Proceedings of the Tenth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2014)

Foreseeing Meaningful Choices

Rogelio E. Cardona-Rivera*, **Justus Robertson***, **Stephen G. Ware**, **Brent Harrison**,
David L. Roberts and **R. Michael Young**
Digital Games Research Center
Department of Computer Science
North Carolina State University
Raleigh, NC, USA 27695-8206
{recardon | jjrobert | sgware | beharri5}@ncsu.edu, {robertsd | young}@csc.ncsu.edu

Abstract

A choice positively contributes to a player's sense of agency when it leads to meaningfully different content. We shed light on what a player may consider meaningfully different by developing a formalism for interactive stories in terms of the change in *situational content* across choices. We hypothesized that a player will feel a higher sense of agency when making a choice if they foresee the available actions lead to meaningfully different states. We experimentally tested our formalism's ability to characterize choices that elicit a higher sense of agency and present evidence that supports our claim. Study participants ($n = 88$) played a choose-your-own-adventure game and reported a higher sense of agency when faced with choices that differed in situational content over choices that didn't, despite these choices differing in non-situational ways. We contend our findings are a step toward principled approaches to the design of interactive stories that target specific cognitive and affective states.

Introduction

In this paper we aim to identify what players perceive as interesting choices within interactive narratives. We want to understand how players perceive distinct and meaningfully different story content, and how that perception affects their sense of *agency*. We adopt Murray's (1997) perspective, defining agency as the satisfying power to take meaningful action and see the results of our decisions and choices. Our concern is agency tied to determining the outcome of a story's development as opposed to other kinds of agency of interest (Harrell and Zhu 2009). We hypothesize that when a player faces a choice with options that elicit meaningfully different outcomes, the player will feel a higher sense of agency than when faced with choices that do not.

Choice, or the ability to take action and effect change, is a central component of gameplay. Salen and Zimmerman (2003) discuss choices in the context of *meaningful play*, which occurs when "the relationships between actions and outcomes are both *discernible* and *integrated* into the larger context of the game" (Salen and Zimmerman 2003, p. 34). A game action is discernible when its results are communicated in a perceivable way (i.e. through short-term feedback), and is integrated when it holds global significance during the play experience by affecting long-term development (i.e. through long-term feedback).

*The first and second authors made equal contributions and should be considered co-first authors.

Related Work

In interactive narrative games, it is natural to conclude that agency is best elicited through highly branching story structures that allow many possible player actions. Fendt et al. (2012) challenged this assumption, and attempted to elicit agency in a mostly-linear choose-your-own-adventure. They hypothesized that the illusion of agency can be elicited by immediately acknowledging results of player actions while progressing through a fairly linear story. To test their hypothesis Fendt et al. prepared three treatments of a choose-your-own-adventure: *S1*) a branching story with short-term and long-term feedback *S2*) a non-branching story with short-term feedback but no long-term feedback and *S3*) a non-branching story with no short-term or long-term feedback. Study participants were asked to play through one of the three treatments and self-report their perceived agency for every choice. One interesting result was that half of the six choices were rated lower in agency across all treatments when compared to the remaining three. For example:

> **Choice 1** You are the great adventurer Stump Junkman. You are employed by the king to hunt down and destroy evil creatures wherever they may be. The king suspects that one of these horrible creatures has stolen his Crown of Power. It is your duty to find this monster, slay it, and retrieve the Crown. You don your armor, your usual adventuring supplies, and reach for a weapon. Do you choose:
>
> 1.) Your sword and shield
>
> 2.) Your crossbow

was consistently rated among all treatments as having lower agency than the second choice:

> **Choice 2** You leave your house and travel to the outskirts of the city. There are two possible locations to explore. Do you choose:
>
> 1.) The forest to the north of town
>
> 2.) The mountains to the east of town

Fendt et al. discussed two possible causes for this phenomenon: *a*) higher agency decisions have more severe consequences for failure or *b*) higher agency decisions offer distinguishably different story paths. In this paper we explore the second of these possibilities. We hypothesized that when making a choice a player will feel a higher sense of agency if they perceive the outcomes of their possible actions to be meaningfully different world states, than when outcomes are perceived as not meaningfully different. We developed a formalism to model how people perceive the outcomes of their actions, tested the predictive power of our model in a new choose-your-own-adventure study, and found that our representation successfully characterizes Fendt et al.'s phenomenon.

A Formalism of Story Content

To test our hypothesis we must identify what is perceived as meaningfully different story content. For this we leverage a cognitive model of narrative comprehension, the Event-Indexing Model (Zwaan, Langston, and Graesser 1995).

Event-Indexing Models

Situation models (van Dijk and Kintsch 1983) are integrated mental models of particular situations in a story world made up from an amalgamation of information explicitly present in the narrative as well as information inferred by the human consumer (Zwaan and Radvansky 1998). The Event-Indexing Model (EIM) is an empirically supported situation model theory, which posits that when people consume a narrative they discretize it into constituent events. Human story consumers segment events in the same manner (Newtson 1973), and segmentation occurs when there is a discontinuity in one of the EIM dimensions (Radvansky and Zacks 2011), which (at least) include 1) the *time* the event took place, 2) the *space* in which it took place, 3) the event's *causal status* with regards to prior events, 4) the event's relatedness to the *goals* of characters, and 5) the main *characters* for the event itself. Story consumers mentally index events along the EIM dimensions.

Interactive Narrative EIM Formalism

A *situation vector* models a person's mental representation of a story situation.

Definition 1 (Situation Vector). *A* situation vector *is a quadruple $\mathcal{SV} = \langle S, T, G, C\rangle$ where S denotes the perceived space of the situation, T denotes the perceived time of the situation, G denotes the perceived goals of the player character, and C denotes the perceived characters that are involved in the situation.*

An example situtation vector $\mathcal{SV}_2$ of Fendt et al.'s second choice might be: S) the outskirts of the city G) find the Crown of Power and C) the protagonist. The T component of the situation vector is left unbound because nothing in the prompt specifies time information. This definition allows us to distinguish any two situations as different according to the EIM. Two situations are different when $\Delta(\mathcal{SV}_1, \mathcal{SV}_2) > 0$.

Definition 2 (Situation Change Function). *A* situation change function *is a function $\Delta(\mathcal{SV}_1, \mathcal{SV}_2) \in [0, 4]$ that takes as input two situation vectors and returns the number of unequal indices between the two vectors.*

EIM indices represent important factors that contribute to how people segment and store narrative events. We do not model perceived causality in our situation vector because the actions in our choose-your-own-adventure game are along the same causal path. Our hypothesis is that people will feel a higher sense of agency when they foresee their choice options lead to meaningfully different story content. Humans segment a story when two events differ along one of the EIM's indices. We reason that through the process of segmentation human story consumers make a meaningful distinction between story events that are perceived as having a discontinuity on any EIM index. For this reason we propose that in the context of an interactive narrative's story *two story situations are meaningfully different when they differ along at least one of the EIM dimensions* or $\Delta(\mathcal{SV}_1, \mathcal{SV}_2) > 0$. This allows us to operationalize and experimentally test our hypothesis.

Experiment

Our hypothesis distinguishes two types of choices: 1) choices that lead to outcomes that are perceived as "not different" (*Class ND choices*), and 2) choices that lead to outcomes that are perceived as "different" (*Class D choices*). The perceived difference we refer to is a person's perception of the situational content. We hypothesize that when players foresee a choice's options lead to outcomes that are perceived as different in their situational content (Class D choices), we predict they will self-report a higher sense of agency than when they foresee outcomes as equal (Class ND choices). Formally:

> H_A: Participants will self-report a lower sense of agency when they are faced with a choice whose options are perceived to lead to outcomes with a zero-weighted shared index edge between them (Class ND choices) when compared to outcomes connected with a non-zero-weighted shared index edge (Class D choices).

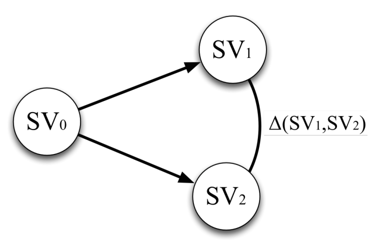

Figure 1: An example choice. $\mathcal{SV}_0$ represents the state of the world in which the player makes a choice. Directed edges represent two possible actions the player can take. $\mathcal{SV}_1$ and $\mathcal{SV}_2$ represent situations the player foresees to be the result of their possible actions. The undirected edge represents the situational change between $\mathcal{SV}_1$ and $\mathcal{SV}_2$. If $\Delta(\mathcal{SV}_1, \mathcal{SV}_2) > 0$ we predict the choice will have high relative agency.

Table 1: Statements used to assess whether choose-your-own-adventure game choice structure options were perceived as leading to different situations. Judges rated each question as "agree" or "disagree" (except the last question, rated as "yes or no") for each option within the choices; every in-game choice had 2 options.

Q1: These actions would occur in the same time.
Q2: These actions would occur in the same space.
Q3: These actions would both achieve the same goal.
Q4: These actions would involve the same characters.
Q5: These actions would involve the same items.
Q6: Is there some difference between these choices that was not covered above?

Setup

We make two key assumptions: 1) participants can readily identify when a choice structure's options lead to distinct outcomes and 2) participants agree on the general form of those distinct outcomes. We assume that when players engage with a choice structure, they will perceive the situations that result from their choices in a consistent manner as computed from their mental Situation-Change function as discussed in Definition 2. While we do not offer any claim as to the internal structure of such a function, we acknowledge it is an essential component and assume that it is consistent across people. To help support these assumptions, the designed choose-your-own-adventure game was validated by a panel of nine judges. These judges compared the options for every choice structure in the game, by agreeing or disagreeing to six statements (outlined in Table 1), designed to evaluate the similarity between the situations that result from a player's choice. The statements evaluate similarity between options as a proxy for evaluating similarity between outcomes. The judges were all from the authors' institution, and were not trained to perform the judging task in any way prior to issuing the judgments. We computed inter-rater reliability metrics (Fleiss 1971) across the EIM dimensions individually, as well as across the EIM dimensions grouped by choice structure. As noted in the second column of Table 2, we achieved a significant, fair to moderate level of inter-rater reliability for judgments on how choice outcomes compared in terms of their situational content. We are confident this agreement is representative of how players experience options in our choose-your-own-adventure.

To decide the number of situation shifts present at each choice structure we looked at individual judgments and tallied agreements and disagreements: if a simple majority (five or more) of judges disagreed that an index remained the same then we noted the index as changed. As noted in the last column of Table 3, choices 1, 2, and 4 were rated as not differing on any situational dimension, and therefore form our Class ND of choices. Additionally, choices 0, 3, and 5 were rated as differing on at least 1 situational dimension, and therefore form our Class D of choices.

Although our analysis here focuses on the four dimensions of time, space, goal, and characters, we instructed judges to rate situational dimensions beyond our formalism. We coded additional dimensions to address the possibility that participants see choice options that do not result in different perceived situations as a "false choice," which we expect to have intrinsic low agency. We explicitly asked judges to code for a change in items since it may potentially be important for event-indexing (Radvansky 2012). However, the number of other dimensions on which a situation could vary is potentially infinite, and we therefore tasked raters to identify other potential differences in a general fashion. All three options that were identified as not different per our model were rated by judges as having at least one other difference. Because of this, we feel that choice options are not perceived as being the same.

Table 2: Inter-rater reliability metrics computed for nine (9) raters, for situational dimensions (in the upper half), and the six choices (in the lower half), over the labels "agree" / "disagree" to the questions in Table 1. The "yes" / "no" question was reverse coded to preserve the labeling of "agree" / "disagree." The judging served as a proxy for assessing whether people could agree on how different choice structure options lead to distinct outcomes. Values are statistically significant to the 0.05 level, unless otherwise noted.

Agreement Criterion	**Fleiss'** $\kappa \in [0, 1]$	**Z-score**	**Interpretation** (Landis and Koch 1977)
Space change	0.250	3.67	Fair agreement
Time change	0.201	2.95	Slight agreement
Goal change	0.437	6.43	Moderage agreement
Character change	0.400	5.88	Fair agreement
Item change	0.330	4.84	Fair agreement
Other change	0.114	1.67	Slight agreement ($p < 0.1$)
C0 dimension change	0.215	3.16	Fair agreement
C1 dimension change	0.262	3.84	Fair agreement
C2 dimension change	0.199	2.93	Slight agreement
C3 dimension change	0.662	9.74	Substantial agreement
C4 dimension change	0.137	2.01	Slight agreement
C5 dimension change	0.417	6.14	Moderate agreement

Method

The experiment evaluated whether or not choices that lead to meaningfully different situations could be used to predict a participant's self-reported sense of agency in a choose-your-own-adventure game. The experiment was a within-subjects repeated-measures design with all participants completing the same choose-your-own-adventure game. The game had three (Class ND) choices with no situational difference between outcomes and three (Class D) choices with a non-zero situational difference between outcomes. The order in which participants experienced the different classes of choices was consistent and had no discernible pattern. Upon exiting the game participants answered a survey regarding their experience. The independent variable that was fixed was the number of index values that choice outcomes shared. The dependent variables were the surveys collected after participants finished playing.

Participants 88 participants (63% M, 33% F, 4% other) 18 years or older ($M = 27.5$, $SD = 10.42$) took part in this experiment. Of these 95% reported having played games for more than five years, 63% considered themselves "gamers," and 84% rated themselves as having native proficiency with the English language. Participants were recruited online, and were encouraged to recruit others for the study. Participants were offered no compensation for participating.

Table 3: The choice structure options in our choose-your-own-adventure game were judged through answers to the questions of Table 1. If a simple majority (five or more) of raters agreed that a particular (space, time, goal, or character) index dimension was different between both choice options, then it was annotated as changed, and it counted toward the total reported in the "EIM changes (Δ)" column.

	Judges felt the options differed in						EIM
Choice structure options	**space**	**time**	**goal**	**char.**	**item**	**other**	**Δ**
0a) Take aim and throw the reading lamp at the intruder. 0b) Sneak up on the intruder and knock him out with the lamp.	1	3	3	0	0	6	0
1a) Stay here and help the guard nurse his wound while the intruder escapes. 1b) Leave the museum and help the injured guard chase down the intruder.	2	5	9	6	3	1	3
2a) Trust the man claiming to be a holy knight and let him explain everything to you. 2b) Spend the rest of the day reading the book for yourself to find an explanation.	6	7	1	7	6	3	3
3a) Take the ancient spellbook with you. 3b) Load your gun and take that with you.	0	0	0	0	8	0	0
4a) Let Alexander charge in and fight the goblins here on the dock so that they can't steal the crown. 4b) Sneak aboard the ship and search for the crown before the goblins have a chance to find it.	1	7	4	5	4	1	2
5a) Melt down the Crown of Power before someone else tries to use it for evil purposes. 5b) Burn the spell book so that the Crown of Power's secrets can never be discovered.	1	2	1	0	8	1	0

Apparatus The choose-your-own-adventure was hosted online, and appeared as centralized text at 14 point font with choice options placed at the bottom of the choice framing as hyperlinked text. The game used JavaScript to reveal the text of the story by paragraph blocks on command. When the last block was revealed, so too were the choice options. Participants accessed the game via their browser. The game was designed to monitor whether or not participants attempted to experience the game out of linear order and were warned against doing so. If participants insisted they were removed from the data set.

Stimuli The game was a text-based choose-your-own-adventure game with six binary choices. Choice option pairs were kept within one syllable of each other. The stimuli acknowledged player choices through both short-term and long-term feedback. Short-term feedback was implemented by referencing the player's choice in the outcome text that followed the selection. Long-term feedback for the n-th choice was implemented by referencing the player's choice in the outcome text that followed the selection at choice $n + 2$; i.e. the long-term effect of a player's choice was evident two choices later. Players received no long-term feedback for choices 4 and 5.

Procedure After providing informed consent, players filled out a demographic survey. They were subsequently instructed to fill out a self-efficacy survey (Chen, Gully, and Eden 2001). Upon completion, participants were asked to play the stimuli. Afterwards, they completed an intrinsic motivation survey (Deci and Ryan 1985), and a modified version of the interactive story assessment survey developed by Vermeulen et al. (2010). We adapted the latter survey to assess feelings of agency at an individual choice level. These statements (hereafter *agency scales*) are five-point Likert scales, ranging from "Strongly Disagree" to "Strongly Agree." The agency scales are the primary focus of analysis of this paper and are presented in Table 4. Like Fendt et al. (2012), individual choice agency was assessed by presenting to players their choice as they experienced it, and then having players answer survey items at that choice point.

Table 4: Agency scales used to assess a player's self-reported sense of agency, modified from the survey instrument developed by Vermeulen et al. (2010). These statements were presented as 5-point Likert-scale prompts (ranging from "Strongly Disagree" to "Strongly Agree") for every choice point that players experienced.

S1:	This choice had considerable impact on the events in the story.
S2:	The consequences of this choice were clearly visible.
S3:	I could recognize which events in the story I have caused with this choice.
S4:	My decision here clearly influenced how the story went on.
S5:	I discovered how this action influenced what happened later in the story.

Results, Analysis, and Discussion

While coders were able to agree that there existed choices that were intrinsically different (Class D v. Class ND) as defined by our formalism, we wanted to see whether our data supported the judgments as well. If there exist two different classes of choices then we should expect that intra-class choices will not differ in terms of their median agency scale reports in a statistically significant way. Because the Likert-scale assessments for the 3 Class D and 3 Class ND choices were obtained within-subjects (each participant was independent, but provided five ratings for each of the six choices), we analyzed the data using a Friedman (1937) Two-Way ANOVA. Every choice (in Table 3) was rated by a participant five times, once for every agency scale in Table 4. Both choice classes (D and ND) had three members each. The ANOVA compared the median score of every agency scale, for all intra-class choices against each other (for a total of three comparisons per choice class). In this analysis we were looking for data to support the null hypothesis as opposed to reject it, because we wanted there to be no difference at the intra-class level.

The ANOVA results suggest that Class ND choices formed a cohesive class, with two of the three intra-class comparisons resulting statistically indistinguishable. Class D choices, however, seemed to not form a cohesive class, with two of the three intra-class comparisons resulting statistically different: the agency reports for choices 1 and 2,

Table 5: Friedman Two-Way ANOVA for Intra-class choices across agency scales. If a cell has a ✗, it means that for that row's choice pair, we reject the null hypothesis ($p < 0.05$) that the median scores for that column's agency scale are drawn from the same population (i.e. the choices of the pair are statistically different). Conversely, if a cell has a ✓, it means that for that row's choice pair, we fail to reject the null (i.e. the choices of the pair are statistically indistinguishable).

Intra-class Choice Pairs		S1	S2	S3	S4	S5
Class ND	C0-C3	✗	✓	✗	✓	✗
	C0-C5	✓	✓	✓	✓	✓
	C3-C5	✓	✓	✓	✓	✓
Class D	C1-C2	✗	✗	✗	✓	✗
	C1-C4	✓	✓	✓	✓	✓
	C2-C4	✗	✗	✗	✗	✗

as well as the agency reports for choices 2 and 4 were statistically different (see Table 5 for all intra-class agency scale comparisons). This assessment does not mean that the choices could not form part of the same class, but rather that it is possible that within the same class there are intrinsic agency differences that are not accounted for by our formalism. However, this assessment does imply that (for the purpose of statistically verifying our hypothesis) we could not group the choices by their initially judged class. To better understand our data, we graphed (Figure 2) five parallel coordinate plots, one for each statement in our agency scales. Each horizontal line in this plot corresponds to a player's trajectory of choice agency ratings, through the respective agency scale. Through visual inspection we identified that the median agency scores for Class D choices (choices 1, 2, and 4) were consistently reported as "Agree" for all agency scales. Class ND choices were consistently rated at or below Class D choices, lending support to our hypothesis. Given the ANOVA data of Table 5, and the plots in Figure 2, we operationalized our hypothesis as follows:

$$\text{H}_0\text{: } Md_{Ci} = 0$$

$$\text{H}_A\text{: } \underbrace{Md_{C0} < Md_{C5} < Md_{C3}}_{\text{Class ND Choices}} < \underbrace{Md_{C1} < Md_{C2} < Md_{C4}}_{\text{Class D Choices}}$$

where Md_{Ci} is the median rating for choice structure Ci.

We calculated the Page (1963) trend test for ordered alternatives, given H_A, for all agency scales. The Page test has more statistical power than the Friedman test, since the latter only considers the alternative hypothesis that the central tendencies (in our case, the median ratings) are different

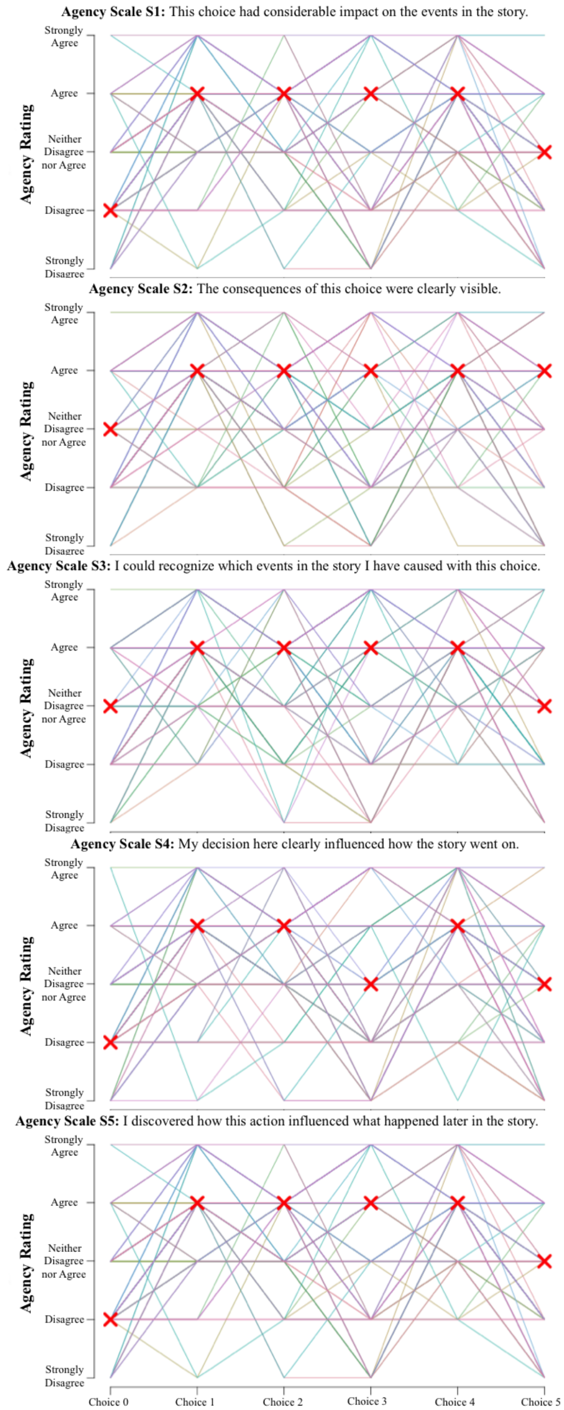

Figure 2: Agency scale parallel coordinate plots. Each plot identifies (×) median agency across the choices in Table 3.

without specifying their order. The results of the test can be seen in Table 6, and were conclusive and positive: we reject the null hypothesis for agency scales S1 through S3. Upon closer inspection, we realized that agency scale S5 (in Table 4) was targeting long-term feedback within the story. Since choices 4 and 5 do not receive long-term feedback by design, we dropped them from the analysis of S5, and were

Table 6: Page trend test for ordered alternatives for each of the agency scales in Table 4. We were able to reject the null hypothesis in favor of our alternate hypothesis for agency scales S1, S2, S3, and S5′.

Agency Scale	L-statistic	Z-score	p-value
S1	6353	3.09	0.001
S2	6519	11.94	≈0
S3	6464	13.79	≈0
S4	6256	-6.27	0.999
S5	6342	-5.03	0.999
S5′	2922	138.71	0

able to reject the null hypothesis. Thus, only for S4 ("My decision here clearly influenced how the story went on.") were we unable to reject the null. Despite this, Figure 2 illustrates that S4 was the only scale on which our prediction of Class ND choices all rated (in terms of median score) below Class D choices is true. Because all choice points had short-term feedback that acknowledged the player's choice immediately, all choice points clearly indicated how they influenced story going forward, possibly skewing the data in a way that makes the trend undetectable. In light of the evidence presented here, we reject the null hypothesis in favor of our alternate: *participants do self-report a lower sense of agency when they are faced with Class ND choice, when compared to a Class D choice*.

Limitations and Future Work

While we have made progress toward explaining how people cognitively engage with interactive stories we recognize that several factors could potentially be at play beyond what was measured and recorded here. Future work should expand this study to include longer interactive story experiences, structural EIM equivalents in other story genres, as well as more heterogeneous populations.

We also recognize the epistemic limitations of our study. Participants reported their sense of agency for each choice by revisiting the choices they made after they had already made them. In other words, when players were asked to revisit their choice and self-assess their sense of agency, they already had the benefit of seeing how the story resulted after they made their choice. While we feel participants are able to project themselves to respond as if they were at the moment of choice, knowing the partial outcome of the story undoubtedly has some effect in their sense of agency. Participants only experienced one trajectory through the interactive narrative, such that when they revisit a choice in the survey, they still must project an outcome for the choice they have not experienced. We posit that this cognitive engagement serves as an adequate proxy for how participants generally engage with interactive narrative choose-your-own-adventure games. Future work should address online agency assessment, by (e.g.) interrupting participants before they commit to a choice in the game. We note, however, that interrupting participants during gameplay may adversely affect their cognitive engagement, and possibly their sense of agency.

Notwithstanding the aforementioned limitations, we feel that there are ways to leverage the insights gathered here. One promising path is through intelligent expansion of a story graph during interactive narrative generation. In interactive narrative systems, modeling interesting plot sequences that respond to a variety of user interactions produces highly branching graphs of story content (Bruckman 1990). A process called *mediation* (Robertson and Young 2013) is one artificial intelligence planning-based method (Young et al. Forthcoming 2014) for story graph creation. Mediation builds a branching story graph using *accommodation* and *intervention*. Accommodation expands a story graph by adding new content, and intervention prunes graph branches by preventing user actions. Incorporating declarative knowledge of intrinsic high-agency versus low-agency choices into mediation may help answer the open research question (Young 1999) of when best to use accommodation or intervention while building a story graph: if a choice has low intrinsic agency the cost of intervening may be negligible compared to the benefits of preserved coherence, and a decreased branching factor. Another promising path is using our formalism to generate discourse that elicits in players *narrative affordances* (Young and Cardona-Rivera 2011), sequences of story content that players envision as potential completions to their current story experience. Introducing feedback during gameplay may prompt players to think of an intrinsically low-agency choice as a high-agency choice, and vice-versa. Players may then pursue or abandon choices depending on their perceived sense of control during the course of gameplay, and interactive narrative designers could design stories to take advantage of how players engage.

Conclusion

In this study we have provided evidence that people self-report higher feelings of agency when faced with choice options that are perceived to result in meaningfully different states, in contrast to choice options that are perceived to result in equivalent states. We characterize "meaningfully different" in terms of situational content, inspired by efforts in cognitive psychology that attempt to understand the mind in the context of non-interactive stories. In essence, we have argued and provided evidence to suggest that the mental effort required from a person to distinguish choice outcomes as equivalent or not is sufficient to characterize that person's sense of meaningful action in interactive stories.

We conclude that we have successfully presented a formalism that serves as a tool for characterizing how people cognitively engage with interactive narrative choices. In addition, we have demonstrated the formalism's effectiveness in explaining intrinsic high-agency and low-agency choices as originally identified by Fendt et al. (2012). We contend that our work is an important step toward the automated design and adaptation of interactive stories. By characterizing interactive narrative choices in terms of their perceived outcomes, and using that characterization to predict a player's sense of agency, we have begun to accomplish what Szilas (2010) has suggested: a computational model of interactive stories that goes beyond story structure, and accounts for the effect of the artifact on the human consumer.

Acknowledgements

This work was supported in part by the U.S. Department of Energy under grant number DE-FG02-97ER25308.

References

Bruckman, A. 1990. The Combinatorics of Storytelling: Mystery Train Interactive. Technical report, MIT Media Lab.

Chen, G.; Gully, S. M.; and Eden, D. 2001. Validation of a New General Self-Efficacy Scale. *Organizational Research Methods* 4(1):62–83.

Deci, E. L., and Ryan, R. M. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press.

Fendt, M. W.; Harrison, B.; Ware, S. G.; Cardona-Rivera, R. E.; and Roberts, D. L. 2012. Achieving the Illusion of Agency. In *Proceedings of the 5th International Conference on Interactive Digital Storytelling*, 114–125.

Fleiss, J. L. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin* 76(5):378–382.

Friedman, M. 1937. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *Journal of the American Statistical Association* 32(200):675–701.

Harrell, D. F., and Zhu, J. 2009. Agency Play: Dimensions of Agency for Interactive Narrative Design. In *Proceedings of the AAAI Spring Symposium on Intelligent Narrative Technologies*, 44–52.

Johnson-Laird, P. N. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, MA, USA: Harvard University Press.

Landis, R. J., and Koch, G. G. 1977. The Measurement of Observer Agreement for Categorial Data. *Biometrics* 33(1):159–174.

Mawhorter, P.; Mateas, M.; Wardrip-Fruin, N.; and Jhala, A. 2014. Towards a Theory of Choice Poetics. In *Proceedings of the 9th International Conference on the Foundations of Digital Games*.

Murray, J. H. 1997. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. Simon and Schuster.

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

Page, E. B. 1963. Ordered Hypotheses for Multiple Treatments: A Significance Test for Linear Ranks. *Journal of the American Statistical Association* 58(301).

Radvansky, G. A., and Zacks, J. M. 2011. Event Perception. *Wiley Interdisciplinary Reviews: Cognitive Science* 2(6):608–620.

Radvansky, G. A. 2012. Across the Event Horizon. *Current Directions in Psychological Science* 21(4):269–272.

Riedl, M. O., and Young, R. M. 2006. From Linear Story Generation to Branching Story Graphs. *IEEE Computer Graphics and Applications* 26(3):23–31.

Robertson, J., and Young, R. M. 2013. Modelling Character Knowledge in Plan-Based Interactive Narrative to Extend Accomodative Mediation. In *Proceedings of the 6th Workshop on Intelligent Narrative Technologies*, 93–96.

Salen, K., and Zimmerman, E. 2003. *Rules of Play: Game Design Fundamentals*. MIT Press.

Szilas, N. 2010. Requirements for Computational Models of Interactive Narrative. In *AAAI Fall Symposium on Computational Models of Narrative*, 62–68.

van Dijk, T. A., and Kintsch, W. 1983. *Strategies of Discourse Comprehension*. Academic Press: New York.

Vermeulen, I. E.; Roth, C.; Vorderer, P.; and Klimmt, C. 2010. Measuring User Responses to Interactive Stories: Towards a Standardized Assessment Tool. In *Proceedings of the 3rd International Conference on Interactive Digital Storytelling*.

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

Young, R. M.; Ware, S. G.; Cassell, B. A.; and Robertson, J. Forthcoming 2014. Plans and Planning in Narrative Generation: A Review of Plan-Based Approaches to the Generation of Story, Discourse, and Interactivity in Narratives. *SDV. Sprache und Datenverarbeitung*.

Young, R. M. 1999. Notes on the Use of Plan Structures in the Creation of Interactive Plot. In *AAAI Fall Symposium on Narrative Intelligence*, 164–167.

Zwaan, R. A., and Radvansky, G. A. 1998. Situation Models in Language Comprehension and Memory. *Psychological Bulletin* 123(2):162–185.

Zwaan, R. A.; Langston, M. C.; and Graesser, A. C. 1995. The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science* 6(5):292–297.