Characterizing Gameplay in a Player Model of Game Story Comprehension

Rogelio E. Cardona-Rivera
recardon@ncsu.edu

R. Michael Young
young@csc.ncsu.edu

Digital Games Research Center
North Carolina State University
Campus Box 8206
Raleigh, NC 27695

ABSTRACT

This work seeks to validate the use of a situation model (a model for the comprehension processes in non-interactive narrative) for use modeling a player's comprehension process within interactive narratives and games. Unlike conventional narratives, games contain actions and events that might not carry any narrative significance; consequently, a cognitive model of player understanding may require a new distinction between "narratively important" and "narratively unimportant" events. This hypothesis was tested in an experiment using Event Segmentation Theory. Our results failed to reject the null hypothesis (Fisher's exact test, $p = 1$) but were insightful. Narrative importance seems to be perceptually oriented, and dependent on the outcomes of the event in question. Although our data failed to reject the null hypothesis, more work must be done before rejecting the necessity of the distinction between narratively important and unimportant events when modeling a player's comprehension process of an unfolding story-focused game.

Categories and Subject Descriptors

I.2.0 [**Artificial Intelligence**]: General—*Cognitive simulation*; I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems—*Games*

General Terms

Experimentation, Human Factors, Measurement

Keywords

Interactive Narrative, Player Model, Event Segmentation Theory

1. INTRODUCTION

Previous work has described an initial model of the cognitive processes involved in a player's understanding of an unfolding story as the player plays in an interactive narrative virtual environment (i.e. in a video game) [9]. This previous work posited that *situation models* [16], which have primarily been used to model the cognitive processes involved in story comprehension for non-interactive narrative, are a reasonable proxy for modeling story comprehension in an interactive narrative setting. As defined by psychologists, situation models index every perceived event (defined by verbs or actions) along 5 situational dimensions: Time, Space, Intention, Causality, and Character/Protagonist [16]. The model can be used to predict the strength of recall of events; events that share the same index (e.g. events that take place in the same location will share a Space index) are more likely to be recalled due to their tight integration than events with different indices.

From a narratological perspective, it makes sense to apply situation models for modeling the cognitive processes that are active when experiencing a narrative. Conventional narratives can be viewed as collections of *communicative acts* that observe Grice's Cooperativity Principle [4]; the audience expects that every story element inserted by an author of a narrative will ultimately demonstrate its relevance. In short, everything that is present in a narrative is understood as serving some authorial purpose. Thus, from a reader's perspective, it is not unreasonable to index every event in a narrative as part of his or her cognitive model of the story.

When considering an interactive narrative, some of the same assumptions hold. Young [8] posits that Grice's Cooperativity Principle is still applicable in this situation; the audience (i.e. the video game player) relies on the game to communicate clues regarding how to act in order to further the story. The game figuratively acts on behalf of the game designer, whom is the parallel figure to the author in non-interactive narratives. The player, in turn, is expected to use her knowledge of the environment and her perception of other contextual elements to act in a way that doesn't break the story's development.

It is not clear, however, that assumptions of relevance and cooperativity are valid for every element of a story that is experienced by a player within interactive narratives like games. We feel that it is possible for certain events to carry no narrative significance in the development of the story. A player's ludic behavior (e.g. exploration), which is often undirected, may or may not bear relevance to the interactive narrative's actual story. Further, gameplay in which a player makes an error in play (e.g. when a player unintentionally falls off of a bridge that she needs to cross to finish a level)

FDG'12 May 29-June 1, 2012 Raleigh, NC, USA

clearly doesn't advance the story plot but is still an aspect of a player's interaction with the environment.

In order to use situation models to characterize a player's comprehension of story in video games, it would seem that we must distinguish between two classes of events: one class of events that are "narratively important" and another class of events that are "narratively unimportant", that is, that bear no relevance to the development of the story. This is the issue we explore in this paper: is it necessary to distinguish between two classes of narrative events (important vs. unimportant) in a model of player story comprehension?

If this distinction is necessary, it implies that in order to use situation models for modeling story comprehension in interactive virtual environments, we must be able to detect narratively unimportant events as the player plays the game. These events would then be filtered from the player model, as they wouldn't inform the model. Without this characterization, we could potentially introduce unnecessary noise in the model that would lead to misrepresenting the player's mental state regarding the progression of the game's story.

To address this question, we looked at how people perceive events in interactive narrative virtual environments (i.e. games). If people draw a distinction between narratively important and unimportant events when understanding gameplay, we would expect people to be able to individually distinguish these types of events within a game. To test whether or not this is the case, we designed an experiment that relied on event segmentation theory [12] to gauge the ways that people categorize events in gameplay.

2. EVENT PERCEPTION IN INTERACTIVE NARRATIVE VIRTUAL ENVIRONMENTS

Event segmentation theory (EST) covers a wide collection of research on the human perception of ongoing action. Zacks, Speer and Reynolds [12] characterize EST as: "a computationally and neurophysiologically explicit account of event structure perception." The EST framework posits that continuous streams of incoming information are segmented into a series of discrete units [6, 12]. Each unit, or *event model* is made up of information about the current state of the world, as well as information from other similar event models. Zacks and Swallow [14] note: "segmenting ongoing activity into meaningful events is a core component of perception and this has consequences for memory and learning".

Most experimental research concerning EST has used a variant of the experimental design developed by Newtson [6], in which participants were tasked with viewing a film and as they viewed, they were instructed to press a button when, in their opinion, one meaningful event ended and another event began. This task produced consistent event segments for all participants across time [6, 12] for two types of events:

1. Coarse-grained events - large meaningful segments of activity
2. Fine-grained events - small meaningful segments of activity, which are contained hierarchically within coarse-grained events

Fine-grained events experimentally correlate with changes in space (i.e. movement), whereas coarse-grained events experimentally correlate with conceptual changes such as changes in intentional or goal-directed activity [10].

A concern which besets the validity of EST as an actual model of the cognitive processes involved in event perception is that most experiments explicitly instruct participants to segment the events. It is possible that event segmentation suffers from a priming effect, that is, participants don't naturally segment events into meaningful units. However, Zacks, Braver et al. [11] studied changes in brain activity using functional magnetic resonance imaging (fMRI) while participants viewed movies without explicit segmentation instructions. When asked to view the movie again with explicit segmentation instructions, the researchers found significant activation overlap between the uninstructed and instructed conditions' brain patterns. As Zacks and Swallow note,

> Because the critical brain data were acquired before participants learned of the segmentation procedure, these changes cannot be attributed to overt or covert performance of a laboratory-specific task. These results strongly imply that brain processes correlated with event segmentation are a normal part of ongoing perception. [10].

The fMRI study presents a convincing argument that event segmentation is an automatic activity, but it doesn't address the criteria that humans use to segment events. Segmentation may be guided by information-seeking goals, and there is evidence that observers can adjust the grain at which they segment events [15].

2.1 EST and Situation Models

The primitive units of narrative used by situation models are events, which are defined by verb phrases in text or actions in film. Zwaan and Radvansky [16] distinguish between three types of situation models that work at different stages of the narrative comprehension process:

1. The current situation model - refers to the model of the event that is currently being perceived. This is the model at time t_n, for a given event e_n.
2. The integrated situation model - refers to the model of the events that have been perceived up until right before the event currently being perceived. This is the model for times t_1 through t_{n-1}, for events e_1 through e_{n-1}.
3. The complete situation model - refers to the model that is stored in long-term memory after all events have been processed. This is the model for times t_1 through t_x, for events e_1 through e_x, where the index x is arbitrary and indicates the reader's or viewer's perceived end of the story. The complete situation model combines the current and the integrated situation models.

We adopt the view espoused by Zacks, Speer and Reynolds that the terms *current situation model* from situation model literature and *event model* from EST literature describe functionally equivalent mental representations [12]. Practically, this characterization licenses us to evaluate situation model phenomena via EST. Our experiment operationalization does precisely that.

3. OVERVIEW OF EXPERIMENT

In this experiment, we sought to determine whether or not it is appropriate to use situation models – previously used for modeling comprehension processes in non-interactive narratives – for modeling the story comprehension processes within games. Because there could be events in a game that don't play any role in the player's comprehension of the game's narrative, we expected that some events might not be included in the construction of game story situation models during gameplay.

We hypothesized that a distinction between narratively important and unimportant events should be recognizable by humans. Our operationalization of the hypothesis was based on EST: if tasked with segmenting the action in a video game, humans should segment actions taking narrative importance in to account. We expected that narratively important events within games would be segmented, and narratively unimportant events would be unsegmented. Instead, those narratively unimportant events would be aggregated with the narratively important ones. The segmentation of events constitutes a player's online comprehension process, that is, the comprehension of the game's story as the player plays the game.

We designed and executed an experiment to test the validity of the idea that some elements of a player's experience during game play are not incorporated into a situation model. In this experiment, we required participants to watch a video and segment the action they saw into coarse-grained units, which correlate to conceptual changes such as changes in intentional or goal-directed activity. [1] Subjects were divided into two groups and each group saw a different version of a video showing near-identical gameplay from the video game *Halo: Reach* [2]. Both videos record a third-person view of a player-controlled avatar attempting to complete a level by first crossing a bridge to a landing platform which contains a helicopter and then boarding the helicopter to escape the levelÕs NPC opponents. In the videos, the camera follows the player using one continuous shot without edits. The shot tracks a view directly behind the player's forward-looking vector, identical to typical third-person follow-cam perspectives.

In both videos, the player initially comes under rocket fire from an NPC. This NPC is positioned inside a fort near the landing platform, located across a chasm. The player begins to move across the only bridge that spans the chasm, but falls off the bridge at the near edge. The player lands on a ledge which is right below the start of the bridge. The ledge is narrow and contains a route back up to the start of the bridge. After landing on the ledge, the player looks down, re-orients her view, and climbs back to the start of the bridge. At that point, the player crosses the bridge past enemy fire and gets inside the escape helicopter. The helicopter takes off and is shot several times by the enemy on the landing platform. The gameplay concludes with the helicopter exploding when it is struck by a rocket fired by the enemy NPC.

The two videos were identical except that one version of the gameplay (version NR) contained an event that was narratively unimportant while in the other version (version R), the same action was promoted to being narratively important. For the purposes of this experiment, we considered an event to be narratively important if it was relevant to the development of the story, that is, if it was either initiated/caused by the game itself or if the event furthered the player's goals. [2]

For the experiment, we manipulated the narrative importance of the event of the player falling off the bridge onto the ledge. In version NR of the video, the player falls off the bridge as described above. In version R of the video, the same falling sequence from the NR version is immediately preceded by a rocket which is fired at the player by an enemy avatar near the landing platform. The rocket misses its target and hits the ground, exploding behind the player. The player falls off the edge of the bridge at that point in a manner identical to the fall in the NR version.

The experiment was designed around the following hypothesis:

> $H1$: Participants viewing the NR version of gameplay will *not* segment the event of the game avatar falling off, whereas the same fall will be segmented in the R version of gameplay.
>
> $H0$: There will be no significant difference in the segmentation of the event of the game avatar falling off across both versions of the gameplay.

The gameplay shown in the NR version of the video was designed to appear as if the player made a navigation error while in crossing the bridge, moving off the side of the walkway and falling to the ledge. Since there is no apparent reason for why the player falls, the fall is narratively unimportant. The fall does *not* correspond to a change in intentional or goal-directed activity, neither from the player's point of view, nor from the game designer's point of view (because the fall includes no events that are initiated by the game).

In contrast, the gameplay shown in the R version of the video introduces a potential system-initiated cause for the player's fall, which promotes the event to a narratively important one (since it is now game initiated). While it does not correspond to a change in the player's intentional or goal-directed activity, we feel it indicates a change in the intentional activity of the game designer, because the game initiates an action which appears to have obstructed the player's progress. Thus, our prediction is that, in the R version of gameplay, a participant's segmentation of coarse-grained units will reflect the fact that there *is* a change in the author's intentional/goal-directed activity. In summary, we expect participants in the NR condition will not segment the player falling off because it is not narratively important, whereas participants in the R condition will segment the player falling off because it is narratively important.

3.1 Method

3.1.1 Participants

Sixteen students from North Carolina State University (ages 18-28 years, 12 men, 4 women) participated in this

[1]While we may not be able to decidedly affirm that an external viewer is capable of exactly reconstructing the event/situation model of the player who actually played the game, research in experimental psychology suggests that we may be justified in doing so [7].

[2]The definition was contrived to exclude a player's error in gameplay; we make no commitment to exploratory behavior or variants thereof.

experiment; 14 participants received course credit and 2 participants volunteered without compensation. Seventeen additional participants were excluded due to:

- familiarity with the experimental materials ($n = 12$). These participants detected the staged nature of the experimental materials and noted that the gameplay they segmented didn't seem "natural".
- failure to follow the task instructions ($n = 5$). These participants segmented events that were too narrow, contrary to experimental instructions. The shaping procedure we used to determine this qualification is discussed in Section 3.1.3.

3.1.2 Materials

Three gameplay traces from the video game Halo: Reach [2] were created and used as stimuli for participants. All videos last 62 seconds. All game levels were created by humans in the game's level editor: The Forge. The gameplay traces were created in the game's Custom Game (multiplayer) mode. The first trace was used to practice the experimental task. Participants then watched one of the other two videos; each of those two videos represent a distinct experimental condition. In the experimental condition gameplay traces, the player is trying to beat the level by escaping in a helicopter. Both experiment traces contain very similar action sequences and were labeled NR and R, respectively. The perspective shown in the video is a third-person over-the-shoulder view, which follows the player.

In the NR (short for "no-rocket") version of gameplay, the player navigates to the top of a hill where she sees a salvo from a laser cannon impact against a rock formation near her. The player navigates towards a narrow bridge, which connects the top of the hill to a landing platform. The landing platform contains a helicopter and a watch tower with an enemy at the top. As the player begins crossing the bridge, she moves off the edge of the walkway without looking and lands on the ledge below. The player looks down for orientation, looks back up, and proceeds to climb an auxiliary path which leads back to the start of the bridge.

When the player gets close to the top of the hill (for the second time), she re-orients her line of sight to align it with the narrow bridge and begins walking towards the helicopter. Upon taking a few steps on the bridge, the watch tower enemy launches a rocket which explodes in front of the player without any damage. The player continues down the narrow bridge without swerving or turning to engage the enemy.

The enemy launches three more rockets, two of which explode in front of the player by making contact with the bridge. The last rocket misses the bridge completely and doesn't explode. The player runs to the side of the helicopter, gets in, and begins throttling the engine for take off. As the helicopter lifts, the enemy shoots the helicopter with a high-impact weapon. The enemy continues to fire even as the helicopter successfully takes off and flies away. As the helicopter recovers from the barrage of shots, a single rocket is fired from the watch tower which tracks the helicopter, eventually making impact and causing the vehicle to explode. The player then falls to her death.

In the R (short for "rocket") version of gameplay, the action is very similar save for one difference: When the player begins crossing the bridge for the first time, the enemy in the watch tower shoots a rocket which misses the player and makes contact behind her. The gameplay trace shows the feedback the player received when the rocket hit: the screen glows red and an arrow indicating the direction of the blast appears. The rocket launch and the player beginning to cross start simultaneously, such that the rocket makes impact just when the player abruptly navigates off it and lands on the ledge below. The rocket action was inserted to make it seem like the player fell because of the rocket (something that was initiated by the game) as opposed to falling because of a player mistake. The rest of the video continues as the NR version.

All participants watched the videos on a MacBook Pro laptop computer with a 15 inch screen running QuickTime Player version 10.0 [1]. Participants were given over-ear headphones to hear the videos without distraction. An external USB keyboard was provided to help participants with the experimental task.

3.1.3 Design and procedure

Participants were seated at a table, facing the laptop placed in front of them. After providing informed consent, participants were alternately placed into one of the two experimental groups (R or NR) and were given the following instructions:

> What I am interested in here are the units that people use to organize or break up gameplay. By that I mean that people may break up gameplay in different ways.
>
> For example, in a game, I might:
> turn, walk over, push the door closed, turn, and walk back, and you might see each of those actions as a discrete, meaningful act.
>
> Or, you might see them as just one action such as: closing the door
>
> What I want you to do is to mark off the gameplay you'll be seeing into the <u>largest</u> units that seem natural and meaningful <u>to you</u>. There are no right or wrong ways to do this; I just want to know how you do it.
>
> To identify a unit, press the `spacebar` to pause the video when, in your judgment, **one unit ends and a different one begins**. After you pause, record the timestamp on the video for the moment you paused it, and write down a brief description of the unit. Then, press `spacebar` again to continue the video.
>
> You will be given a practice video so you can familiarize yourself with the controls and the process.
>
> Both the practice video and the experimental video are recordings of a player playing a **single-player mod** of the game <u>Halo: Reach</u> for the Xbox 360. In both videos, the player is trying to finish the custom level. Both videos are approximately 1 minute long.

The instructions that we presented were varied only slightly from the instructions given to participants in the original event segmentation task experiment run by Newtson [6] in that the use the term *gameplay* instead of *film* when describing the video. In our study, participants were given a pen and two sheets of 8.5" by 11" white paper – one sheet for the practice video and one for the experiment trial. These sheets were used by the participants to help them organize the segment timestamps; those timestamps represent the breakpoints between the action segments. Participants were verbally instructed that they would be watching both the practice and experimental videos twice. The first time, they should simply observe the unfolding action. The second time, they should segment the action as instructed. Since event segmentation is a mental process that happens automatically [11], we felt that having participants watch the videos twice would help segmentation and not create an adversarial priming effect.

To reduce idiosyncratic variability in the segmentation procedure, we used a shaping process similar to one used by Zacks [13]: if a participant segmented less than four or more than ten units of action in the practice trial, we asked participants to repeat the segmentation procedure with the following comment: "Historically, participants have segmented [more/less] units than you have segmented. I would like you to adjust your grain of segmentation to be a little more [exclusive/inclusive] and repeat the practice." This was repeated at most one time. If the participant segmented less than four or more than ten units of action for the experimental trial, their data was flagged as unusable. The event thresholds of four and ten were determined by the experimenters based on extreme cases of the pilot data compiled in the design of this experiment.

In this design, we asked subjects to watch a gameplay video rather than to play a video game. This helped maintain the scientific validity of the results. If participants were given the freedom to play a game as they saw fit, there would by no way to control for each participant's experience to ensure that a basis for comparison across participants within the same group existed. By using a video that is explicitly labeled as gameplay, we assume that participants understand that the video represents an unedited recording of a player's experience in a game. No film devices which might carry narrative significance (such as "flashforwards", "flashbacks" and scene cuts) were employed. Therefore the disposition of participants towards the video should be comparable to their disposition towards video games.

After finishing the segmentation task, participants completed a short exit survey which assessed their familiarity with the game and their confidence in their own segmentation of units. Finally, the experimenter orally asked the participants: "What do you think caused the player's fall?".

3.2 Results

The results for the segmentation task are shown in Tables 1 and 2.

3.2.1 Inter-rater segmentation agreement

Before testing the hypothesis directly, we were interested in verifying whether or not the participants were segmenting in a reasonably comparable manner within their respective experimental condition groups. This verification was necessary in order to rule out the possibility that the segmentation of narratively important versus narratively unimportant events was simply due to participants segmenting the videos according to some experimentally uncontrolled criterion. To assess inter-rater segmentation agreement, we first grouped the breakpoint data into clusters. We used a k-means clustering algorithm [5] with $k = 4$. The parameter k was set to 4 since it was the lower bound used in our participant shaping process.

Table 1: Breakpoints for 8 participants in R condition

Participants							
1	2	3	4	5	6	7	8
0:14	0:10	0:10	0:02	0:09	0:16	0:09	0:13
0:17	0:17	0:18	0:09	0:23	0:23	0:16	0:17
0:28	0:26	0:43	0:16	0:42	0:40	0:43	0:26
0:42	0:42	1:01	0:29	1:02	0:43	1:02	0:39
1:00	1:02		0:39		0:49		0:43
			0:44		1:02		0:48
			1:00				0:51
							1:01

Table 2: Breakpoints for 8 participants in NR condition

Participants							
1	2	3	4	5	6	7	8
0:15	0:08	0:16	0:15	0:14	0:16	0:16	0:15
0:35	0:15	0:29	0:28	0:24	0:27	0:41	0:27
0:41	0:26	0:42	0:39	0:38	0:41	0:46	0:41
1:02	0:39	0:49	0:51	0:43	1:00	1:02	0:54
	0:48	1:02	1:02	0:49			1:02
	1:00			0:57			
				1:00			

We used each of the centroids as the respective mean of a normally distributed data cluster, i.e. we calculated a sample normal distribution about each centroid using each cluster of data. Our criterion for significant agreement followed that used by Newtson [6]: if the majority of the breakpoints for each cluster are within one standard deviation of the mean for that cluster, then we can reasonably say that participants generally agree on that area as a breakpoint for segmenting action. [3] The breakpoint data plots for each experimental condition are shown in Figures 1 and 2.

An analysis of the data about the normally distributed centroids reveals that for seven of the eight centroids, at least 66% of the data points are within one standard deviation of their respective centroid. For the one centroid that was exceptional, the corresponding cluster yielded exactly 50% of its data points within one standard deviation. We feel that this represents idiosyncrasies in segmentation rather than a sign of disagreement; the next closest points to this centroid/mean are both 1.028 standard deviations distant. Based on the previous analysis, we feel confident in stating that intra-group participants generally agree on the breakpoints for segmenting the video they saw in their respective experimental condition.

3.2.2 Segmenting narratively important events

This hypothesis depends on participants inferring that the

[3]Even though it is possible for a participant to annotate several breakpoints within the same vicinity, we feel that this effect does not manifest itself in our data in a meaningful way.

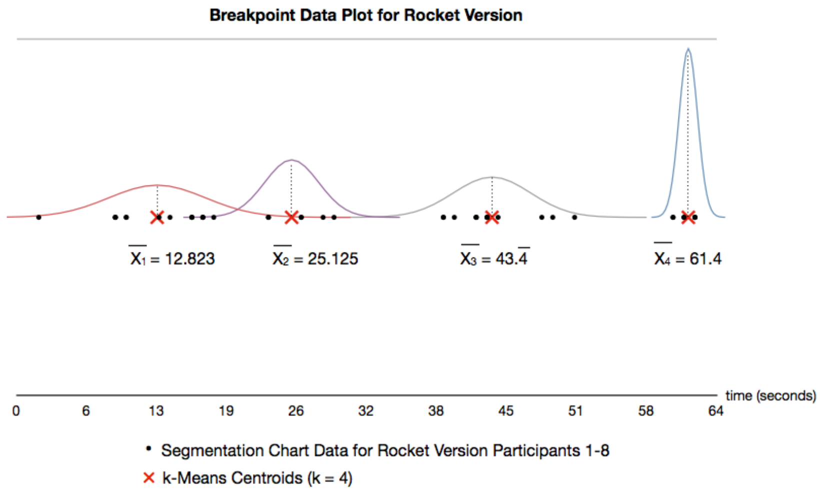

Figure 1: Breakpoint data for participants in the R condition.

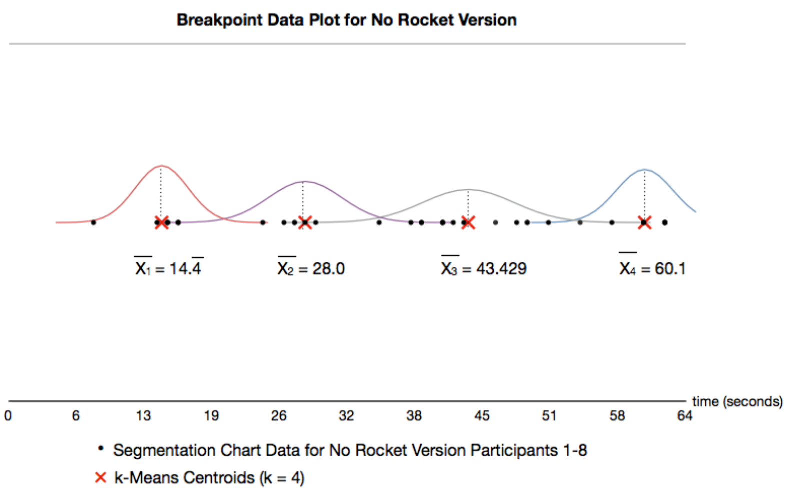

Figure 2: Breakpoint data for participants in the NR condition.

player's fall in the NR video is due to a player's mistake. This inference should then prompt the participants to recognize that there is no change in the game's intentional/goal-directed activity, and consequently, that the fall is not narratively important. In both videos, the player falls and resumes the goal-directed activity of crossing the bridge at time 0:18. For the NR version, the player finishes crossing the bridge at time 0:38 and for the R version the player finishes crossing the bridge time 0:39. Our hypothesis is operationalized as follows. For the NR version, a participant's segmenting the fall as a significant event is done by annotating a breakpoint at any time during the period from time 0:18 to time 0:38; we predicted NR participants *would not* segment during this period. For the R version, segmenting the fall as a significant event is done by annotating a breakpoint at any time during the period from time 0:18 to time 0:39; we predicted R participants *would* segment during this period. Our data can be summarized by the 2×2 contingency table, shown in table 3. By inspection, we can see that our data does not support our hypothesis. Indeed, Fisher's Exact Test [3] yields $p = 1$ and thus we fail to reject the null hypothesis $H0$.

Table 3: Contingency table for segmentation of player's fall

	Did Segment	**Did Not Segment**
R **Version**	(0:18-0:39) 6	(0:18-0:39) 2
NR **Version**	(0:18-0:38) 7	(0:18-0:38) 1

This contingency table illustrates the segmentation of the player's fall. We considered the fall as a narratively unimportant event.

3.2.3 *"What do you think caused the player's fall?"*

As mentioned in Section 3.1.3, at the end of the experiment, participants were asked an open ended question: "What do you think caused the player's fall?" Despite the question being open ended, participants roughly responded with one of the following three answers: [4]

- The player fell due to the player's mistake.
- The player fell due to the incoming rocket.
- I am not sure why the player fell.

Even though our experiment was not designed around an analysis of the open ended question specifically, our original hypothesis depended on participants recognizing that the player's fall in the NR video is due to a player's mistake. Thus, we present the open-ended question data summarized in the 2×3 contingency table shown in Table 4.

Table 4: Contingency table for interpretation of player's fall

	The player fell due to...		
	Mistake	**Rocket**	**Unsure**
R **Version**	2	5	1
NR **Version**	7	0	1

This contingency table illustrates participant interpretation of the player's fall. We considered the fall as a narratively unimportant event.

Calculating Fisher's Exact Test [3] on this data yields a p-value of: $p = 0.01259$. This implies that there is at least significant interaction, and the data suggests that participants are capable of identifying the player's mistake for the set of experimental materials (videos) we used.

4. DISCUSSION

We identified several interesting phenomena during the course of this experiment. Firstly, participants generally agreed on the segmentation of events in their respective experiment condition groups. This is consistent with previous work in event segmentation and serves to further validate the applicability of EST for understanding the perception of ongoing activity. Secondly, while the participants did not distinguish between narratively important/unimportant events as we defined them, they were still capable of determining when the player fell due to error versus an external cause. This is interesting because it suggests that an external reviewer can accurately attribute causal structure in a virtual environment that he or she is not participating in. Lastly, participants did not distinguish between narratively important and narratively unimportant events, as we framed them in this experiment. There are several potential explanations for this:

1. Narrative importance is perceptual, as opposed to structural – We considered an event to be narratively important if it was either initiated/caused by the game itself or if the event furthered the player's goals. However, this definition of narrative importance is structural, and does not consider what criterion a player uses to determine narrative importance. Despite an event not being initiated/caused by the game, nor furthering the player's goals, it is possible a player perceives an event to be important due to some other interest in the event. It remains to be seen as to what contributes to this interest, which may be due to attributes such as personal preference or game story context.

2. The player's mistake in our experiment should be considered narratively important – As counterintuitive as this point may be, it is possible that, because the user mistake in our experiment led to a change in state (the player now has to overcome having fallen off the bridge), it is actually narratively important perceptually, as opposed to structurally, as was described in the previous point. Since the player in an interactive narrative is, in essence, creating a narrative as she plays, all events that she effects or that happen to her could be narratively important perceptually. It may be the case that a player mistake becomes narratively important if it causally contributes to a change in state; thus, inconsequential errors in gameplay might be unimportant enough to be ignored in event segmentation. For exploratory behavior, the aforementioned revision would imply that if the exploration leads to a change in state, then the behavior is narratively important.

3. Participants perceived the gameplay videos in the same way they would conventional narrative videos – Despite the plausible assumption that participants should perceive the non-interactive video in the context of the game experience it records, it is possible that, because they are not able to exert direct control over the experience they see, participants might still view the stimulus as a film. In this case, participants would not discriminate between events they see. Everything in the film would have narrative significance.

4. Participants focused their segmentation tasks on *gameplay* rather than on narrative significance. – For this explanation, we will revisit a portion of the instructions participants received in the experiment:

   > What I am interested in here are the units that people use to organize or break up *gameplay*. By that I mean that people may break up *gameplay* in different ways.
   >
   > (Emphasis added here.)

   It is possible that falling to the ledge in the NR condition is considered by participants as an important part of the gameplay in the video. Falling bears no relevance to the progression of the story, but it is still an obstacle that interrupts a player's gameplay experience and might therefore be considered an important break in the action.

5. CONCLUSIONS AND FUTURE WORK

Our interest in the segmentation of events and EST [6, 12] originated in our desire to directly use the situation model

[4] While the individual answers varied in the level of detail, the intent of all participants' responses was clear enough to categorize into these three groups.

work [16] as a model of the player's story comprehension processes in interactive narratives and video games. We were interested in determining whether or not, for purposes of situation models, a distinction between "narratively important" and "narratively unimportant" events was necessary. Without this characterization, our model of the player's mental state as he or she experiences the unfolding story of a game could be inaccurate due to the introduction of model-inappropriate elements. This would could lead to misrepresenting the player's mental state and could interfere with a system's ability to effectively manipulate the player's cognitive processes in the feedback loop that leads

1. From the game to the player - The game presents the story such that the player understands how to act in the environment.
2. From the player to the game - The player acts in the environment according to what she perceives her role is, how the story has progressed so far and other contextual elements.
3. The game reasons about the actions the player has taken in the environment and adjusts the story presentation to facilitate the player's comprehension and engagement.

   (back to 1)

Unfortunately, given our data, we still cannot make a conclusive statement. Our original distinction was structural, but future work will be oriented towards a more perceptual account of narrative importance. Although our data failed to reject the null hypothesis, more work must be done before rejecting the necessity of the distinction between narratively important and unimportant events when modeling a player's comprehension process of an unfolding story-focused game.

6. ACKNOWLEDGMENTS

We would like to thank Francisco J. Cardona-Rivera and Allison G. Martínez-Arocho for their help in creating the experimental videos. We would also like to thank the members of the Liquid Narrative Group for their insightful discussions, as well as Dr. Jeffrey M. Zacks for his brief comments. Finally, we thank the anonymous reviewers for their comments on an earlier version of this paper. This research was supported by the Department of Energy Computational Science Graduate Fellowship under DoE Grant DE-FG02-97ER25308.

7. REFERENCES

[1] Apple Inc. Quicktime player [software].

[2] Bungie. Halo: Reach [video game], September 2010.

[3] R. Fisher. On the interpretation of χ 2 from contingency tables, and the calculation of p. *Journal of the Royal Statistical Society*, 85(1):87–94, 1922.

[4] H. P. Grice. Meaning. *The Philosophical Review*, 66(3):377–388, 1957.

[5] J. A. Hartigan and M. A. Wong. A K-means clustering algorithm. *Applied Statistics*, 28:100–108, 1979.

[6] D. Newtson. Attribution and the Unit of Perception of Ongoing Behavior. *Journal of Personality and Social Psychology*, 28(1):28–38, 1973.

[7] J. Rall and P. L. Harris. In Cinderella's slippers? Story comprehension from the protagonist's point of view. *Developmental Psychology*, 36(2), 2000.

[8] R. M. Young. The Cooperative Contract in Interactive Entertainment. In *Socially Intelligent Agents*, chapter 1, pages 229–234. Kluwer Academic Press, 2002.

[9] R. M. Young and R. E. Cardona-Rivera. Approaching a Player Model of Game Story Comprehension Through Affordance in Interactive Narrative. In *Proceedings of the 4th Workshop on Intelligent Narrative Technologies*, pages 123–130, 2011.

[10] J. M. Zacks. Using movement and intentions to understand simple events. *Cognitive Science*, 28:979–1008, 2004.

[11] J. M. Zacks, T. S. Braver, M. A. Sheridan, D. I. Donaldson, A. Z. Snyder, J. M. Ollinger, R. L. Buckner, and M. E. Raichle. Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4(6):651–655, 2001.

[12] J. M. Zacks, N. K. Speer, and J. R. Reynolds. Segmentation in Reading and Film Comprehension. *Journal of Experimental Psychology: General*, 138(2):307–27, May 2009.

[13] J. M. Zacks, N. K. Speer, J. M. Vettel, and L. L. Jacoby. Event Understanding and Memory in Healthy Aging and Dementia of the Alzheimer Type. *Psychology and Aging*, 21(3):446–482, 2006.

[14] J. M. Zacks and K. M. Swallow. Event Segmentation. *Current Directions is Psychological Science*, 16(2):80–84, Oct. 2007.

[15] J. M. Zacks and B. Tversky. Event structure in perception and conception. *Psychological Bulletin*, 127:3–21, 2001.

[16] R. A. Zwaan and G. A. Radvansky. Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2):162–85, Mar. 1998.