An Enjoyment Metric for the Evaluation of Alternate Reality Games

Andrew P. Macvean
School of Mathematical and Computer Sciences
Heriot-Watt University
apm8@hw.ac.uk

Mark O. Riedl
School of Interactive Computing
Georgia Institute of Technology
riedl@cc.gatech.edu

ABSTRACT

Alternate Reality Games layer a fictional world over the real world in order to provide players with a location-based interactive narrative experience. Building off previous work on game flow and enjoyment metrics in games, we present a metric based on the key elements that empirical studies suggest make for enjoyable ARG gameplay. We empirically validate our metric and call out key elements of ARGs that are most likely to have bearing on the success of a game.

Categories and Subject Descriptors

K.8.0 [**Computing Milieux**]: Personal Computing - Games; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Services - Artificial, augmented and virtual realities; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces - Evaluation/methodology, User-centred design

General Terms

Human Factors, Performance

Keywords

Alternate Reality Games, Pervasive Games, Enjoyment, Flow, Evaluation, Game Play

1. INTRODUCTION

Recently, we have seen the emergence of a new genre of gaming, known as *Pervasive Gaming*. According to Montola [5] "Pervasive gaming is a genre of gaming systematically blurring and breaking traditional boundaries of game". We are investigating one specific sub-genre of pervasive gaming known as *Alternate Reality Gaming (ARGs)*. Utilizing the principles of location-aware gaming, ARGs can be described as an interactive narrative experience played out in the real world of the player. ARGs create an illusion of a hidden truth, such as a conspiracy or mystery, by distributing narrative elements throughout the real world, thus creating the appearance that the game is unfolding around the player [8]. Although ARGs are a sub-genre of pervasive gaming, this additional focus on narrative brings with it unique and novel considerations for enjoyment.

How do we evaluate a particular instance of an ARG to determine whether it will be enjoyable or not? Leveraging Csikszentmihalyi's work on optimal experience [2], the theory of game flow [7] has been used to evaluate a broad range of games. Although pervasive games are a relatively new genre, Jegers' devised the Pervasive Game Flow (PGF) model [3] to extend game flow theory to specifically capture the essence of enjoyment within pervasive games. The unique design considerations required for ARGs leads us to question the suitability of preexisting enjoyment metrics for accurately capturing the key elements of an ARG experience. In this paper, we present our work on understanding enjoyment within ARGs. We describe a new enjoyment metric, PGF-ARG, which extends PGF in order to accurately assess the content of an ARG and allow conclusions to be drawn on how enjoyable the experience is likely to be.

2. ARG ENJOYMENT

ARGS are, by definition, predominantly story-based games and the primary enjoyment of ARGs arise from story based elements [6]. ARGs have a modular story structure, with the game split into small independent pieces, and subsequently scattered throughout the real world, requiring players to reassemble the pieces into a coherent whole [6]. In our earlier work [4] we highlighted 5 key principles we believe underpin the important content of an ARG story. In brief:

Strong Narrative Structure (I)

a) No dead ends, the final outcome should always be complete and meaningful.
b) The game should make use of traditional narrative elements, Propp and Aristotle.
c) The game should contain a number of problems which the player is required to solve.
d) The narrative should support the creation of both the overall goal of the game as well as a number of sub-goals, the resolution of which will lead to the fulfilment of the overall goal.
e) The narrative should support the player in their progress towards the goal, making it clear how they are progressing.

Modular Structure (II)

a) The game should be presented in small modular pieces.

b) It should be the players job to assemble these pieces in whatever manner they wish.

Meaningful Story Pieces (III)

a) Each piece should play a role in the overall outcome of the game.
b) The pieces should appear in a logical order.
c) Players should feel all elements of the game fit within the alternate reality created by the game.

Interactivity (IV)

a) The decisions and actions of the player should influence the overall outcome of the game.
b) The player should feel that the decisions had an effect on the game they experience.
c) While the player should have control, there should be constraints such that the player does not stray too greatly or feel lost in what they are trying to achieve.

Skill Level (V)

a) The game should be suitable for play by a general audience. Therefore, clues should not be personal to the developer or contain references, which make no sense to a wider audience.
b) The game should show progression in difficulty, allowing for a basic introduction and culminating in a climatic ending.

Based on our key requirements, we created a metric to assess player enjoyment within an ARG. In Table 1 we present the metric as a set of 11 key principles. The 5 original requirements were diluted into 11 principles in order to produce a more digestible and usable format. Along with each of the principles we attach a link back to the requirements presented above so that a link between the metric and the key requirements for enjoyment within ARGs can be established.

3. EVALUATION

With the metric devised, we ran two user studies in order to assess the success of the PGF-ARG enjoyment metric at encapsulating what makes for an enjoyable ARG experience. If our metric has value, this score should reflect how enjoyable a player would find the gameplay experience. In order to test our hypothesis, a two-stage evaluation process was followed. To evaluate the metric, we had to apply it to actual instances of ARGs with differing qualities of enjoyment. As it is hard to find two instances of ARGs that are similar enough in content to be comparable while having different enjoyment qualities, we were forced to create our own games. The games were created using the WeQuest authoring tool [1]. Thus, the first stage of our evaluation assessed whether instances of ARGs that we created had qualities we required to evaluate the metric. As games are a subjective topic, we wanted to ensure that the metric score we assigned to a game was not influenced by our role as both developers and evaluators.

The results of the first study can be found in Macvean and Riedl [4], confirming the games met our requirements as tools to evaluate the PGF-ARG metric. The second stage

Table 1: PGF-ARG Enjoyment Metric.

Criteria	Link
The game contained an appropriate introduction	**I** (b)(d)(e)
The game contained appropriate sub goals at each location	**I** (c)(d) **II** (a)(b) **III** (b)(c) **V** (a)
The game characters were interesting and relevant	**I** (b) **III** (a)(c) **V** (a)
The game locations were relevant	**I** (a) **III** (a)(b)(c)
The game contained a satisfying ending	**I** (a)(b)(d) **II** (b) **III** (a) **IV** (a)(b)(c) **V** (b)
The game contained no dead ends	**I** (a)(b) **III** (c) **IV** (b)
The content at each location was relevant to the overall narrative	**II** (a) **III** (a)(b)(c)
The game contained interesting and varied sub-goals	**I** (b)(c) **II** (b) **V** (a)
The player was afforded control over the game	**I** (c) **II** (b) **III** (a) **IV** (a)(b)(c)
The story was modular in nature	**IV** (a)(b) **V** (b)
The game was of appropriate difficulty	**I** (c)(e) **II** (a)(b) **III** (a) **IV** (c) **V** (a)

of our study, reported below, ascertains whether games that score well when marked using the metric will be judged to be more enjoyable than games that scored relatively poorly.

Seven participants were recruited to play each of the games, in a randomized order, using our web based game engine, Figure 1. A web based game engine was used due to the onerous nature of playing two ARGs in the real world. Each participant was asked to play each game twice, in order to explore various paths through the narrative structure. We asked our participants to rate each of the games they played for overall enjoyment using a 5-point scale. Participants then rated both games using a 5-point scale for each of the following criteria: clarity of goals, structure, outcome and non-player characters (NPCs). These criteria were selected based on the 7 significantly different metric criteria established in our earlier study [4]. By asking participants to score aspects of the game we know are significantly different we could then establish how certain criteria are evaluated when the participants are not explicitly told to look for them.

Results of our study found that the 'good' game was statistically more likely to score higher for enjoyment ($t = 5.46$, $p < 0.002$) than the 'bad' game. When asked to comment on
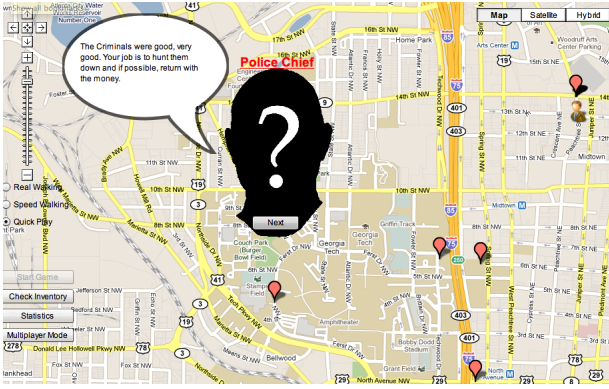
Figure 1: The game play interface showing a dialogue between the player and a NPC.

Table 2: Results.

Criteria(Link to PGF-ARG)	'Good' Avg	'Bad' Avg	T Test	Sig.
Clarity of Goals (1, 8)	4.6	1.6	9.72	$p < 0.001$
Structure (9, 11)	3.6	2.1	4.26	$p < 0.01$
Outcome (5, 6)	4.3	2.0	5.43	$p < 0.01$
NPCs (3)	3.9	2.6	3.06	$p < 0.05$

the extent to which they preferred one game over the other, using a separate 5-point scale, all seven participants marked that they preferred the 'good' game. On average quantifying this as a 1.87 difference (representing a slight preference).

Table 2 shows the ratings of the two games for the four key criteria. The participants were also given open-ended questions where they were asked to describe the reasons they preferred one game over the other. In total, nine distinct qualitative reasons were gathered, with a number of the reasons highlighted by multiple participants.

- The story was more interesting
- I had more control over the story / the story was less linear with more user decision points
- The game was longer
- My actions had consequences
- The sub-goals linking locations had more meaning
- The characters and events were more realistic
- Everything within the game played a role in creating the overall narrative / everything had meaning to the narrative
- There was more variety in actions
- My own role within the narrative was more interesting

In total, 6 of the 7 participants commented that they preferred the 'good' game as they felt they had more control over the story and the path through the narrative was less linear. Linked to this, 4 of the 7 participants felt that their actions had more consequences within the 'good' game and this contributed to their enjoyment. We can see from the qualitative feedback and the results in Table 2, a link between the structure and outcome of the game, and the way in which players reacted favorably to the various paths through the narrative and the way in which their decisions could influence the conclusion of the game. We see that the four key criteria were scored as significantly higher in the good game, correlating with the results from our first evaluation.

4. DISCUSSION AND CONCLUSION

This study affirmed our conclusion that the game we designed to be better was in fact found to be so. We note that there is only a slight preference for the 'good' game, which can be attributed to the fact that there were intentionally designed similarities between games. Players identified that control, outcome and a non-linear story structure were the most influential aspects on their enjoyment. Between the results of both our studies, we hypothesize that one can split the original PGF-ARG metric into two clusters; (a) the base elements which are required in order to create an ARG experience, and (b) the optional elements that have greater potential for direct influence on the overall enjoyment experienced by the players. The base elements are: modular structure, relevant locations, appropriate sub-goals, relevant location content, and difficulty. While these base elements ensure a playable ARG experience, the enjoyment elements are the aspects which guarantee for an enjoyable ARG experience. These elements, primarily linked to interactivity and narrative content are; appropriate introduction, interesting and relevant characters, satisfying ending, no dead ends, interesting sub-goals and overall player control.

Our results suggest that our metric is effective at predicting enjoyment within an ARG and that there is also potential for correlation between metric scores and the enjoyment experienced by the player, further evaluation is required to confirm this hypothesis. The limitation of our findings comes from the use of the web-based interface; real world evaluation of our games is required for a complete validation of the metric.

In this work we focus on investigating what makes for an enjoyable ARG experience. Building off prior work on game flow theory, we have evidence to suggest that the PGF-ARG can be a useful metric for evaluating ARG games. This work is one step toward providing empirically validated tools to help designers create better, more enjoyable experiences in pervasive, narrative based games.

5. REFERENCES

[1] C. Barve, S. Hajarnis, D. Karnik, and M. Riedl. A mobile alternate reality gaming platform and intelligent end-user authoring tool. In *Proceedings of Artificial Intelligence for Interactive Digital Entertainment*, 2010.

[2] M. Csikszentmihalyi. *The Psychology of Optimal Experience*. Harper Perennial, 1990.

[3] K. Jegers. Pervasive game flow: understanding player enjoyment in pervasive gaming. *Comput. Entertain.*, 5, January 2007.

[4] A. Macvean and M. Riedl. Evaluating enjoyment within alternate reality games. In *Proceedings of Computer Graphics and Interactive Techniques*, 2011.

[5] M. Montola. Exploring the edge of the magic circle. In *Proceedings of Digital Arts and Culture*, 2005.

[6] S. Stewart. *Alternate Reality Games*. http://www.seanstewart.org/interactive/args/.

[7] P. Sweetser and P. Wyeth. Gameflow: a model for evaluating player enjoyment in games. *Comput. Entertain.*, 3, July 2005.

[8] A. Waern, M. Montola, and J. Stenros. The three-sixty illusion: designing for immersion in pervasive games. In *Proceedings of the conference on Human Factors in Computing Systems*, 2009.