The Serious Games Showcase and Challenge: Learning from the Evaluations

Jennifer McNamara
BreakAway Games
Hunt Valley, MD
jmcnamara@breakawayltd.com

Dr. Peter Smith
Katmai in Support of the ADL Initiative
Orlando, FL
peter.smith.ctr@adlnet.gov

ABSTRACT

Since 2006, the Serious Games Showcase & Challenge (SGS&C) has encouraged student, government and commercial game developers to submit their serious games for review by a panel of military, academia and industry gaming experts. For the SGS&C, an entry is considered a serious game if it incorporates game play dynamics in a product to educate or train a learner at any stage of the learning continuum from middle school through adult. A wide range of information is collected when a product is submitted to the SGS&C and as each entry navigates through the evaluation process. In 2012, we analyzed and presented data provided in the submission forms, videos and playable games provided by entrants. This data gave us insight into the technologies, subject matter addressed and techniques developers have applied to their existing projects (McNamara, Smith, Gritton and Smith, 2012). However, the original paper stopped short of sharing data regarding what evaluators found important in the design of a best-in-class serious game. This paper analyzes the data provided annually via the SGS&C standard evaluation rubric and unstructured comments provided by the SGS&C evaluator pool, comprised of over 70 professionals representing users, developers, and researchers of best-in-class serious games. Previously, the rich evaluative data collected has only been shared with individual entrants to provide them feedback from the process and possibly inspire improvements in their games. By analyzing the evaluator comments across a series of dimensions, we've identified practices that are relevant to anyone interested in the application or development of serious learning games. This paper maps evaluator comments to identify successful instructional interventions, design principles and technology factors that translate into best practices for serious games designers and developers.

ABOUT THE AUTHORS

Jennifer McNamara is Vice President of Serious Games at BreakAway, Ltd. providing serious games serving the corporate, defense, homeland security and medical communities. Jenn holds a B.S. in Cognitive Psychology from Drexel University and a M.Ed. in Instructional Systems from The Pennsylvania State University. Jenn strives to help professionalize and advance the state of the serious games industry through serving as the Industry Lead for the Serious Games Showcase and Challenge IPT in addition she serves as the coordinator of Special Events for I/ITSEC 2013 and member of the technology and standards subcommittee for the Society for Simulation in Healthcare.

Dr. Peter Smith is the lead for emerging technologies at OSD's ADL Co-Lab. He is responsible for heading up efforts in games, virtual worlds, and social media. He has a rich background in serious games with both research and development efforts ranging from massively multiplayer online games to small scale web games and everything in between. Peter earned his PhD in Modeling and Simulation from the University of Central Florida. He was one of the founders and current IPT member of the Serious Games Showcase & Challenge held at the I/ITSEC conference, and is very active in both the serious games and games for health communities.

The Serious Games Showcase and Challenge: Learning from the Evaluations

Jennifer McNamara
BreakAway Games
Hunt Valley, MD
jmcnamara@breakawayltd.com

Dr. Peter Smith
Katmai in Support of the ADL Initiative
Orlando, FL
peter.smith.ctr@adlnet.gov

BACKGROUND

The Serious Games Showcase and Challenge (SGS&C) provides a venue to educate members of the education, training and simulation communities on the potential of serious learning games. The SGS&C venue also "showcases" serious learning games while challenging developers to compete for the title of best serious game across a series of categories. To qualify for the SGS&C, serious game entries should incorporate game play dynamics in a product to educate or train a learner at any stage of the learning continuum from middle school through adult. A previous paper (McNamara, Smith, Gritton and Smith, 2012) presented in depth background on the origin of the challenge and the results of an analysis of trends in serious learning gaming based upon the first six years of the showcase and challenge. This paper will focus on sharing information about the evaluation process, evaluators and common noted strengths and criticisms of serious learning games. While the discussion here is based solely on the evaluation of the finalists in the SGS&C, we believe it is representative and applies broadly to the serious games space. Further, the SGS&C's expert evaluators, worthy incentives and a marketing/awareness campaign attract competitive entries that are representative of the industry.

EVALUATION TEAM

The evaluation team is comprised of two parts. Both have an essential role in curating the best possible showcase and ensuring the best contributors are recognized. The evaluation team includes the SGS&C Steering Committee integrated process team (IPT), and the Evaluator Panel. Additional evaluation committees are stood up for various special awards including the People's Choice Award and the Adaptive Force Award.

The SGS&C Steering Committee IPT

The approximately 30 Steering Committee IPT members are volunteers. The IPT is comprised of approximately 45% government members and 55% industry representation. This group is responsible for maintaining the integrity of the challenge and showcase while managing all aspects from soliciting sponsors and submissions, marketing the event, the submission and evaluation processes and the awards ceremony. With regards to the evaluations, this group narrows the submissions to the finalists during the annual down select meeting.

The Evaluator Panel

To preclude over-weighting any one assessment area based on community biases, evaluators are recruited with with the intention of balancing between the developer, user, and research communities. This balanced approach to evaluation has strengthened the credibility of the SGS&C. The formal SGS&C evaluation team is comprised of a multi-disciplinary team of individuals from government, academia and industry. The evaluation team includes educators, developers (artists, engineers, and designers), program managers, business developers, and senior leaders with a wide range of perspectives for how these technologies are used around the world. To provide an example of the evaluators' backgrounds, the 2012

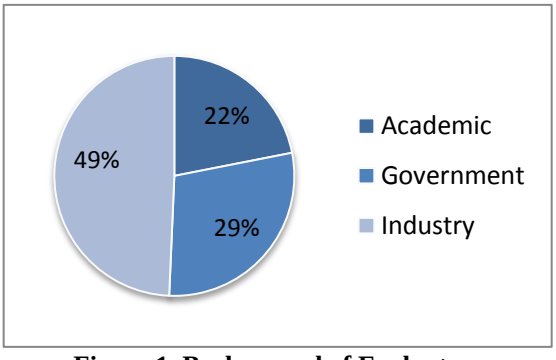

Figure 1. Background of Evaluators

pool of evaluators will be discussed. It included 16 people from academia, 21 people from government, and 36 people from industry for a total of 73 evaluators, see Figure 1. Twenty-three percent of the evaluators were women. No specific data was collected on the age of the evaluators, but ages spanned from early twenties through early sixties.

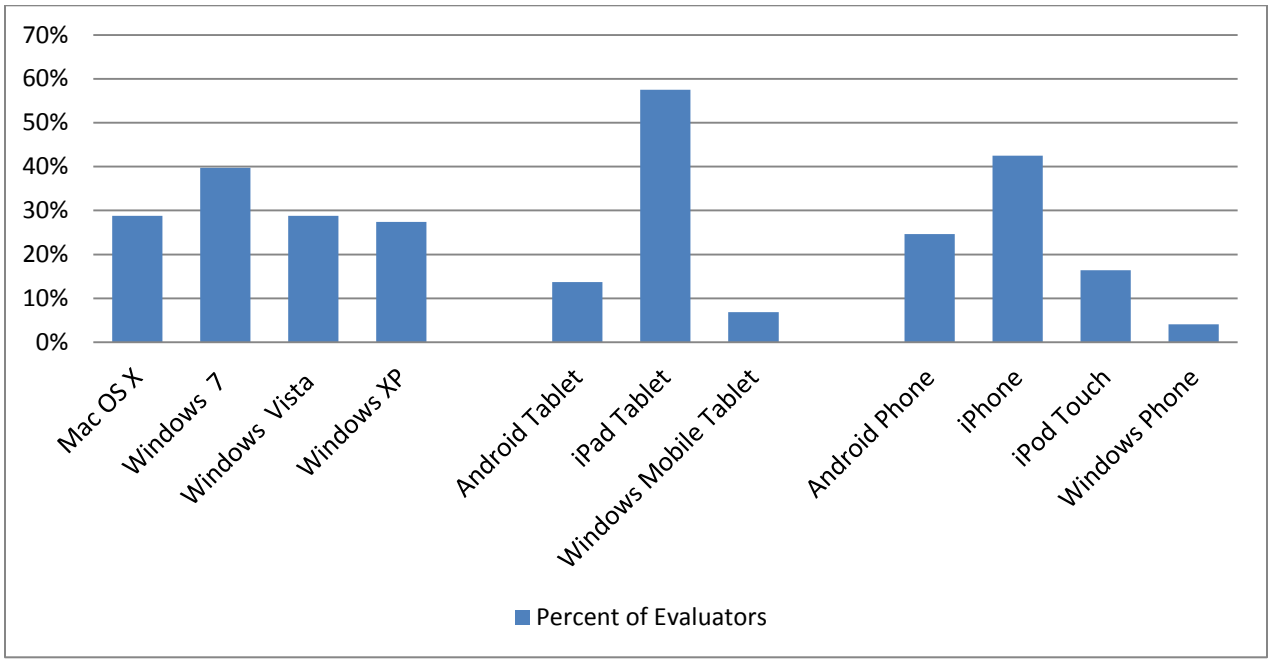

Figure 2. Evaluator Desktop/Laptop Operating Systems, Mobile Tablets and Mobile Phones

Part of the challenge for a competition judging serious games is that the games run on a variety of computer hardware and software platforms requiring many computers and devices for judging the games. To facilitate judging games on these diverse platforms appropriately, we conduct an annual technology survey of the evaluator pool. Figure 2 serves as an interesting survey of the evaluators' computers and mobile devices in use within the I/ITSEC community. Note that the percent totals will not add up to 100% because people can have more than one device or none.

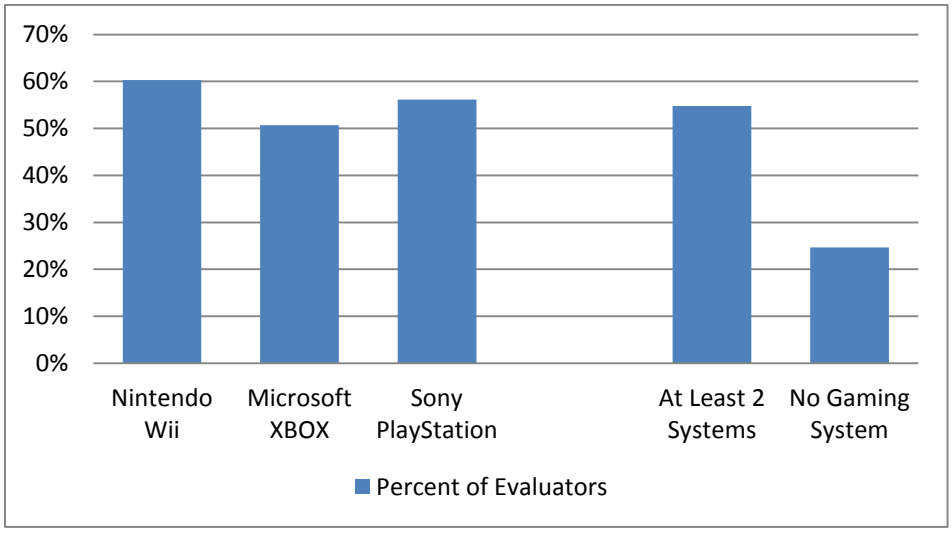

Figure 3. Evaluator Gaming Systems

A majority of the evaluators use a computer with a Windows Operating System (96%) spread across Windows 7 (40%), Windows Vista (29%) and Windows XP (27%). Twenty-nine percent use a Mac Operating System. With

increased customer desire for mobile games, and so many games designed for mobile devices, developers are always challenged to select devices for which to develop. In our evaluator pool, the iPad (58%) and iPhone (42%) were the most common devices followed by Android tablets (14%) and phones (25%) with very few Windows mobile users (7% tablet and 4% phone). Some serious games target gaming systems. In our evaluator pool, the Nintendo Wii was used by the largest number of evaluators (60%) followed by the Sony PlayStation 2 and 3 combined (56%) and Microsoft XBOX and XBOX 360 combined (51%). Note that 25% of the evaluators don't have any gaming systems and 55% have at least 2 systems. Of those with at least two systems, 88% included a Wii among the systems (see Figure 3). According to the Entertainment Software Association's data, 51% of U.S. households own a game console, and those that do own on average 2 consoles (ESA, 2013).

The People's Choice Award

One of the available awards, The People's Choice Award, is judged by the attendees on the exhibit floor that visit the SGS&C booth in the STEM Pavilion to play the games and vote for their top serious game selection from among the finalists.

EVALUATION PROCESS

The Evaluation process begins when games are submitted to the SGS&C. The submitting organization is asked to provide information about the game including: a description of the game; the standout features of the game; the primary audience; goal or challenge; how skills/ideas are taught and how they are measured; and how feedback is provided. The responses provided to these questions are a key piece in the game's evaluation. These responses are shared with the IPT for use during the finalist down select meeting and to the finalist evaluators. All reviewers are asked to judge the game with regard to the stated goals and objectives.

Evaluation Rubrics

A well-designed evaluation rubric is crucial to attract expert evaluators and ensure a challenging and fair competition. Since serious games run the gamut of training objectives and game methodologies, designing a single evaluation rubric capable of assessing all entrants equitably is crucial. The essence of a well-designed evaluation rubric is simplicity. At its' core, a serious learning game is a game merged with instructional elements. Accordingly, the rubric concentrates on three main areas: the physical operation of the game, the enjoyment of the game, and its ability to impart the intended training. The SGS&C evaluation process collects both qualitative and quantitative data about each entry. As detailed above, as each entry is submitted, information is collected about the instructional goals and objectives of the application. This information is collected via free-form responses on the entry form and provides insight into the target audience for the application, the technologies used to develop the application, the various instructional feedback/remediation strategies used within the application, and how the game is designed to meet the stated instructional objectives. As entries are received, this information is used to populate the rubrics, and the evaluators are asked to evaluate each game with regard to this information.

Assessing the pedagogy of training objectives as diverse as culture and language, aircraft maintenance, and animal husbandry requires each entrant to clearly articulate the training objective of his or her game. Armed with the goals for the serious game, the rubric assesses how well it achieves the entrant's stated training objective while employing key education and training principles (i.e., after action review, real-time tutoring, remedial training). A final aspect of the rubric covers the unexpected – extra bonus points for innovative aspects of the game within any of the aforementioned criteria.

Measures were put in place to further ensure equitable assessments. To guard against individual evaluator biases, all assessment areas are refined with anchor statements so evaluator deviations will be minimized. Thus, the mixture of objective criteria and anchored subjective measures provide a rubric able to competitively select clear winners without controversy.

A final aspect of the rubric which has afforded great insight into the evaluator's likes/dislikes of a serious game is the free-form comments section. These objective comments have become an excellent source of constructive criticism for the game developers. All comments are collected and collated at the conclusion of the SGS&C, and forwarded to the entrants (less evaluator's name) as feedback for review. Many of the serious games developers reported that this feedback is incentive enough for becoming a finalist.

The Down Select

The first gate for all games consists of a systematic vetting (called the "down select") to select the finalists. The down select judges are the approximately 30 SGS&C Steering Committee IPT members. The down select rubric is used to consistently evaluate all entries submitted to the challenge across the appropriate submission/award categories. This rubric provides a standard mechanism to measure the production quality, playability, engagement and instructional effectiveness of each entry while also identifying any innovative qualities a particular entry incorporates into its design. The down select rubric is populated during a two to three day meeting where all submitted entries are played, discussed and reviewed. As the committee plays and evaluates each entry, scores are entered into an electronic spreadsheet and comments are collected. These scores are used to determine the finalists for each SGS&C category. Once selected, finalist games and the evaluation rubric are forwarded to the worldwide corps of evaluators.

The Finalists' Evaluation

After finalists have been identified, they are assigned to the SGS&C evaluators using a complex matrix ensuring no conflicts of interest between the submitters and evaluators. The evaluation rubric used during this stage of the evaluation establishes a clear and concise score based on how well the entry meets the established criteria that comprise the overall categories of instructional quality, technical quality, and usability/playability. Each evaluator is assigned three to four games to evaluate, but given access to all finalist games, and asked to evaluate as many as possible. Evaluators provide their completed quantitative scores and free response comments to be shared with the developers. At the conclusion of a six-week evaluation cycle, all evaluations are independently tabulated by three people with controls in place to ensure a single high or low score can't unnecessarily bias a game's final score. The winners of the various serious games categories are announced at the SGS&C awards ceremony.

Special Awards

Historically, two special awards have been judged outside of the normal evaluation team, these are the People's Choice Award and the Adaptive Force Special Emphasis Award. The People's Choice is judged by the conference attendees who visit the SGS&C booth in the STEM Pavilion to play the games and vote for their favorite game. Votes are closed and tallied the night before the conference closes in order to announce the winner at the Awards Ceremony on the morning of the last day of the conference. The Adaptive Force Special Emphasis Award finalists are forwarded to a panel of evaluators at the Office of the Secretary of Defense (OSD) for judging and winner selection. Beginning in 2013 there is also the Students' Choice Award through which games targeting students are to be judged by age appropriate high school students.

QUANTITATIVE EVALUATION DATA

Scoring data available includes the final average score for each game and each evaluator's individual score for each game for the years from 2009 through 2012. Here, a discussion of the average scores by game will be shared to give an idea of the range of scores achieved using the evaluation rubric. The rubric allows a total score of 100 points with an additional 15 innovation bonus points available for a maximum score of 115 points. Across the 4 years of data, the average game score was 74 points with the lowest score of 47 points and highest score of 95 points demonstrating that the rubric produces meaningful stratification amongst the games. There were no meaningful trends in terms of quantity of free response data corresponding to high versus low scoring games, i.e., the evaluators do not seem to provide more (or less) feedback based on the overall quality of the game.

QUALITATIVE TRENDS IDENTIFIED THROUGH THE EVALUATION PROCESS

The free response comment data compiled and analyzed for this discussion includes 200 pages of evaluator comments collected across all finalist games from 2007 through 2012. These comments were collected to provide additional evaluation information and to be presented back to the developers after the challenge. Each comment was read and categorized into relevant topic areas. Unless otherwise noted, these topics are non-directional. For example, the topic graphics contains all comments related to graphics regardless of whether the comment suggested the graphics were of high or low quality. Given the varying opinions of individual evaluators, the topics covered in the

comments varied greatly within and across games, and often comments contained more than one topic area. There were 773 unique comments across the six years. These represented 2,279 topic mentions for an average of nearly three topics per comment. Comments also varied greatly in length with the longest comment being over 3 pages long and shortest comment being a single, one-line sentence. Figure 4 shows the percentage of comments that related to each topic across all six years of data analyzed. Following these topics year to year, and evaluating the content of the comments, trends began to emerge. While not every topic was represented year to year, several overarching trends were present throughout every year of the SGS&C: instructional interventions, design principles, and technology. These groupings can be seen in Figure 4 represented by the dark, medium, and light coloring respectively, and will be discussed in greater detail below.

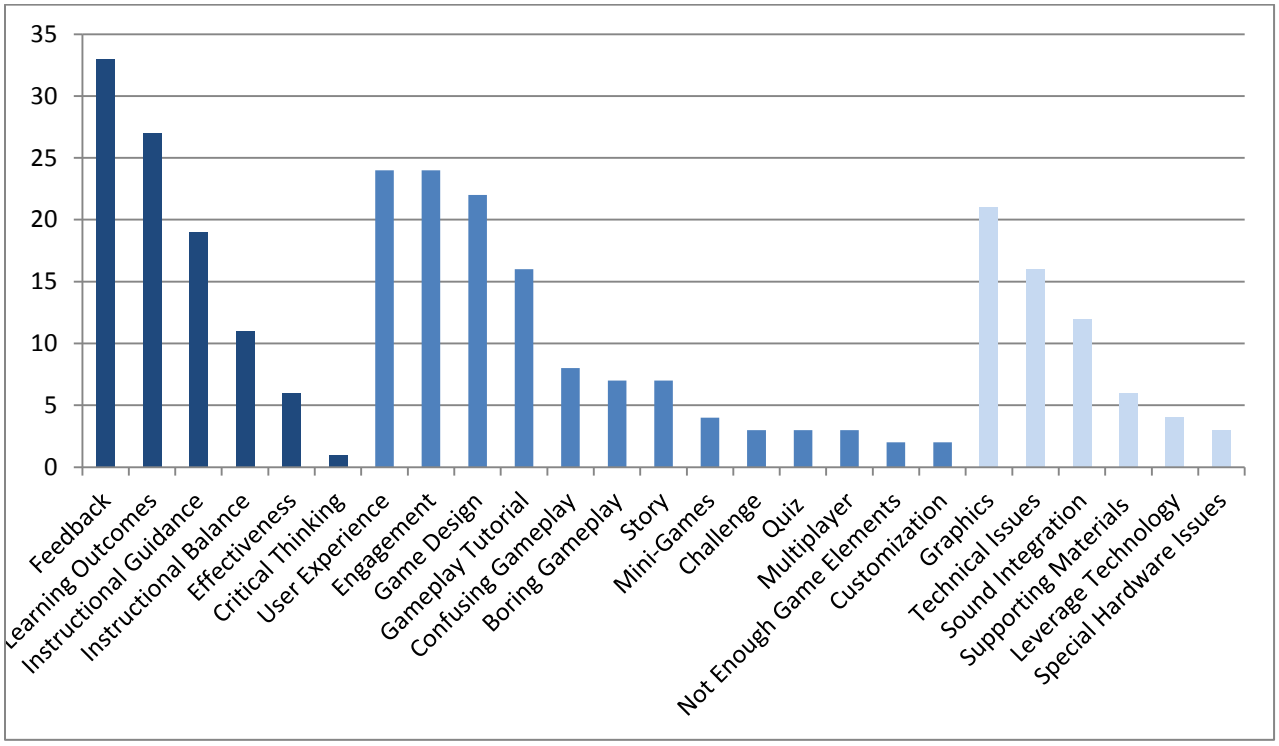

Figure 4. Percentage of Comments by Topic

Instructional Interventions

As the SGS&C is focused on serious learning games, instruction plays an important role in all finalists' games and the evaluators frequently provided feedback related to instructional interventions utilized within the games. The instructional elements most frequently discussed included: feedback, learning objectives, instructional guidance, instructional balance, effectiveness and critical thinking. Different aspects of instructional interventions are discussed further below.

Feedback

It is interesting to note that player feedback is consistently the most mentioned topic in the comments. It is common to think of score as the obvious feedback mechanism in games, but the evaluators are looking for more. Comments suggest that games should ensure that players know when they have done something in the game and how their actions have changed the game world. They expect that feedback should drive learning by clearly indicating that players have learned something or have committed an error. Feedback should also be used for remediation from errors. On average feedback is mentioned in 33% of all comments. Table 1 identifies the number and percentage of comments related to feedback.

Table 1. Comments Related to Feedback

	2007	2008	2009	2010	2011	2012
Feedback Comments	13	28	44	40	75	57
Percent out of Total Comments	56	41	38	29	28	33

While the percentage of comments related to feedback is going down annually, the total number of comments mentioning feedback is on the rise indicating that it is not diminishing in importance. It seems that more and more evaluators understand that feedback can make or break serious games. Comments are also transitioning from merely noting the need to add feedback to more specific ideas on how to implement feedback. For example, evaluators suggest that feedback needs to be explicit and consistent in presentation media and timing. In addition, specific comments suggest that feedback should not be reserved only for an After Action Review (AAR) provided after the game is completed. Feedback should be given immediately as an action is taken and guidance should be provided on how to leverage this feedback to be successful. AARs are also supported by the evaluators as a way to provide more in depth feedback and provide an opportunity to reflect on what was learned after the game session has completed.

Learning Outcomes and Effectiveness

Having a specific and identifiable learning outcome is consistently considered important with mentions comprising 24% of the comments. Learning outcome comments come in both positive and negative forms. The positive comments are comprised of evaluators expressing having gained a complete understanding of what is being taught by the game and feeling as though students would be able to gain the appropriate knowledge by playing. Negative comments involve the evaluator not understanding what is being taught in the game, or understanding it, but believing that students will not be able to learn from playing the game. It is important for serious game designers to consider what lessons they need to impart and how they plan to impart them.

Effectiveness is mentioned in 6% of all comments and refers to how successfully a game achieves stated learning outcomes. In order for a serious game to be effective, the game must have explicit goals for the learning and well-structured learning objectives that are easily understood and clearly connected to the overall game. In general, a serious game with strong effectiveness achieves a higher level of success than one which merely addresses learning outcomes.

Instructional Guidance

Instructional guidance is represented in 19% of all comments. Instructional guidance is exemplified in the need to bring learners, that are not experts in the subject matter, into the game as easily as possible. Allowing a low starting level of subject matter understanding helps learners successfully progress through the game without feeling intimidated by the subject matter. Some instructional strategies for initial instructional guidance may include presenting a series of well scaffolded levels that ease a player into the learning experience. as Another method to help the learner is to augment the game with external materials, such as a book or video that helps guide the player through the required instruction.

Instructional guidance does not stop with bringing the player into the game, but remains important throughout the game. Players should be provided with guidance and feedback throughout their learning experience with a serious game. This guidance might be in the form of imbedded literature, like a guidebook or diary for the player, or be provided by non-player characters (NPCs). NPCs are a crucial part of providing instructional guidance, feedback, and narrative throughout many serious games.

Instructional Balance

Instructional balance is mentioned explicitly by the evaluators 11% of the time. Good instructional balance implies that a game is equally engaging and instructional. Instructional balance is possibly the single most important issue regarding serious games. This concept is tied into almost every facet of the design and instructional categories in our analysis – thus this 11% taken alone does not fully represent the importance of the topic in general. A perfectly balanced game is the holy grail of serious games. Very few games get instructional balance perfect. The ways in which games fail to achieve balance will be discussed more in the design sections below. Games that are not well balanced are either considered entertainment games with instruction sprinkled in or are criticized as being more of a simulation than a game.

Serious game developers should consider the implications of this issue during the design of their games. If balance can't be achieved, erring towards instruction over gameplay is probably the best strategy to ensure learning objectives are met. For those developers that have managed to get balancing right, the payoff is worth the work, both as reflected in the evaluators' comments and in the learning achieved by the players.

Design Principles

Design principles are the second major category by which we classify the comments. Design is how everything in a serious game is intended to be implemented including instruction. As such, many of the issues related to design will also have direct effect on the instructional effectiveness of a serious game. Players of a serious game should be learning the intended learning outcomes and not spending too much time learning how to play the game. The most appropriate design solution is to have a smart level of design that imparts learning outcomes through learning the game system, as discussed in the instructional balance section above. These same issues can be viewed through the mirror of gameplay balance. The medium hued group in Figure 4 includes the topics that are considered important to design principles including: user experience, engagement, game design, gameplay tutorial, confusing gameplay, boring gameplay, story, mini-games, challenge, quiz, multiplayer, not enough game elements and customization. These are discussed below in the following categories: game design, engagement, and emerging gameplay features.

Game Design

Game design is the third largest area for design principles referenced in 22% of all comments. Game design dictates the system in which the game operates specifying the content, gameplay and rules, game environment, storyline, and characters. It has a part to play in instruction, and must be integrated fully with the technology being used. Evaluators often cite game genre standards that are exploited (or violated), and explore decisions made with game mechanics in their comments. Interestingly, evaluators will often comment about game design changes that could allow the game to cover additional material, or be repurposed for a new audience when they are making positive comments about game design.

A common negative comment about game design is that a specific game is really just a glorified quiz. Serious games are often considered analogous to glorified quizzes by the main stream game development community (Creighton, 2012). This is mostly due to experiences with older edutainment titles (Ito, 2006). The premise is that the player is just answering quiz questions and given graphical rewards for performing well. In 2009, 15% of all comments were about the games being quizzes. This has dropped off to only 1% in 2012. Many games in the past would contain gameplay that hopefully taught the material, and then followed that up with a quiz to test for learning. Across the serious learning games entries, this practice has been replaced at some level by the use of embedded mini-games. Mini-games are small games that provide a single learning objective within a larger gaming system. As seen in Table 2, mini-game comments are on the rise, while quizzes are tapering off. These mini-games have been found to provide a more engaging opportunity to assess or instruct learners within the context of a larger gaming infrastructure (Smith & Sanchez, 2009).

Table 2. Total Comment between Mini-Game and Quiz

	2007	2008	2009	2010	2011	2012
Quiz Comments	0	4	17	1	3	2
Quiz Percent out of Total Comments	0	5	15	0	1	1
Mini-Game Comments	0	1	3	2	13	12
Mini-Game Percent of Total Comments	0	1	2	1	4	7

Unfortunately, in some cases evaluators have seen that there are not enough game play elements in a finalist's game. This happens only 2% of the time on average, but was more of a problem in earlier evaluations. For example comments about not enough gameplay elements comprised 11% of all comments in 2008 and have been at 1% or less since 2010.

Engagement and User Experience

Engagement is sometimes a synonym for fun when discussing serious games. While fun should never take precedence to learning, it is still very important to serous game design. Engagement comprised 24% of all comments made. Engagement is also identified through the use of story. Story comprises 7% of all comments with closer to 10% in the last 4 years, suggesting the use of story is on the rise within the SGS&C finalists. The evaluators are also, to a lesser extent, looking to be engaged through challenge. Challenge was referenced only 3% of the time, though it is still an important feature of serious games (Quick, 2011).

The importance of user experience cannot be overlooked. Each year, it has been the third largest category by percentage of comments from the evaluators, with 24% of all comments referencing user experience. Many

comments that directly reference user experience are related to user interface. These comments highlight the importance of a well-designed interface. The user experience also covers in-game guidance, tutorials, and the need for serious game designers to spend more time reducing confusion. In particular there is a need to spend more time on designing tutorials to bring new players into a game (Salen & Zimmerman, 2004).

Gameplay Tutorials

While most entertainment games start with a simple tutorial for how to use the game system this is something many serious games lack. This is disconcerting when considering that many players of serious games have not been exposed to other gaming systems, a luxury that most entertainment games have, and are therefore having to learn both how to play the game and the content within the game.

The need for tutorials is apparent, as they are mentioned in 16% of all comments. Serious game designers should consider novice players who may not understand the vocabulary, standard devices, and procedures of a topic domain or of a gameplay mechanic. This need for general guidance is important as players are attempting to learn two things at once, i.e., the topic being taught and the game play. Making sure that they are not frustrated or confused by the game's system is paramount.

Emerging Design Features

Throughout the years certain new design features have emerged in importance as developers have learned to utilize them to their full extent. One such feature is customization which was discussed in 2% of the comments. Evaluators have good things to say about customization especially when provided the ability to customize avatars. Customization, however, can be a double-edged sword. If evaluators have the ability to change a character's clothes they may also be inclined to change a character's race, sex, or other attributes.

Another feature that is currently emerging is multiplayer capabilities. Multiplayer comments were represented in 5% of the comments in 2012, but it was not even mentioned until 2010. The reason multiplayer has been slow to emerge is technical issues related to getting games networked, and more time being spent on getting single player games right. Now that multiplayer games are starting to emerge, there was even a comment suggesting that a single player game should include multiplayer capabilities. Multiplayer games encourage players to share skills and knowledge, possibly leading to deeper learning (Gee, 2007).

Technology

The technology used in SGS&C entries has varied greatly. While historically the SGS&C has only allowed PC games, and has only recently opened to mobile devices, submitters have often included specially designed pieces of hardware and have leveraged every bit of available technology from their existing computer systems. Everything from galvanic skin response sensors, to full virtual reality head gear, the technology around the SGS&C continually impresses the evaluators.

Presentation

Leveraging the technology available on a standard device is as important as using special hardware. The overall presentation of a serious game is important and should utilize as many capabilities of a platform as necessary to achieve objectives, but at a minimum both the graphics and sound capabilities of a platform should be used. The presentation is commonly the first aspect of a serious game that the player will experience—making attention to details in this category of the utmost importance.

Table 3. Comments Related to Graphics

	2007	2008	2009	2010	2011	2012
Graphics Comments	3	18	21	25	45	51
Percent out of Total Comments	13	26	18	18	17	30

Graphics were referenced 21% of the time. Table 3 shows that graphics was referenced more in 2008 and again in 2012. It is difficult to understand why by just looking at the numbers. Reading the comments it is clear that there was a period early in the SGS&C existence where the evaluators were critical of the quality of game graphics. In 2012, the evaluators were commenting on how improved the graphics have become over the last few years. Research suggests that graphics matter a lot to players, . In one particular instance the graphics were so distractingly

bad that students felt deceived by the developers and refused to learn from the game (Elliott, Adams, Bruckman, 2002). See also Chandler et al (2009) for more discussion of fidelity in serious games. It is clear from the evaluator comments that graphics do not need to match commercial game quality, but they do need to support the theme of the game and be clean.

While it is easy to imagine that the evaluators are interested in graphics fidelity, a more surprising result is the importance of sound integration. This is the use of sounds, both sound effects and music, in a serious game. Specific comments related to a game having annoying repetitive sound or very rich, well integrated sound. In one particular instance a game was noted as being very well done but lacking any sound at all. The evaluators were quick to identify that this was mistakenly a silent game, highlighting the importance of good sound design in a serious game. Sound was mentioned on almost 60% of all games, and composed 12% of all comments illustrating a need for sound to be integrated into a serious game with attention paid that it will not become repetitive or otherwise annoying.

Leveraging Technology and Special Hardware

There are two main types of technology issues. The first are issues within the game, and the second is external. Within a game, any bugs will be found by the evaluators, no matter how obscure they are. In fact, submitting a game to the SGS&C is similar to releasing a game commercially, because the planning committee spends time understanding the capabilities of the evaluators and matching them with the best games for them. The submitter cannot know beforehand what system the evaluator will have. Technical issues are the focus of 16% of all comments.

External issues usually relate to special hardware provided by the submitters with examples including galvanic skin sensors and virtual reality headgear. Some submitters think that they will wow the evaluators with an external device. This is almost never the case. Special hardware is usually mentioned by the evaluators due to issues that are created. Special hardware issues comprise 3% of all comments, while leveraging technology is 4% (and includes good uses of both hardware and software technologies). The fact of the matter is that special hardware seems to be more of a detriment rather than a positive factor for the evaluators.

CONCLUSIONS

Over the last eight years the evaluators have provided a great service to the community. The evaluations from the SGS&C have helped the game developers improve their games, hopefully the lessons we have shared in this paper will help the larger serious learning games community improve the quality of its games and consequently improve overall learning outcomes. Here we share a summary of some best practices derived from the evaluators' comments.

Helpful Hints for Success in the SGS&C

For serious learning game developers hoping to compete in future SGS&Cs, the evaluator comments provide some helpful hints. Evaluators pay attention to external supporting materials and instructions. When submitting a game, be sure that everything is filled out and submitted along with the game. The documentation must be supported by the game design. Games where things start out too slowly may lose the evaluators interest; try to ensure you capture attention early. There will be a diverse range of game and subject matter expertise among the evaluators, so ensure the game and documentation supports game and content novices with additional materials. Mobile games should utilize special features of the hardware, not just be a computer game transferred to a mobile device.

Best Practices for Serious Learning Game Design and Development

Serious learning game designers and developers should pay explicit attention to the selection and implementation of instructional interventions, design principles and technology choices. The intertwining of learning and gameplay should receive maximum design attention. Specifically, our evaluators make the following recommendations: (1) designers must ensure that gameplay does not interfere with the instructional intervention itself, (2) instruction and narrative should not be at a higher level of knowledge as compared to gameplay, (3) game activity and learning should be coupled (4) unneeded elements including mechanics that don't relate to learning should be removed (5) game-specific jargon/fiction should be explained in narrative to avoid confusion, (6) pressure is important but the game play should not be too hard to achieve the learning goal, (7) be mindful of the engagement-to-complexity ratio, (8) failing to support novice players is a mistake. Even when higher level players are your

target, the first rounds should be made easier to facilitate success and an understanding of the overall system, (9) strive for consistency in theme, (10) strive for clean and consistent graphics that match the game style, do not worry about producing AAA entertainment title level quality, (11) pay attention to incorporation of appropriate sound, but personal taste in music is not important, (12) don't assume players will have instructors, peers, videos, or other external resources or that they will refer to them if they do; players expect everything in the game, (13) some players pay attention to external supporting materials and instructions, so include as much guidance as possible and (14) it is clear through countless hours of interacting with the evaluators that there is a need for clear instructions on how to use serious games, provide them. In terms of technology choices, ensure that appropriate use is made of the selected technology. Particularly for mobile games, the game should be designed with the mobile interface in mind and make use of its affordances including touch screen, accelerometer, etc. Finally, special hardware intended to enhance player experience should be used only where there is significant return on investment and used sparingly in controlled deployment environments with appropriate technical support where possible.

ACKNOWLEDGEMENTS

We would like to acknowledge NTSA for giving us the opportunity to put on the SGS&C as part of the STEM Pavilion for the last eight years; all of the many sponsors that have provided various levels of support be it financial or infrastructure based; the huge number of evaluators from across the industry who volunteer their time to support us; the other members of the planning committee and evaluators that have spent countless hours organizing and planning everything behind the scenes; and of course everyone who has entered a game for the chance to be recognized as part of the SGS&C.

REFERENCES

Chandler, T., Anthony, M. and Klinger, D. (2009). Applying High Cognitive vs. High Physical Fidelity within Serious Games. Proceedings of the Interservice/Industry Training, Simulation, and Education Conference. [CDROM]. Arlington, VA: National Training Systems Association.

Creighton, H. (2012). It's Not an Educational Game. Untold Entertainment Inc. Blog. Available online: http://www.untoldentertainment.com/blog/2012/01/26/its-not-an-educational-game/

Elliott, J., Adams, L., and Bruckman, A. (2002). No Magic Bullet: 3D Video Games in Education. Proceedings of ICLS 2002, International Conference of the Learning Sciences, Seattle, WA

ESA. (2013). Essential Facts About the Computer and Video Game Industry: 2013 Sales, Demographics, and Usage Data. ESA. Available Online: http://www.theesa.com/facts/pdfs/ESA_EF_2013.pdf

Gee, J. P. (2007). What video games have to teach us about learning and literacy: Palgrave Macmillan.

Ito, M. (2006). Engineering Play: Children. Discourse: Studies in the Cultural Politics of Education, 27(2), 22.

McNamara, J., Smith, P. A., Smith, B., Gritton, K. (2012). Serious Games Showcase & Challenge Distilled. Proceedings of the Interservice/Industry Training, Simulation, and Education Conference. [CDROM]. Arlington, VA: National Training Systems Association.

Quick, J., & Atkinson, R. (2011). A Data-Driven Taxonomy of Undergraduate Student Video Game Enjoyment.

Salen, K., & Zimmerman, E. (2004). Rules of play: Game design fundamentals: The MIT Press.

Smith, P. A. & Sanchez, A. (2009). Bite Sized Learning, Major Outcomes. Proceedings of the Interservice/Industry Training, Simulation, and Education Conference. [CDROM]. Arlington, VA: National Training Systems Association.