of the course design), outline existing families or ‘ideal types’ of design approaches, draw recommendations and caveats for system creators, and describe further research needs.

**Motivating K-12 Math Students with Special Needs with MathLand**

Kate Fanelli, Beacon Day Treatment Center, katefanelli@gmail.com

I teach Algebra 1 and 2 to students with severe emotional impairments. Although my students receive special education services, they are on a diploma track and accountable for the same skills and
credits as their general education peers. They work below grade level in mathematics (average grade level is 10.2, average achievement level is 5.5) and have significant gaps in learning. My students are generally capable students who are unable to function in school because of their emotional issues. This manifests itself through school or work avoidance, acting-out behaviors, lack of focus, learned helplessness and/or poor relationships with school personnel. 87% of students at my school receive free or reduced lunch. 90% carry a psychiatric diagnosis, and 68% take medication(s). 25% of our students have a history of court-involved placements or interventions.

In teaching these students, I struggled with continuity (attendance is a major problem), skill mastery (the tendency was toward work production rather than learning), and motivation (students showed minimal motivation to do endless piles of work). I wanted students to be more self-directed, independent, and focused on forward movement and skill development. A solution was to re-design my classroom using a game-based approach, which I call “MathLand.” I use a cycle of formative assessment, self-assessment, and summative assessment to help students learn, check for their own understanding, and demonstrate both long and short-term mastery. Points are awarded for passing mastery tests, and final grades are based on number of points earned. The system is reinforced by the use of student avatars, which earn status and track student movement on an avatar board that is displayed in the classroom.

The curriculum is split into “levels.” Students begin at level one and work until all levels are completed or until the school year ends. If students earn 100 points in one marking period, they enjoy free time until the new quarter, or keep working and get a head start on the next quarter. One junior finished an entire course early and had the opportunity to begin her senior math class or be a teacher’s assistant for the last marking period. She worked hard to learn the curriculum because she wanted to finish the course, be the first student to finish so quickly, and choose how to spend the rest of the school year. Her accomplishment is legendary and frequently recounted by newer generations of MathLand students. Levels have three parts. The “lesson” has explanations and a few exercises. “Practice” is optional. If a student does not feel ready for the mastery test after the lesson exercises, or repeatedly fails the mastery test, he/she may choose to do practice problems until ready to continue. The “mastery test” is completed independently and must be done 100% correctly to pass. Students also take a summative assessment as a pre-evaluation and then at the end of each of four quarters to show long-term gains and retention.

My presentation will describe basic elements of the program (alignment to standards, lesson/practice/mastery), the grading system (cumulative point system), and assessment mechanisms (formative and summative). I will discuss effective motivating elements, and game design elements I use to maintain student engagement and program structure. I will discuss pros and cons of this program, implications for instruction and classroom management, and present anecdotal and statistical evidence of the program’s success.

Gaming on the Ground: Assessment at the PlayMaker School

Lucien Vattel, Tanner Higgin, Katerina Schenke, GameDesk
lucievattel@gamedesk.org, tannerh@gamedesk.org, katerinaschenke@gmail.com

Launched in September 2012, the PlayMaker School is an attempt to infuse a sixth grade curriculum in Los Angeles with play, making, and discovery-based activities. In this sense, it is one of a handful of attempts to make games and game-like elements a fundamental part of the classroom experience (e.g., Salen, et al., 2010). These attempts are motivated by the potential of games to motivate, engage, and facilitate deep, conceptual learning, and exploration of complex systems (Klopfer, Osterweil, & Salen, 2009). Mindful of the fact that games are not the perfect tools for every situation, the school is also interested in tracking the learning circumstances and outcomes suited to games and play (Squire, 2011). In this context, a question of great importance becomes: What does game-based learning and assessment look like when not an exception but a common practice, and how does it work over the course of the year, and throughout all content areas in a functioning school? In our presentation, we outline specific challenges, and solutions we have iteratively developed, including a larger assessment framework, as well as specific assessment strategies designed to better capture learning with games.
Over the course of the past year, through pilot studies of educational games and curriculum in after-school programs, we have found that one of the greatest challenges facing game-based learning in classrooms, and gameful curricular design, is designing and employing assessment approaches that both facilitate student learning and adequately capture that learning (Vattel & Riconscente, 2012; Vendlinski, et. al., 2010). It’s not enough to integrate games into traditional classroom approaches, because familiar assessment approaches often inadequately capture or outright misrepresent game-based learning; rather, we need radical new means of assessment to be built and used in conjunction with new playful classroom activities. At our school, this assessment framework manifests in curricular and assessment design, course management, and teaching practice.

Tackling this issue directly, we’ve developed processes and tools to help researchers and educators capture observations and design on-the-spot and extended assessments around emergent learning experiences. Our approach—developed iteratively and through collaboration between researchers, educators, administrators, and curriculum developers—consists of a “character sheet” which allows students to be assessed formatively and tracked across a range of non-traditional measures, a modular “adventure map” that dismantles the class period structure blending disciplines and allowing differentiated student pathways, and a “learning tool” course management system designed at New Roads School which captures and orchestrates the curriculum in our school.

Beyond these larger-scale initiatives, we’ve found a need for philosophical and conceptual changes at the pedagogical and curricular level. Educators need to articulate broader learning outcomes and cognitive processes, and to situate and classify these around particular emergent activities within the school, and to view student learning as developing slowly over time rather than in discrete situations. We’ve seen in our school, and within our curriculum development sessions, how educators aware of these concerns can build, refine, and share rubrics, assessment frameworks, and facilitation techniques for effective integration of physical and/or digital educational activities.

One example of how assessments around a game have been developed in this context is Newton’s Playground, a digital physics sandbox game and open design environment that requires players to “draw in” machines (levers, pulleys, and springs, etc.) in order to move a ball across a series of obstacles (Shute & Ventura, 2013). In Newton’s Playground, students have an opportunity to discover physics concepts such as conservation of angular momentum, torque, and potential and kinetic energy. There were no available assessment approaches for this game, or any similar physics sandbox style game. Consequently, we designed a set of assessment approaches unique to the learning process witnessed in the game. These strategies include talk-aloud sessions that extract vocabulary from the player during play, assessment through public and private exhibition of play strategies, problem solving of challenges both solo, in pairs, and in small groups, and annotation, review, and reflection of video captured play.

With these on-the-ground and on-the-spot educator strategies, placed in conversation with the larger assessment structures in place at our school, we hope to provide a model from which to discuss not only how games fit into the classroom, but how we can design assessment frameworks and practices that appropriately work from what games do well, and what students and educators do with them.

**Competition + Collaboration: Keys to Intrinsic Rewards in Higher Education**

Lee Sheldon, Rensselaer Polytechnic Institute, sheldc2@rpi.edu

It is not enough to simply change letter grades to experience points, or offer extrinsic rewards such as leveling, badges or leaderboards. As in any good game design, it is essential to celebrate the social aspects of multiplayer gaming, and build them into any assessment system.

When I designed my first multiplayer classroom in 2009, I first concentrated on a simple mapping of game elements to instruction and learning. But game design is a process of iterating, testing, then iterating again. It became clear as that class progressed that there were several elements that worked particularly well: grading by attrition, learning by failing, but most importantly the social interaction multiplayer games provide. This presentation will detail how outcomes and assessment evolved hand-in-hand to create a super-charged atmosphere of engagement that produced almost perfect attendance, higher class grades, and deeper retention of subject matter. Following are some examples that will be explored during the presentation.
Dividing students into guilds allowed them to both compete with other guilds and also collaborate. I began simply with a secret ballot peer review that allowed guild members to offer input on how well they felt their fellow guild mates were doing. I began designing exams with a section of questions which, if answered by any guild member, all members received credit. Most enlightening (and fun) were exam prep classes designed as competitions between guilds where again guild members could individually help their guild mates. Soon, emergent behavior appeared. In one case students found “better” ways to compete under the competition’s rules. Two other examples involved guilds that were doing better on the prep competition, helping weaker guilds so everyone could win that game. Later, during a class teaching Mandarin Chinese, students again, of their own volition, collaborated to play the game on a more complex level than I had designed. What occurred was not some students using the assistance as a crutch, as I had at first feared, but all students doing much better when assessed. Woven through the presentation is the narrative of how my own learning was enhanced by the depth of the students’ engagement; the ways in which they made the classes their own; and my seriously tardy realization of the importance of sustained narrative in what I now think of as “collateral learning.” I’ll introduce a new class teaching engineering called “These Far Hills,” the saga of a multi-generational Irish family emigrating to the New World: Mars. I am a writer. I should have known this from page one!

Proceduralized Gameful Course Design with Queso

Clayton Ewing, University of Miami, c.ewing@miami.edu

When we constrain ourselves to a set of rules, the choices we make are guided by the imposed system. This effect can be productively used in the design of “white label” learning platforms: By following a design-based research paradigm (Wang & Hannafin, 2005), I created a learning management system (LMS), named Queso, that adheres to a rigid gameful implementation. Therefore, any classroom using my software is guided by principles based on the experiences of myself and other users of my LMS in creating a gameful course. This presents an opportunity for instructors not well versed in gameful design to easily adopt these principles as a means of introspection to challenge the “traditional” model of pedagogy and allow for new models to emerge.

The software is built as a three-tier system composed of quests, skills and grades that mimic traditional role-playing game (RPG) leveling systems. An instructor starts by determining the skills a successful student should have when they complete the course. For example, a journalism instructor might value integrity, writing and research. Quests incorporate the various skills for the class: an article might be awarded points for Writing and Research while an in-class discussion of an article increases Integrity. Finally, an instructor chooses point thresholds for grades that are assigned to their lowest skill level. This approach encourages students to be well-rounded rather than excel at one particular skill. It also frames the course as a heroic effort of amassing enough points to achieve an A rather than not losing enough points to receive a failing grade.

Whether submitted digitally by the student or by the instructor, all quests are tracked within the software, which provides basic gradebook functionality for instructors and assignment tracking for students. The information is constantly shown to the student as a progress bar to increase their motivation (Lewis, Wardrip-Fruin, & Whitehead 2012), as well as letting each student be constantly aware of their standing in the course. The quest data is also visualized through charts and lists of completed assignments with instructor feedback for improvement. Allowing for the benefits of self-paced learning (Tatum, 2012) unattempted quests are presented to the student providing them with an opportunity to choose what interests them at that moment. The software does not confine an instructor to due dates and as a result quests do not need to be mandatory. A superfluous amount of quests can be created and provides a student multiple paths for achieving a top grade.

A central premise of video games is the freedom of failure. Once submitted, a quest does not end. By creating a psychosocial moratorium on quests, students don’t have to fear receiving a bad grade (Gee, 2003). A student can attempt a quest as many times as they want during the course in order to achieve a higher skill point total. Borrowing from the ideas of boss battles in games, quests can also be set to only allow students of a certain skill level to attempt them. The student practices with small quests and is rewarded with a big skill boost when completing a master quest.
This software has been used in various educational settings for courses on math, programming, visual design and game design. Once a course is created, an instructor can iterate upon its design by adding new quests as well as modifying and removing old ones while they slowly perfect the classroom experience. With further research, we hope to expand the types of courses taught and examine the metrics behind those courses to find ideal scenarios for gameful classrooms.

GradeCraft: A Tool to Support Gameful Teaching & Learning
Barry Fishman, Caitlin Holman, Stephen Aguilar, University of Michigan
fishman@umich.edu, cholma@umich.edu, aguilars@umich.edu

Gameful assessment systems are potentially motivating for learners, but also potentially challenging for instructors. Part of the challenge is related to the change in approach; new or different pedagogies present challenges to teachers who are used to organizing instruction and assessment in a particular way. Pedagogies that present more choice to learners and result in a broader variety of representations of learning, such as many gameful approaches, are naturally more complex and more difficult to manage than “traditional” didactic pedagogies (e.g., Crawford, 2000). And there are additional questions about what design elements (and in what combinations) in gameful approaches are most likely to be effective in motivating learners. How does one strike a balance between the extrinsic motivation that is typical of standard assessment approaches (Jürges, Schneider, Senkbeil, & Carstensen, 2012; Shepard, 2000) and the more desirable intrinsic motivation that gameful approaches are thought to inspire (Connected Learning, 2012)? Is it possible to design a Learning Management System (LMS) that increases students feelings of autonomy, belonging, and competence — all key elements of self-determination theory (Ryan & Deci, 2000)? If so, will this invariably lead to improved student motivation? To explore these questions, we constructed a LMS to both support the implementation of and support research on gameful approaches to teaching.

Our gameful LMS is designed with the goal of supporting grading systems that give learners more choice and control over pathways towards accomplishing course goals, and providing greater feedback to learners regarding their progress towards those goals. On the instructor side, our LMS also makes it easier for teachers to monitor the progress of individual students and groups of students, to organize and support both collaborative and competitive work, and to provide feedback on assignments that are linked to different kinds of recognition for student work. In its most basic form, the tool can be thought of as a replacement for the assignment and gradebook tools that are central components of typical LMS environments. We conceptualize the process of building this tool as a design-based research endeavor (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003), both extending the theory of how to support learner engagement through the structure of the course grading system and assignment structure, and supporting empirical research to better understand how the system can be improved through iterative design (e.g., Fishman & Aguilar, 2012).

A central feature of our design is the “Grade Predictor,” a visualization that lets students examine the points (grades) they have received in the course, and then literally “game the course” by exploring various pathways toward their desired final score or course grade. Students can compare their progress to classmates’ progress, viewed as a box and whisker plot of the mean, median, and range of progress both overall and on individual assignments. We hypothesize that the Grade Predictor increases students’ agency in the course by making the consequences of different choices clearer, encouraging students to experiment with different learning opportunities and assignments. Other features of GradeCraft include a framework for creating, awarding, and displaying, badges (Hickey, 2012) as a part of the assessment environment; tools for the interpretation of learning analytics data (both for instructors and for students); tools to support the formation and management of teams (both instructor- and student-determined); and assignment creation, collection, and grading tools, including rubrics that are linked directly to badges and point determinations.

We have examined the use of GradeCraft in several different university instructional settings, including large lecture-style courses and smaller seminar courses. We are interested in studying the use of GradeCraft in a broader range of courses and will do so as we refine both the tools and our understanding of how the design and use of the tools are related to greater student engagement in general as well as greater student feelings of autonomy, belonging, and competence in particular.
References


Kapp, Karl M. (2012). The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education. San Francisco: Pfeiffer.


