4PEG: A Structured Rating System for Games for Learning

Katrin Becker
Department of Mathematics & Computing
Mount Royal University
Calgary, Alberta
kbeker@mtroyal.ca

Abstract

Videogames are interactive by nature - people proceed in games by doing things, and this experiential quality lies at the very core of game design. Without interaction, it isn't a game. Videogames are popular precisely because of the experience - games designed for learning can do no less. However, to be feasible for use in formal educational settings, they must do more, and while we are making progress studying games in classrooms, there remain few structured approaches to analysing games that do not include classroom testing. This paper outlines the author's Four Pillars of Game-Based Learning (4PEG) which can be used to perform a structured analysis of both COTS and serious games to assess whether or not a game has potential for use in the classroom.

1. Introduction

Digital games are being used more and more often as teaching resources in the classroom [1]. What makes a game suitable for use in the classroom? The answer to this question is not nearly as clear-cut as we might like. The criteria that make a game suitable for use in the classroom are as varied as the uses to which it can be put. The criteria are different if we are using a COTS game, a serious game, or something in between. The length of play also changes things. This is further complicated by the fact that game scholars don't all agree on what is important in a game for learning. Jim Gee states that games for education must be replayable and immersive, that they must build expertise and identity, be transdisciplinary, and allow for different trajectories of expertise. They should also connect with the real world and create a community of learners [2]. These worthy goals but it is unclear if are always necessary, and some argue that they aren't always desirable. According to Carrie Heeter an educational game can be "an elegant experience that encourages and enables the target player to achieve the intended learning goal(s)" [3].

We are getting much better at designing and delivering authentic, meaningful education using technology, but many of the methodologies we employ to examine technology-enhanced interventions look primarily at how the learner is affected or changed [4] [5]. We've become better at researching games for learning over the last decade. According to a 2011 review of the state of game based learning, many early studies were flawed and of limited use, but more researchers are now paying close attention to the design of their studies as well as the kinds of games they choose to study [6]. While the importance of solid research into the efficacy of games is indisputable, it is also important to have ways of evaluating the object itself apart from the learner, especially as the design of instructional objects becomes more complicated and more expensive. If we are truly interested in promoting the use of games in the classroom, then we must address the needs of the teachers. The predictive evaluation model described in this paper addresses all of these.

2. What's Important in a Game for Learning?

A recent study conducted by Games and Learning found that teachers' choices of games in the classroom are influenced first by what other teachers say about the game, then whether or not it includes assessment within the game that connects with curriculum, then their own experience. Information on efficacy was ranked fourth [7]. In 2005, John Kirriemuir asked teachers what was important in games for learning and his list of requirements is as relevant now as it was in 2005:

- Examples from other teachers on how to use the game and their experiences when they did use it.
- Games should not contain advertising, long expositions, videos, or narrations.
- The game should be able to be started at a point (position) useful to the teacher.
- It should have well-defined goals and objectives that teachers can use as homework or in-class tasks.
- The game must be accurate in the process and facts it conveys, and should avoid political or scientific controversy [8].

Additionally, gamesandlearning.org also found that costs, student opinions, ratings, and reviews were important to teachers [7]. When teachers use COTS games in the classroom are usually left to make the necessary connections to the curriculum and the learning goals and outcomes themselves. Unfortunately, many games designed for educational purposes still do not contain the kind of support that is typical of, say, textbooks. Teachers need a way to evaluate the potential of a game before actually committing to using it in the classroom - in other words they need predictive assessment models.

3. The Four Pillars of Educational Games

The following model allows teachers and others to perform a structured analysis of both COTS and serious games to assess whether or not a game has potential for use in the classroom.

These four pillars are:

1 **Gameplay** - This pillar considers the game itself: How is it as a game? Is it fun? Is it Interesting? How does it measure up esthetically?
2 **Educational Content** - This pillar is the one that addresses the learning component. Are there one or more recognizable educational objectives to be found in this game that are discernible either from the game itself or from the accompanying support materials?
3 **Teacher Support** - An often overlooked aspect, the third pillar has to do with the nature of the support available to teachers who might want to use this game. Is there adequate teacher support to make it viable for use in a formal setting?
4 **Balance** - Finally, this section examines the game through the lens of the Magic Bullet model to see how well the various learning elements are balanced.

Together these four pillars highlight the key issues associated with the use of games in the classroom. There are a total of twenty elements to consider, each of which is given a maximum value of 5 to approximate a lichert scale. This means that the end result of the analysis will be a number between 0-100. This total is also shown as a number out of 5 (i.e. the total score divided by 20). These twenty elements are sorted into the four pillars. Each pillar is given its own rating as well as shown in Figure 1.

Game Review V6		
<enter name of game here>		

Overall Rating	OK [3]	2.6	50	/	95

Summaries

Game Overview	3.0	15 25	Educational Overview	2.5	35 70
Gameplay	4.0	12 15	Teacher Support	2.5	10 20
Art & Audio	1.5	3 10	Educational Content	2.5	15 30
			Magic Bullet Rating	2.5	10 20

Game Overview	3.0	15 25	Educational Content	2.5	15 30
Content & Originality	5 /5		Instructional Strategies	5 /5	
Game Mechanics	4 /5		Instructional Design	4 /5	
Game Progression	3 /5		Objectives	3 /5	
Artistic Design	2 /5		Integration	2 /5	
Setting & Characters	1 /5		Accuracy	1 /5	
Audio	-1 /5		Assessment	0 /5	

Teacher Support	2.5	10 20	Magic Bullet Rating	2.5	10 20
Guides	1 /5		Overall Balance	1 /5	
Plug N' Play	2 /5		Can Learn vs Must Learn	2 /5	
Resources	3 /5		Operational vs Educational Learning	3 /5	
Community	4 /5		Educational vs Discretionary Learning	4 /5	

NOTE: Items marked -1 (n/a) are not counted

Figure 1 4PEG Summary Scores

Elements that are missing fall into two kinds: either they are irrelevant and the fact that they are missing doesn't matter, or they should have been there, but weren't. It is important to distinguish between missing items and those not applicable to this game. Something that is not applicable should not count against the overall score. For example in a game built for hearing impaired players the audio aspects should not count for or against the game. Even though these elements might be missing, they should not be rated as missing but n/a. When it comes to the scoring of that pillar only those elements that are applicable will be counted. Any item marked n/a does not count into the total. This is why the scores out of 5 are useful in each pillar as well as overall - that number can be compared to the scores of other games, regardless of whether or not all twenty elements were scored.

3.1 The Gameplay Pillar

The first pillar has to do with gameplay. The gameplay ratings are intended to assess the quality of the game as a game independent of its potential as an

educational object. This pillar contributes 30% to the overall rating, because the value of the game *as a game* is still very important even for an educational one. There are many ways to design interactive educational experiences that are not games. Like many things, good games are challenging to design well so if it's not much good as a game, then it might as well have been designed as something else. Even worse, a good lesson wrapped up in a bad game doesn't help anyone. The gameplay pillar is rated according to six different aspects. Each one is intended to provide information about the game that will indicate whether it is likely to be a good fit for our purposes.

Content & Originality: Are the game elements well developed and appropriate for the game? Does it follow accepted norms for the genre? ~OR~ Does it have new take on known genre? Players quickly learn to expect certain kinds of things from certain genres of games.

Game Mechanics: What can you do in the game? Are the controls logical and easy to use? The amount of time it takes to learn the game relative to the amount of time one can play also affects its usefulness. A short form game should be very easy to learn. Frustrations in learning the game very quickly begin to interfere with educational learning goals.

Game Progression: The transitions between levels (which need not be traditional) should go from simple to challenging and should be smooth and appropriate for the game.

Artistic Design: Is it overall visually attractive? Often educators think of big-budget games with rich realistic characters and assume that a game must have the same artistic qualities in order to be good. Independent games and educational games rarely have budgets that allow for those kinds of graphics, but they should still be visually pleasing. Does it seem to be appropriate for the game? Whatever the artistic style, it should be a good fit for the kind of game it is.

Set, Settings, Characters & Costumes: Does it seem to be appropriate for the game? Here again it is important that the appearance of the surrounding environment as well as the characters in the game all work in harmony with each other to create a coherent experience.

Audio: Does it seem to be appropriate for the game? Linear media like film and television have a known length, but games can go on and on. As a result any music in a game will have to loop somehow, or be procedurally generated. Most games have scores that loop, and these can become irritating rather quickly - especially to people around the game who are not actually playing. Is the music necessary to the game? Can it be easily switched off?

3.2 The Educational Content Pillar

The next pillar is Educational Content, which also carries a weight of 30%. This pillar addresses the quality and extent of the educational potential in the game. Normally, this should be considerably easier to do in a game designed specifically for education, but this is not always the case.

Instructional Strategies: An instructional strategy is a plan for what will happen during the course or lesson. When applied to a game this category tries to identify what kind of strategy is being used in the game to help people achieve the ELOs. This relates to gameplay, but is specifically focused on how well the gameplay matches the intended ELOs. For example, a guessing game or drill and practice may be appropriate for learning anatomy, but not for Mendelian genetics. Are the instructional strategies appropriate for the learning outcome(s)?

Instructional Design: Instructional design is the process of creating instruction through the analysis of learning needs and the systematic development of learning materials. Because of its relative simplicity, it was decided to use David Merrill's 1st Principles of Instruction [9] as the benchmark against which the instructional design component of the game would be measured. This category assesses whether the design is in keeping with Merrill's 1st Principles of Instruction.

- **Problem**: Does it engage learners in solving real-world problems, or if not, are the problems ones that can be applied to real life problems?
- **Activation**: Does it activate existing knowledge as a foundation for new knowledge?
- **Demonstration**: Does it demonstrate new knowledge to the learner?
- **Application**: When the player learns something new in the game, is she provided with opportunities to use it within the game or is she simply given 'factoids' as an afterthought?
- **Integration**: Does the game help learners integrate that new knowledge into the learner's world?

Objectives: This category rates the extent to which the game supports the objectives that have been determined by the rater. Since we are not testing this game in the classroom, we are making a reasoned judgment here. Even with classroom testing, it is not possible to guarantee that any particular objective will be met in an educational object like this, but it is

possible to assess whether or not it provides the necessary 'raw materials'.

Integration: In a serious game it is essential that the desired learning outcomes be part of the required interactions of the game. Does it pass Becker's Lazy Test (BLT)? This is a test to determine if it is possible to get through the game by brute force or by random chance. It should not be possible to get through the game while ignoring the learning objectives. A game that passes Becker's Lazy Test fails as an educational game.

Accuracy: Does the game contain accurate information? Most of the time we would want those parts of the game that relate to our ELOs to be accurate. Even though no game can be completely accurate, it is crucial that all of the facts associated with the ELOs be correct, and that the needed concepts and principles are clear. While factual inaccuracies can create teachable moments, they should not be a surprise.

Assessment: Most popular commercial games are already pretty good at doing assessment. Some games use a simple score but many have a fairly complex set of measures that players can use to determine their standing. However, in some games a losing score can sometimes be just as valuable for meeting educational objectives as a winning one, and the actions required to achieve a positive score in some games may have little to do with what players are supposed to be learning [10]. Just like the other categories in this pillar, it is important to remember that the score tells us about the connections of the game to our ELOs. A low score does not always indicate that the game is not suitable, but it does however help us to understand what we will need to do in order to make use of this game in a classroom setting.

3.3 The Teacher Support Pillar

The second to last pillar addresses the level and quality of the teacher support that is easily accessible, either with the game itself, or elsewhere. Information on how to use games in the classroom can sometimes be very difficult to find, and it really helps no-one if there is a stellar teacher's guide that no-one can find. It contributes 20% to the overall score. The following are elements to consider when assessing the quality of the teacher support for a game.

Teacher's Guide: Teacher's guides should be clear and easy to locate. Support materials should include such things as:
- A description of game play.
- Content description (documentation) that is well organized.

- Any required special permissions/skills to install or run that are clearly identified.
- Installation and execution processes that are clearly identified and easy to read and follow.

You should be able to see how the game will play.

Plug N' Play: Does it include lesson plans with thorough instructions for using them in the classroom (or other target environment)? It should not require a large time investment to make it "teacher-ready".

Supplementary Resources: This includes any additional information specifically for teachers, such as background on both the game and the topics it addresses, ways to use the game and where to get help. For full points, these resources should be complete and readable to your satisfaction.

A community: Does a community exist where teachers can go for help, support, and to share ideas, experiences, and ways to use the game?

3.4 The Balance Pillar

The final pillar relates to the Magic Bullet model and contributes the final 20% towards the overall score. This model was originally developed while analyzing several strictly commercial videogames using another methodology I devised known as instructional ethology [11]. In the process of producing extensive gameplay logs it became apparent that one perspective for looking at videogames is from the point of view of what players are learning as part of the game experience. It turns out that all learning in and around a game can be classified into four broad categories. It is known that not all learning in a game is necessary to win, although some always is. It is also true that sometimes learning occurs that was never intended by the designers, while other times players learn things outside the game that help them inside the game.

So it is that all learning in games can be classified as (non-exclusive) members of at least one of these sets. In the process of trying to show the relationships between these different sets, several visualizations of the interrelationships of these four sets were created, and the final picture ended up being somewhat bullet-shaped. Thus, it earned the moniker "Magic Bullet".

The four categories of learning are as follows:
1 **Things we CAN learn** - as deliberately designed by those who created the game.
2 **Things we MUST learn** - this will almost always be a subset of the first category, and includes only those items that are necessary in order to win or get to the end.
3 **External Learning** – This category includes learning that happens outside of the game: in

fan sites, and other social venues. This category also includes 'cheats'.

4 **Coincidental Learning** - other things we can learn. These are not necessarily designed into the game, although sometimes designers may hope that players choose to take these up.

When used in an educational context, this model includes an additional layer that is specific to educational contexts [12], so each of these categories becomes a sub-category of one of the following contexts:

Operational Component - Game controls & some mechanics and other necessary overhead. How much is reasonable depends to some extent on how the game will be used.

Educational Component - How does the teacher plan to use the game? This is the critical piece.

Elective Component - These are any other elements of the game that don't fit into one of the other two categories.

This pillar makes use of the Magic Bullet to consider four key perspectives and thus there are four categories. In each of these categories the rating is determined by how well the balance of the elements fits the type of game it is, its intended use and audience.

Overall Balance: This category looks at the overall balance of the learning in the game in light of the intended application. Here we seek to answer the question of how well the relationship between the 4 main categories matches its intended use.

Can vs. Must: Is the balance of things we can learn versus things we must learn appropriate for the intended use of the game?

Operational vs Educational: How much of the time spent in the game has to do with just learning to play the game? Is the required operational learning appropriate for the game's intended purpose?

Educational vs Discretionary: Finally, is there an appropriate balance of learning and fun? We need to keep in mind that we are talking about games - and they should be at least engaging even if they are not exactly fun.

4. Conclusion

Our understanding of educational game design is still evolving, and one of the great challenges is how to combine approaches to game design with approaches to instructional design in a way that results in a game that works both as a game and as instruction. Another common problem happens when a game is built around a single mechanic when that one mechanic represents an action out of context from the concept/skill to be learned.

It is important to remember though that final numeric score that represents the sums of all the other scores is not to be viewed like an exam score. A low score doesn't necessarily mean that the game has no potential as a learning tool. The point of the variety of categories in each of the pillars is that it provides us with a detailed, but straight-forward analysis of the game's strengths and weaknesses. A low game score may be tolerable if it has a high content and support rating. Alternately, a low support rating may not be important to you if you have the experience and the time to develop your own materials. This is a subjective rating tool, so it is important to include additional comments whenever possible. Once a game has received a sufficient number of ratings from various sources it will become possible to see trends in the evaluations. Wildly different ratings in the same category would indicate that this part deserves a closer look - perhaps the evaluators had different goals in mind. On the other hand the more consistent the ratings, the more confidence we can have.

All games involve learning. According to Raph Koster, author of "A Theory of Fun for Game Design" enjoyment in games triggers the release of endorphins, which actually comes from learning. "One of the subtlest releases of chemicals is at that moment of triumph when we learn something or master a task. This almost always causes us to break out into a smile. After all, it is important to the survival of the species that we learn - therefore our bodies reward us for it with moments of pleasure. …. Fun in games arises out of mastery. It arises out of comprehension. It is the act of solving puzzles that makes games fun. In other words with games, learning is the drug." [13] When asked why players continue to play a game after they have beaten it, they will often say they do this because there are still more things to discover – more things to explore, different endings, and so on. In order to take best advantage of that games have to offer, we need a variety of ways to assess them.

5. References

[1] M. P. J. Habgood and S. E. Ainsworth, "Motivating Children to Learn Effectively: Exploring the Value of Intrinsic Integration in Educational Games," *Journal of the Learning Sciences,* vol. 20, pp. 169-206, 2011/04/19 2011.

[2] J. P. Gee. (2012, June 20, 2012). Keynote: Big-G Games are Good for Learning. *9th Annual*

Games fro Change Festival (G4C). Available: http://gamingandeducationengagementinlearning.com/2012/07/25/james-gee-says-that-big-g-games-are-good-for-learning/

[3] C. Heeter, "Games 4 Learning Course Notes," Michigan State University, Department of Telecommunication, Information Studies, and Media2013.

[4] W. Dick, L. Carey, and J. O. Carey, *The systematic design of instruction*, 5th ed. New York: Longman, 2001.

[5] P. Pirnay-Dummer, D. Ifenthaler, and J. M. Spector, "Highly integrated model assessment technology and tools " *Educational Technology Research and Development,* vol. 58, pp. 3-18, 2010.

[6] P. Felicia and S. Egenfeld-Nielsen, "Game-Based Learning: A Review of the State of the Art," in *Serious Games in Education : A Global Perspective*, S. Egenfeldt-Nielsen, B. Meyer, and B. H. Sørensen, Eds., ed Aarhus, DNK: Aarhus University Press, 2011, pp. 21-46.

[7] L. M. Takeuchi and S. Vaala. (2014, Oct. 24 2014). Level up learning: A national survey on teaching with digital games. Available: http://www.gamesandlearning.org/2014/10/21/level-up-learning-captures-state-of-digital-games-in-classroom/

[8] J. Kirriemuir, " A Survey of COTS Games Used in Education," in *Serious Games Summit/Game Developers Conference*, ed. San Francisco, CA, 2005.

[9] M. D. Merrill, "First Principles of Instruction," *Educational technology research and development : ETR & D,* vol. 50, pp. 43-60, 2002.

[10] K. Becker and L. Gopin, "Selection Criteria for Using Commercial Off the Shelf Games (COTS) for Learning," in *Learning and Education Games*. vol. Volumne Two: Bringing Games into Educational Contexts, K. Schrier, Ed., ed Pittsburgh, PA: ETC Press, 2016.

[11] K. Becker, "Instructional Ethology: Reverse Engineering for Serious Design of Educational Games," in *Future Play, The International Conference on the Future of Game Design and Technology*, Toronto, Ontario, Canada, 2007.

[12] K. Becker, "A Magic Bullet for Educational Games," in *CNIE Green Aware 2012 – The Canadian Network for Innovation in Education*, ed. Canmore, Alberta, 2012.

[13] R. Koster, *Theory of Fun for Game Design*. Scottsdale, AZ: Paraglyph Press, 2004.