

A Computational Model of Perceived Agency in Video Games

David Thue¹ and Vadim Bulitko¹ and Marcia Spetch² and Trevon Romanuik¹

Departments of Computer Science¹ and Psychology²

University of Alberta

Edmonton, AB, T6G 2E8, CANADA

dthue | bulitko | mspetch | trevon @ ualberta.ca

Abstract

Agency, being one's ability to perform an action and have some influence over the world, is fundamental to interactive entertainment. Although much of the games industry is concerned with providing more agency to its players, what seems to matter more is how much agency each player will actually perceive. In this paper, we present a computational model of this phenomena, based on the notion that the amount of agency that one perceives depends on how much they desire the outcomes that result from their decisions. Using a structure for high-agency stories that we designed specifically for this intent, we present the results of a 141-participant user study that tests our model's ability to select subsequent events in an original interactive story. Using a newly validated survey instrument for measuring both agency and fun, we found with a high degree of confidence that event sequences selected by our model result in players perceiving more agency than players who experience event sequences that our model does not recommend.

Introduction

The ability to have an effect on the world (i.e., a person's agency) is a fundamental craving of human nature, having been linked to emotional well-being, improved performance, and good health (Bandura 2001; Thompson and Spacapan 1991). While progress in commercial video games has attempted to increase the amount of agency that is given to the player, what matters most is the amount of agency that each player *perceives*. Maximizing the latter is complicated in several ways.

First, in video games, many levels of interaction exist, ranging from players' physical interaction with the environment (e.g., knocking barrels in *Call of Duty: Black Ops*) to dialogue interaction with NPCs (e.g., the conversation wheel in *Mass Effect*) to interaction with the underlying plot (e.g., multiple endings in *Fallout: New Vegas*). The amount of agency *perceived* depends on the level of interaction. For example, if the player realizes that they can change the course of the game's story in a fundamental way early on (e.g., killing Agent Navarre in *Deus Ex*), they would arguably feel a great deal of agency. On the other extreme, deciding which side of the corridor to run along in a tunnel-shaped first-person-shooter affords very little perceived agency at all.

Copyright © 2011, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Second, given that many players play the game only once, they may not be able to assess the true amount of agency that the game actually provides (e.g., was there any way to stop Dr. Freeman in *Half Life: Opposing Force*, or was he scripted to jump into the teleport while staying just beyond the player's reach?). Many games exploit the fact that the player will see a particular sequence only once, and create an illusion of agency through carefully crafted exposition.

This paper studies perceived agency at the level of fundamentally changing the progression of a story-based game. We first present a novel, flexible structure for representing and organizing high-agency interactive stories, which we use to guarantee that every player has the same amount of agency while still experiencing any one of several possible stories. We then propose a model in which perceived agency is expressed in terms of the degree to which players desire the outcomes of their decisions. We implemented our model within an Artificial Intelligence system that dynamically modifies the story in an attempt to increase players' perceived agency, and evaluated it with a 141-person user study situated in an original interactive story developed with the Dragon Age: Origins Toolset (BioWare Corp. 2009). Players from both our control group (without our AI system) and our experimental group (with our AI system) were given the same amount of agency, yet the experimental group members perceived higher agency. This evidence lends support to our model of perceived agency.

Related Work

Recent research into the development of computational models of influence or agency has traditionally had the goal of finding some way to surreptitiously coerce the player of an interactive experience to go along with its designer's intent (Roberts, Narayanan, and Isbell 2009; Figueiredo and Paiva 2010); providing agency is wasteful if player desires can be controlled. While encouraging players to perceive their agency is markedly different from player coercion, we view the two efforts as being mutually beneficial. Weyhrauch (1997) discusses the notion of agency in terms of author-specified courses of action ("options"), but any perception of that agency by players is entirely assumed. Harrell and Zhu (2009) discuss agency from an aesthetic perspective, and our view that there are several levels of interaction and agency mirrors their notion of "agency scope".

Research in Social Psychology can provide important insight into the study of perceived agency. In particular, Thompson et al.’s “Control Heuristic” (1998) suggests that one’s perception of control varies with the degree to which they desire the outcomes that occur as a result of their decisions; good outcomes lead to strong feelings of control, while bad outcomes lead to weak feelings of control.

Proposed Approach

Toward increasing the likelihood of players perceiving their agency in an interactive story, we propose to partially select the content of that story in a way that leverages Thompson’s notion of desirable outcomes leading to increased feelings of control; this is similar in spirit to recent work by Wardrip-Fruin et al. (2009), as we ultimately evaluate outcomes with respect to the actions that they allow players to perform. With this goal in mind, we extended our prior work (Thue et al. 2010a), wherein a model of the player’s preferences for different styles of play is used to estimate the desirability of the events in a story. Following a brief overview of the player model, we present two novel contributions: a flexible structure for representing and organizing the events of a high-agency interactive story, and a model that represents perceived agency as the degree to which players desire the story events that occur as a result of their decisions.

Player Model Our previous interactive storytelling system, PaSSAGE, maintains its player model as a vector of five scalars, each representing the current player’s inclination toward a different style of play (e.g., <Fighter = 4, Method Actor = 8, Storyteller = 0, Tactician = 4, Power Gamer = 1>) (Thue et al. 2007). As a story progresses, PaSSAGE learns its model of the player automatically by observing her in-game actions and applying updates to the model based on author-supplied annotations (e.g., the action of attacking a friendly character could be annotated to increase the “Fighter” score in the model). To make use of PaSSAGE’s ability to estimate the desirability of events (which we review later on), our approach adopts its method of player modelling in its entirety.

A Flexible Story Structure

To fairly assess any attempt to increase players’ perceptions of agency, we must first guarantee that every player will be provided with the same amount of agency, regardless of which sequence of story content they experience. Toward satisfying this goal, we extend the *event* and *sub-event* sub-structures from our preliminary design (Thue et al. 2010a).

Sub-events Sub-events place the player in a scene at a particular point in the story, and a series of action unfolds that can involve player actions, non-player character (NPC) actions, or both. In our story structure, sub-events serve two primary purposes in addition to conveying the story: they offer players alternative courses of action so that the player model can be learned, and they present exactly one *key decision* that gives players agency over the story (Figure 1).

As an example of the type of content that a sub-event might contain, consider the following scenario: *The player returns from a journey and discovers that her home has been occupied by bandits. To resolve the situation, she can charge*

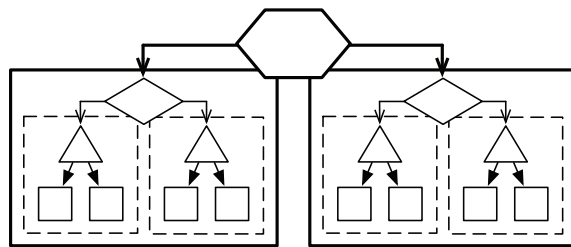


Figure 1: One phase of a story, with two events to select between adaptively. Within each sub-event (dashed boxes), outcomes (small squares) are determined by key player decisions (triangles). Within each event (solid boxes), sub-events are selected by testing the outcome of a prior key player decision (diamonds). An event is selected once the adaptive branching point (hexagon) is reached.

through the door in a frontal assault (Fighter style), or sneak in through the cellar and take their leader by surprise (Tactician style). Upon reaching the leader, she must make a decision: Should she make the leader pay, or let him go free?

Whenever the player makes a key decision, its outcome will have one of two effects. If the current sub-event is the last one in the story, then the outcome of the player’s decision will trigger a particular ending of the story. Otherwise, the outcome will be used to decide between two future sub-events that are considered as a matched pair. We refer to these matched pairs of sub-events as *events*.

Events Represented in Figure 1 by solid boxes around pairs of sub-events, each event provides a common situation in which one of the player’s prior key decisions can have a tangible effect. The tests of such decisions (diamonds in Figure 1) completely determine which of the event’s two sub-events will occur after a common introductory sequence is played. Expanding on our previous example, the player’s realization that bandits had occupied her home (a common introduction) could lead to completely different sub-events depending on whether or not she had managed to befriend the bandits previously. If the bandits were unfriendly, then the sub-event that we outlined above would occur; otherwise, the bandits might invite the player to help them defend against the law-makers of the region. Using previous player decisions to decide between pairs of sub-events is important to satisfying our goal of guaranteeing that every player will have the same degree of agency over the story.

Toward maximizing player agency, we propose that events should be selected adaptively, based on a model of how desirable that event will be for each particular current player. We refer to such opportunities for event selection as *adaptive branching points* (hexagon in Figure 1); they always occur at the start of new *phases* in the story.

Phases and Abstraction Figure 1 represents a possible configuration of a phase of the story: a set of one or more events¹, all of which are viable candidates for continuing the story, from which one must be selected to occur by an event selection procedure. We treat our story as a sequence of phases, each with a potentially different number of events.

¹Having only one event (with a single sub-event) in a phase is useful for learning an initial player model when the story first starts.

As the story progresses from one phase to the next, it can become difficult to guarantee that the outcomes of every possible key decision in one phase will be tested by some event in a phase later on. Without this guarantee, however, it would be impossible to ensure that every player would have the same degree of agency over the story. Furthermore, to ensure that *every* event in a given phase is a viable candidate for selection (which is desirable in terms of promoting varied content), we must guarantee that each event tests a prior player decision that she actually had the opportunity to make (this is a concern because different sub-events tend to offer different key decisions). We propose the use of *shared abstract decisions* and *shared abstract outcomes* as a tractable solution to this problem. By crafting all of the key decisions in a given phase to share a common, abstract theme (e.g., “Should I help or hinder the bandits?”), then the particular outcome of the decision (e.g., letting the bandit leader go free) can be recorded as a common, abstract outcome (e.g., the player helped the bandits). In a different sub-event within the same phase, this same abstract outcome might occur in a completely different way (e.g., if the player decides to help the bandits defend against the law-makers). If we convert all of the prior decision tests in each phase to consider the shared, abstract outcome of a particular prior phase (e.g., having both of our example sub-events test whether the player had befriended the bandits in some way), then every event will be guaranteed to test the outcome of a key decision that has actually been made. By requiring every event in each phase to be viable for selection regardless of which events came previously, we gain the tremendous flexibility of having every possible combination of sub-events (choosing one from each phase) be a complete and viable story.

The overall effect of using shared abstract decisions and outcomes is that every key player decision will eliminate half of the sub-events in the phase whose events test its abstract outcome. This effect guarantees that *as long as every phase’s abstract outcome is either tested in some later phase or triggers an ending of the story, every player will have the same amount of agency*, regardless of which events are selected by the event selection process.

A Computational Model of Perceived Agency

Whenever the adaptive branching point of a given phase is reached, an event selection process is activated to determine which event will happen next. Toward maximizing players’ perceptions of agency while they play, we propose a computational model of perceived agency that can be used to rank all of the events in a phase with respect to estimates of their desirability from the player’s point of view. As suggested by Thompson et al.’s Control Heuristic (1998), our model is based on the notion that players’ perceived agency can be increased by selecting events in a way that causes their decisions to lead to the most desirable outcome available². Referring back to our story structure, given that the outcomes of key player decisions have no immediate effect on the types of actions that players can perform, PaSSAGE’s

²A more thorough treatment of applying the Control Heuristic in story-based games can be found in our previous work (2010a).

player model is unable to estimate the desirability of key decision outcomes directly. However, the outcomes of key decisions do have a *delayed* effect on the types of actions that players can perform; they control which sub-events will be available in any phase that tests their abstract counterparts (our structure guarantees that there will be at least one such phase), and each sub-event offers courses of action that are suited to different styles of play. By treating the available sub-events in a phase as the potential outcomes of a key player decision, we can use PaSSAGE’s player model to estimate the desirability of those outcomes.

Sub-event Desirability Drawing again from our preliminary design (Thue et al. 2010a), we propose that the desirability of a sub-event can be estimated as the maximum of the desirabilities of each of its available courses of action, which are annotated by their authors to describe how well players of each given play style are expected to enjoy playing through them (e.g., a frontal assault on the bandits might have a high annotation for the Fighter style of play). The desirability of each course of action is then estimated as the inner product of two vectors: PaSSAGE’s learned player model, and the annotation vector that was provided for that course of action. Equation 1 gives the calculation for estimating the desirability of a general sub-event, e .

$$desirability(e) = \max_{a \in Actions(e)} [PlayerModel \cdot Annotations(a)] \quad (1)$$

Given the desirability of each sub-event in the current phase, we propose that the final ranking of the events in that phase should be determined by a metric called *relevance*.

Decision Relevance Calculated for a given key decision (d) with respect to each event (E) that tests it, relevance serves as a measure of the player’s missed or avoided opportunities; it is based on the relative desirability between the sub-event that is currently available ($e^+ \in E$), and the N sub-events that are not available ($e_i^- \in E$) (Equation 2)³.

$$relevance(d|E) = desirability(e^+) - \frac{1}{N} \sum_{i=1}^N desirability(e_i^-) \quad (2)$$

Continuing with our example, if a player who detests fighting or sneaking had managed to befriend the bandits earlier on, then having the option to work with them cooperatively instead of attacking might well be a welcome relief. Assuming that the player model had learned about this player correctly, then the relevance for that pair of sub-events ($[e^+ = \text{Cooperate-or-Defend}] - [e_1^- = \text{Attack-or-Sneak}]$) would be a high positive value.

Event Desirability Although we initially recommended Equation 2 with an absolute value operation around its right-hand side, we noted that without special, individual consideration of every ranked event, arbitrarily *undesirable* sub-events could be selected to occur, even when more desirable alternatives were available (e.g., (e^+, e_1^-) values of $(-20, 10)$ and $(20, 5)$ yield absolute relevance values of 30 and 15, but

³This is a generalization of our earlier equation (2010a) to cases where more than one unavailable sub-event exists.

the first event is much less desirable than the second). To simplify the event selection procedure and have it rely less on special cases, we propose that the desirability of an event E (given a prior decision d) should be computed as a linear combination of d 's relevance given E and the desirability of its active sub-event, treating the former as a weighted bonus or penalty to the latter (Equation 3).

$$\text{desirability}(E|d) = \text{desirability}(e^+) + \alpha \times \text{relevance}(d|E) \quad (3)$$

How should the weight α be set? Put another way, how likely is the player to foresee the desirabilities of unavailable sub-events? Given that the answer to this question depends highly on how each particular event and sub-event are authored, we propose that having an additional set of author annotations on each sub-event can help decide when relevance is more or less important to consider. For each sub-event within an event, authors provide values in $[0, 1]$ which estimate how foreseeable each of the other sub-events would be, if the given sub-event occurred. For example, if the player's sneaky journey to surprise the bandit leader revealed that local law-makers were planning to attack, then she may consider that it might have been possible to help the bandits defend against the law-makers, if she had managed to befriend them earlier. As a result of these annotations, each sub-event obtains a vector of N foreseeability values, each referring to one of the other sub-events in the event. After expanding the relevance term, replacing α with $\alpha_i(e^+)$, and simplifying, Equation 3 becomes Equation 4.

$$\text{desirability}(E|d) = \left(\text{desirability}(e^+) \times \left(1 + \sum_{i=1}^N \frac{\alpha_i(e^+)}{N} \right) - \text{desirability}(e_i^-) \cdot \frac{\alpha_i(e^+)}{N} \right) \quad (4)$$

We propose that this estimate of event desirability can serve as a model of the agency that players perceive over the story. When used to rank the events in each story phase, the top-ranked events will represent the most desirable available outcomes of the player's prior key decisions. By dynamically selecting these events as the story unfolds, we aim to increase each player's perception of agency over the story.

Implementation

Toward demonstrating and testing our approach, we used our proposed story structure to design and implement *Lord of the Borderlands*, an original interactive story crafted in a commercially available version of the Dragon Age: Origins Toolset (BioWare Corp. 2009). The story takes roughly forty-five minutes to play, featuring detailed 3D environments, carefully crafted dialogue, and numerous animated cutscenes with music and sound effects. The story is split into four phases, the first two of which each only contain a single sub-event inside a single event, serving as an opportunity to learn the player's model before the first event selection is made. The structure of phase 3 is shown in Figure 1. Along with our implementation of the story, we developed a new event selection algorithm, PaSSAGE-2, which chooses events based on our model of perceived player agency.

To highlight the flexibility of our story structure, we decided to connect the abstract decision outcomes from earlier phases to the abstract decision tests in later phases in a "leap frog" fashion, connecting phase 1's outcomes to phase 3's tests, and phase 2's outcomes to phase 4's tests. To ensure that players' decisions in phase 3 would have an effect on the story, we doubled the number of sub-events in phase 4's events (to 8 in total) and added a second layer of decision tests that target phase 3's decisions; as a result, sub-events become available in phase 4 based on two prior player decisions. At the level of individual courses of action (e.g., conducting a frontal assault on the bandits versus sneaking in through the cellar), there are roughly 1600 unique ways to play *Lord of the Borderlands*, though all of the content at this level could be experienced with sixteen diverse plays.

Empirical Evaluation

We designed an evaluation to test PaSSAGE-2 in two ways. First, we compared PaSSAGE-2's ability to increase player agency versus the traditional approach of delivering interactive stories in a non-adaptive way. In doing so we compared two experimental groups: 'Adaptive' (A) versus 'Fixed' (F), which we describe below. Second, we tested PaSSAGE-2's ability to distinguish between candidate interactive stories with respect to the amount of agency that its players will perceive; this test also required two groups: 'Recommended' (R) versus 'Not Recommended' (NR).

Adaptive vs. Fixed Players in group A experienced stories whose adaptive branching points were controlled by PaSSAGE-2, while players in group F experienced stories that were forced to match a previous A player's experience. Building group F in this way ensured that the event-level content in the story would be experienced with roughly the same frequency in both A and F . Given that the sub-events that occur within each event are determined entirely by the player's prior choices, it was impossible to balance content at the sub-event level while the study was being run.

Recommended vs. Not Recommended Players in group R experienced stories that PaSSAGE-2 *recommended*: stories which its model indicated would result in maximizing the player's perceived agency. R differs from group A in that R additionally includes players from the fixed group (F) who, by random chance, happened to experience the story that PaSSAGE-2 *would* have chosen for them, if it had been able to do so. We refer to such players as *lucky cases*. Players in group NR experienced stories that were different from what PaSSAGE-2 recommended.

Hypotheses We designed a user study to test two hypotheses for each pair of player groups: that players in the first group of each pair (Adaptive or Recommended) would report that their story experience made them feel more agency than players in the second group (Fixed or Not Recommended) reported, and that players in the first group would report that their story experience was more fun than players in the second group reported. We refer to our hypotheses for 'Adaptive vs. Fixed' as **H1:Agency** and **H1:Fun**, and our hypotheses for 'Recommended vs. Not Recommended' as **H2:Agency** and **H2:Fun**.

Survey Instrument Validation In an improvement over previous, similar studies where the psychometric validity of the survey instruments was unknown (e.g., (Thue et al. 2007; 2010b)), we have adopted and validated two recently developed scales for measuring player fun and agency (Vermeulen et al. 2010). In its initial form, the scale for measuring fun included thirteen Likert-inspired items (e.g., “This experience was gratifying” and “This experience was amusing”) to which players respond on a five point scale ranging from “Strongly Disagree” = 1 to “Strongly Agree” = 5 (Likert 1932). To avoid any participants being confused about which experience they were evaluating, we changed the text of each item to begin with “My story experience” instead of “This experience”. The scale for measuring agency included six of the same type of items (e.g., “My inputs had considerable impact on the events in the story”).

To validate these instruments, we used them to measure the (self-reported) fun and agency of 45 players playing a PaSSAGE-2-selected story. By computing the average scores of each item and examining the correlations between them, we determined that one item from the fun scale (“My story experience was melancholy”) was poorly correlated with the rest of the items on the scale. Removing this item resulted in an internal consistency (Cronbach’s alpha) score for the fun instrument of 0.945, giving strong evidence for the validity of the modified scale⁴. All items on the scale for agency were well-correlated with one another, and we adopted them without modification (Cronbach’s alpha = 0.941). A factor analysis of each scale confirmed that their items can be well thought of as elements of a single factor for each, and we suspect that these factors represent common notions of “fun” and “agency”.

Experimental Procedure

We conducted our study with 141 participants (mean age: 19.4, 87 female), gathered from the student body of a university undergraduate Psychology course. After a short briefing, participants were invited to play through a small, in-game tutorial, toward familiarizing themselves with the controls of the Dragon Age: Origins engine. To avoid biasing our participants’ feedback, no mention of PaSSAGE-2’s adaptive technology was made at any point during the study session; participants were instead told that researchers were attempting to discover which aspects of interactive stories can affect how much fun their players have, or how much influence they feel. After the tutorial, each participant was randomly assigned to one of our two experimental groups. For players in the Adaptive group, PaSSAGE-2 learned their preferences for play style while they played, and used our story’s adaptive branching points to select events that might increase each player’s feelings of agency. For players in the Fixed group, the sequence of events that they experienced was automatically selected to match the sequence that the most recent Adaptive group player had experienced while seated at the same workstation in our lab. Following their

⁴One item from the original scale (“This experience was moving”) was accidentally omitted from this validation process. However, the high internal consistency of our final, eleven-item scale suggests that this item may not be needed to reliably measure fun.

experience, players were asked to complete an online questionnaire based on the instruments for fun and agency described above; we additionally recorded their age, gender, and prior experience playing video games.

Data Analysis

We began our analysis by searching for outliers in both of our metrics (Fun and Agency) across both of our targeted groupings (Adaptive vs. Fixed, and Recommended vs. Not Recommended). As the same five outliers were found to be far away from their means in both targeted groupings (tested via Interquartile Ranges), we eliminated them from our data set altogether. We then identified three properties of our players which, if significantly imbalanced across our experimental and control groups, could potentially bias our results: their gender, the particular event sequence that they experienced⁵, and their prior experience playing video games. Using Chi-Square tests on each factor to detect any significant imbalances between our groups, we found that prior player experience was the only factor of the three that warranted concern with respect to bias; this result held across both of our metrics and both of our targeted groupings. To control for prior player experience as a potential source of bias, we tested each of our hypotheses with a two-way Analysis of Variance (ANOVA) procedure using Type III Sums of Squares. In keeping with the design of our prior studies (2010b), we split players’ prior experience ratings into two levels: ‘low’, being players who play video games for at most one hour in an average week, and ‘high’, being players who play video games for an hour or more each week. Before running each ANOVA, we confirmed its assumptions of Normal distributions using z-score analysis of skewness and kurtosis, and homogenous variances using Levene’s Test.

Study Results

Table 1 shows the primary results of our ANOVA tests. Column 1 indicates which hypothesis is being tested, columns 2 and 3 provide the number of participants, mean score, and standard error for our experimental and control groups, respectively, and column 4 provides significance results as confidence in the given hypothesis. The given confidence scores are equal to one minus the p_{value} reported for the main effect of our experimental/control factor in the ANOVA. There were fifteen lucky cases.

H1	Adaptive (70)	Fixed (66)	$A > F$
Agency	4.05 ± 0.09	3.89 ± 0.11	74.9%
Fun	3.17 ± 0.11	3.17 ± 0.12	<i>low conf.</i>
H2	Recom. (85)	Not Rec. (51)	$R > NR$
Agency	4.07 ± 0.08	3.78 ± 0.12	96.1%
Fun	3.16 ± 0.10	3.21 ± 0.14	<i>low conf.</i>

Table 1: Number of players, mean values with standard errors and confidence scores ($1 - p_{\text{value}}$) of ANOVA tests of 4 hypotheses: that Adaptive > Fixed and that Recommended > Not Recommended for ratings of both Agency and Fun. “*low conf.*” replaces scores less than 50%, as little can be said about the given hypothesis.

⁵Although our study design aimed to control for particular event sequences, random technical errors could still cause imbalance.

Prior gameplay experience affected both agency and fun, yielding main effects for this factor in three of the four tests (*A* vs. *F*, Agency: 92.1%, *A* vs *F*, Fun: 94.6%, and *R* vs. *NR*, Agency: 90.7%). This highlights the importance of having statistically controlled for imbalances in prior gameplay experience. No interaction effects were significant.

Discussion & Future Work

As shown in Table 1, the strongest result of this study is that players whose *Lord of the Borderlands* stories are directed by PaSSAGE-2 feel more agency (by $\approx 10\%$) than players whose stories deviated from what PaSSAGE-2 would have arranged; this result supports our hypothesis H2:Agency with 96.1% confidence. Given that all players had an equal amount of agency over their story, this result also confirms our hypothesis that providing agency alone isn't sufficient to maximize its perception by players. Furthermore, it shows that for any particular player, PaSSAGE-2 can distinguish between subtrees of *Lord of the Borderlands* stories with respect to how much agency the current player would be likely to perceive if they were to experience them directly.

The practical implications of our result for H2:Agency are at least twofold. For story designers interested in on-line learning and adaptive gameplay, PaSSAGE-2 could be deployed at runtime to recommend which events would best help players perceive their agency. For story designers who can gather play-style data about their target audience in advance, PaSSAGE-2 could be deployed offline on even an early skeleton version of a story, so long as reliable annotation estimates could be obtained. Given such estimates, along with a representative sample of player models from the target audience, PaSSAGE-2 could rate each of several proposed events with respect to how well they might evoke feelings of agency in that audience, and provide insight into where resources for development might best be invested.

Although our result for H1:Agency (74.9% confidence) is not statistically significant, we believe that this test was confounded by the presence of fifteen lucky cases in the control group for the test (Fixed), which push the means together and make any significant difference more difficult to detect. While PaSSAGE-2 does not attempt to maximize players' fun ratings directly, our results for H1:Fun and H2:Fun (low confidence) are somewhat curious, as both pairs of means are equal within standard error; we can neither confirm nor disconfirm our hypotheses for Fun.

In terms of future work, several questions remain. Why did our method for selecting desirable story events have no discernible effect on players' ratings of fun? Would a longer story with more adaptive branching points provide PaSSAGE-2 with a better chance of influencing those ratings? Could this work be combined with our previous version of PaSSAGE (Thue et al. 2010b), which obtained a fairly high confidence score for improving player fun?

Conclusion

In this paper, we have explored the concept of perceived agency in computer video games. In doing so, we have made several novel contributions: (i) a flexible, practical structure

for representing and organizing the events of high-agency interactive stories, (ii) a computational model of perceived agency based on how desirable players find the events that follow from their decisions to be, (iii) an original, high-agency interactive story that demonstrates our story structure and implements our model, (iv) a validation study of a pre-existing instrument for measuring fun and agency in games, and (v) a 141-participant user study which showed that with 96% confidence, our computational model of perceived agency can, on average, predict which subtrees of a story will afford higher perceived agency to its player.

Acknowledgements

This work was supported by NSERC and iCORE. We are also grateful for Jim Muirhead's help with our analysis and Charles Crittenden's authoring of *Lord of the Borderlands*.

References

- Bandura, A. 2001. Social cognitive theory: An agentic perspective. *Annual Review of Psychology* 52:1–26.
- BioWare Corp. 2009. The Dragon Age: Origins Toolset. <http://dragonage.bioware.com/toolset/>.
- Figueiredo, R., and Paiva, A. 2010. "I want to slay that dragon!": Influencing choice in interactive storytelling. In *ICIDS 2010*, 26–37. Edinburgh, Scotland: Springer-Verlag.
- Harrell, D. F., and Zhu, J. 2009. Agency play: Dimensions of agency for interactive narrative design. In *AAAI Symp. on Intelligent Narrative Tech. II*, 44–52.
- Likert, R. 1932. A technique for the measurement of attitudes. *Archives of Psychology* 22(140):55.
- Roberts, D. L.; Narayanan, H.; and Isbell, C. 2009. Learning to influence emotional responses for interactive storytelling. In *AAAI Symp. on Intelligent Narrative Tech. II*. 95–102.
- Thompson, S. C., and Spacapan, S. 1991. Perceptions of control in vulnerable populations. *Social Issues* 47(4):1–21.
- Thompson, S. C.; Thomas, C.; and Armstrong, W. 1998. Illusions of control, underestimations, and accuracy: A control heuristic explanation. *Psychological Bulletin* 123(2):143–161.
- Thue, D.; Bulitko, V.; Spetch, M.; and Wasylishen, E. 2007. Interactive storytelling: A player modelling approach. In *AIIDE 2007*, 43–48. Palo Alto, California: AAAI Press.
- Thue, D.; Bulitko, V.; Spetch, M.; and Romanuik, T. 2010a. Player agency and the relevance of decisions. In *ICIDS 2010*, 210–215. Edinburgh, Scotland: Springer Verlag.
- Thue, D.; Bulitko, V.; Spetch, M.; and Webb, M. 2010b. Socially consistent characters in player-specific stories. In *AIIDE 2010*, 198–203. Palo Alto, California: AAAI Press.
- Vermeulen, I. E.; Roth, C.; Vorderer, P.; and Klimmt, C. 2010. Measuring user responses to interactive stories: towards a standardized assessment tool. In *ICIDS 2010*, 38–43. Edinburgh, Scotland: Springer-Verlag.
- Wardrip-Fruin, N.; Mateas, M.; Dow, S.; and Sali, S. 2009. Agency reconsidered. In *DIGRA 09*.
- Weyhrauch, P. 1997. *Guiding Interactive Drama*. Ph.D. Dissertation, CMU.