Interaction in Surround Video: The Effect of Auditory Feedback on Enjoyment

Mirjam Vosmeer[1(✉)], Christian Roth[2], and Ben Schouten[1]

[1] Interaction & Games Lab, Amsterdam University of Applied Sciences,
Postbus 1025, 1000 BA Amsterdam, The Netherlands
{m.s.vosmeer, b.a.m.schouten}@hva.nl
[2] Department of Communication Science, VU University Amsterdam,
Amsterdam, Netherlands
roth@spieleforschung.de

Abstract. This study investigates whether an interactive surround video is perceived as more enjoyable when there is some auditory feedback on interactive moments. We constructed a questionnaire that measured presence, effectance, autonomy, flow, enjoyment, system usability, user satisfaction and identification, filled in by two groups of respondents who had either watched an interactive movie on Oculus Rift with feedback sounds, or a version without. Our results show that users rated presence significantly lower in the feedback condition. We rejected our hypothesis, that auditory feedback would increase the perception of effectance.

Keywords: Interactive cinema · Video games · Presence · Immersion · Enjoyment · Flow · Surround video · Oculus rift

1 Introduction

The introduction of a series of media systems that allow users to experience movies in 360° has opened up a whole new range of research questions into the field of Interactive Storytelling (IS). Surround video is the term that is often used for this type of video [1], that can be watched with for instance Oculus Rift, Samsung Gear Port VR or the - technologically less advanced but surprisingly effective - Google Cardboard. As implied by the term, with this kind of media the user is able to look all around - and often also above and underneath -, from a viewpoint that feels as if he or she is actually positioned in the middle of the movie. Some of the challenges that producers face, when aiming to capture this kind of real live video footage, have been pointed out by [2]. However, also media researchers are challenged by new issues that might require them to re-evaluate existing theoretical concepts and to investigate how media productions like this may be enjoyed by lay audiences. After all, only appreciation among a general public will eventually lead to market success. A defining quality of surround video is that it seems to merge characteristics of both movie and video game. The viewer is able to look all around and study elements that are displayed in the movie, such as actors and props, without his or her gaze being guided by the frames that have been decided upon by the director. This point of view reminds of playing an open

world video game. To move the narrative forwards, interactive elements can be added to the video, which is also reminiscent of playing a game.

The difference in engagement style between the passive position of someone who is watching a movie and the active position of a person playing a video game has been referred to as *lean back* versus *lean forward* [3]. Vosmeer and Schouten [2] have proposed to use the term *lean in* for the engagement style with media such as surround video, in which the user position is neither fully lean-back nor completely lean forward. For these lean in media, concepts such as presence, flow, effectance and enjoyment may be related in different ways than previously discussed in the field of Interactive Storytelling. And since the impact of interaction on user engagement might take different forms as well, it is useful to study the user experience with IS from this new perspective. For the current research project, we have focused on the way that the perception of interaction in a surround video affects the user experience. The interactive drama *A Perfect Party* was specifically produced for this study. The video was presented in an experimental version and a control version during a special Oculus Rift-event that was held in January 2015 in Amsterdam. All the visitors who came to user the application were invited to fill in a questionnaire in which the concept enjoyment and seven other dimensions were measured. The application itself will be discussed in further detail in the method section of this paper.

2   Backgrounds

To allow for interaction in a surround video, small elements of code can be implemented in the application that allow the viewer to have some degree of agency over the events that happen on the screen. By looking towards one of these so-called hot-spots (which could be compared to clicking on a link in an HTML page), the user can activate the next part of the scene. In the field of Interactive Storytelling, research has often focused on the freedom that may be given to users, by letting them choose alternative outcomes for the narrative [4]. This kind of "branched" or "multiple pathway" story structure has for a long time been extremely prevalent within the multimedia industry, academia, and popular culture at large [5]. Murtaugh cites Genette [6] in his critique of such narrative structures, stating that "branch point" interactions are disruptive not only because they stop the story to wait for user input, but they also place the consequence of an action performed by the viewer within the diegesis. According to [5], interactivity should serve the narrative by focusing the viewer's attention on the story rather than on the discourse (to use Genette's terminology), or on the storyworld rather than on the medium.

Another more practical problem of branched story structures has been noted by [4], namely that narratives like that must contain vast amounts of story content. While for developers who create stories in VR, this may be an interesting challenge towards the field of automated story generation, for producers of live video content, it inevitably becomes a massive problem after already two or three branch points.

A second option for an interactive story structure is what has been referred to as 'string of pearls' [7]. This is essentially a linear structure, made up of a string of small episodes, in which the user may need to perform a specific action that will allow him to

advance to the next section of the story. With this kind of structure, users follow only one possible route through the storyline, with just one final outcome. Here, the story is no longer 'disrupted' by giving viewers authority over the outcome, but they still have impact on the temporality of the narrative.

When users move through an interactive story, there is a 'risk' that they will not even notice the fact that they had an impact on the events that happened on the screen - or in case of surround video, in the video scene that is displayed around him. When the technological part of the experience runs smoothly, users may not be aware of the fact that by gazing at a certain object, they started the next part of the scene, and might, in the case of a branched story structure, not even realize that alternative endings would have been possible. This signifies, in fact, an ideal kind of interactive experience as discussed by [5], as the users' attention is completely on the product, and not on the process of narrating. However, experiments with fake interactive stories have shown that when users feel that they have some kind of agency, they enjoy this experience, whether the agency is real or not [8]. So while it can be theoretically argued, on one side, that interactivity is 'disruptive' and might lower the enjoyment of an application, awareness of the autonomy, on the other side, might be an important factor in the overall enjoyment as well. In the current study, we therefore aimed to find out if users of interactive stories enjoy it more when they are given some feedback on interactive moments, or whether they prefer to move through the narrative without distraction. To establish this difference, we have measured the user experience of two groups of users, each group watching a different version of the same interactive movie – one in which some auditory feedback is given on successful interactive moments, and one in which this is not the case.

2.1   Measuring the User Experience

To generate empirical and conceptual knowledge about the specific ways that users evaluate interactive stories, a broad range of entertainment-related measures has been proposed. Roth et al. [9] conducted a theoretical analysis which resulted in a dimensional framework of the most important and most likely dimensions of the user experience in IS. For our current study, in which we wanted to focus on the effect that awareness of interactivity might have on the user experience while engaging in a surround video, we have selected the following list of eight measures from this framework:

Presence is the first concept that needs to be explored, when studying the reception of surround video. Presence has been a popular focus of research since the 1990 s with the development of the first virtual reality technologies [10]. The term is connected to immersion and relates to a sense of 'being in the story world'. This is exactly what happens when engaging in a 360° video experience, as the story world is literally displayed all around the user. The second concept that is often measured in IS studies is *effectance*. This measure considers whether the user is able to recognize when and how he has causally affected the story world, a notion that is directly related to our current research question. *Autonomy*, which refers to the freedom to determine one's own actions and behavior, is the third measure we have included in our study. *Flow* is the

fourth concept, and a term that is often discussed in relation to media enjoyment. Flow is described as the process of optimal experience, "the state in which individuals become so absorbed in an activity that nothing else seems to matter" [11]. *Enjoyment* itself was measured as well, as fifth concept on this list [12]. Verdugo et al. [13] have stated that one of the enormous challenges that creators of interactive narratives face is that both the aesthetic potential and the technical aspects must be considered, in order to come up with a media product that will effectively catch the audience's interest. In our current study, we therefore also included *system usability* and *user satisfaction* as our fifth and sixth concepts, investigating whether the technical aspects of the system were unobtrusive enough to allow the user to enjoy the aesthetic aspects of the experience, and feel satisfied about it. Finally, the eighth concept we wanted to investigate is *identification*, for instance to verify whether men and women would connect differently with the male character position and probably therefore also experience the interactivity in a different way.

From previous studies, we are able to formulate a tentative hypothesis. Effectance and presence are important for the enjoyment of interactive media, and people appreciate to hear or feel when they had an impact on the narrative. An agency study showed that feedback sounds increase the users' perception of influencing the course of a story [14]. We might therefore expect that the condition with auditory feedback will increase perception of having influence on the story (effectance) and will therefore be more enjoyed than the version of our Oculus Rift movie in which we expect no or limited awareness of interaction.

3   Method

In this section we will discuss the movie *A Perfect Party* and proceed with the design of the questionnaire and descriptions of the experimental setting and the group of participants that took part in our study.

3.1   Material

As stimulus we used the surround video *A Perfect Party*, produced by students of the MediaLab Amsterdam of the Amsterdam University of Applied Science. This short interactive drama can be experienced using a virtual reality display such as Oculus Rift. It was filmed using a special camera system, developed by the Dutch company WeMakeVR. This camera consists of a particular construction of lenses, supported by specialized software, which allows for recording live action 360° video footage. The movie was part of a research project for which a group of five students explored the narrative, production and design challenges that they would encounter while creating a video for Oculus Rift [15].

The movie narrative itself takes place at a party, which has just started when the user enters the experience. In the role of the main character, the user is made clear that he/she is hosting a get-together for the best friend who plans to propose to his girlfriend that night. The success of the party and therefore the proposal depends on user behavior. Interaction is established by gazing at the various hot-spots that are

implemented within the scene. The first interactive moment, for which the user is asked by one of the characters to locate her missing violin, serves as a tutorial to make the user comfortable with the system. Users are also suggested choices by a voice-over, which speaks to them in a second person, as if to verbalize one's own inner thoughts. The voice-over might, for instance, state that you would like to hear some music and therefore look at the DJ, whereupon the DJ steps up to the turntable and puts on a record. This then causes the other guests in the room to start dancing. Later in the movie, by making a choice whether to focus on the entrance door or not, the viewer can either evoke scene A in which an uninvited guest enters the room and starts a fight, or scene B in which the door will be locked by a friend and the fight does not happen. Other interactions include finding objects in time, as for instance, the engagement ring. In total, *A Perfect Party* features 5 interaction points, in a story structure that is partly branched, and partly linear. The narrative leads to four different outcomes and final scenes. In all of them, the prospective groom gets to propose to his girlfriend, but dependent on the previous user choices, the other guests have fled the room because of a smoke alarm that has gone off, or the proposal has to be conducted without the engagement ring.

An experimental version of the application was developed, featuring auditory feedback, a guitar riff, for successful interactions. The guitar riff sound was chosen to fit the scenery, which shows one party guest with a guitar, sitting on one side of the party room.

3.2 Design

Our experiment used a 2×8 between-subjects design to compare eight user experiences of interactive surround video between two conditions. For the experimental group, *auditory feedback* was implemented as a feedback for successful interactions. This independent variable was not present for the control group, which experienced the same interactions without auditory feedback. The dependent variables were eight user experience dimensions taken from the short-scale version of the measurement battery for Interactive Storytelling evaluation [9]: enjoyment (2 items); autonomy, presence, identification, system usability (3 items each); flow, effectance, and user satisfaction (with 5 items each). The questionnaire asked participants to rate how much they agree with statements regarding their user experience (e.g. "The experience was entertaining."). All dimensions were measured using a 5-point-Likert scale ranging from 1 (strongly disagree) over 3 (neutral) to 5 (strongly agree). Apart from the items that measured the user experience, we also inquired about age, gender and previous experience with video games and VR applications, in order to investigate whether any of these would significantly influence the overall experience as well.

3.3 Procedure

Upon entering the research lab, participants were randomly assigned to one of the two conditions and then waited for their turn, since only two (identical) VR systems were available. Then, participants were briefly instructed and helped into the head-mounted

Oculus displays and headphones, before the *A Perfect Party* application was started. The experience lasted for an average of 10 to 11 min and participants experienced the application while standing to fit the posture of their character in the story, who is also standing. Afterwards, participants were helped out of the VR system and instructed to fill out our paper-and-pencil evaluation questionnaire.

3.4   Participants

A sample of 287 guests of the *A Perfect Party* presentation event voluntarily participated in our study over the course of five days. 13 participants did not fill out the second page of the paper questionnaire and were excluded from further analysis. The final sample consisted of $N = 274$ participants; 144 males and 130 females between the ages of 9 and 72, with an average age of $M = 32.9$ years ($SD = 13.5$ years). Participants had a low degree of computer game literacy ($M = 2.22$, $SD = 1.06$) and a low degree of VR experience ($M = 1.65$, $SD = .81$), both on a scale from 1(no experience) to 5 (a lot of experience). We controlled for an equal distribution of gender to experimental and control group.

4   Results

In a first step, we ensured that experimental and control group were similar. We compared mean values with independent samples T-tests and found no significant differences regarding age, gender, video game and VR experience.

Reliability of the short-scales was satisfying. Only the recoded autonomy item 3 ("I felt strong limitations to my decisions on how the story should proceed.") was removed for a better reliability of the autonomy scale. See Table 1 for the reliability of each scale.

Multiple regression analysis was used to test if user experience dimensions significantly predicted participants' ratings of enjoyment. The results of the regression indicated that three predictors explained 58 % of the variance ($R^2 = .35$, $F(3, 265) = 46.9$, p < .001). It was found that presence significantly predicted enjoyment ($\beta = .33$, $p < .001$), as did autonomy ($\beta = .27$, $p < .001$) and usability ($\beta = .25$, $p < .001$).

Within-subject comparison by means of independent samples T-tests revealed that auditory feedback mainly affected the perception of presence (see Table 1 for results). Users rated presence significantly lower when getting the guitar riff sound as feedback for their actions. As a result, we have to reject our hypothesis, that auditory feedback will increase the perception of effectance.

Overall, participants enjoyed both versions of the application and gave the applications positive ratings (over 4) regarding system usability, perceived presence and user satisfaction. Albeit non-significant, we can see a trend of users liking the version without auditory feedback slightly better. Autonomy got the lowest rating from both groups; the score of ca. 3 equals a neutral user rating, which suggests that users perceived limited interaction possibilities.

Table 1. Comparison of user ratings between experimental and control group ($N = 274$)

	Reliability	No auditory feedback		Auditory feedback		
		M	SD	M	SD	P
System usability	α = .76	4.29	.51	4.19	.54	.10
Effectance	α = .89	3.51	.83	3.54	.88	.78
Autonomy	r = .54**	2.95	.94	3.08	.89	.23
Flow	α = .75	3.59	.63	3.58	.71	.92
Presence	α = .72	4.10	.65	3.93	.69	.034*
Identification	α = .77	3.37	.80	3.25	.82	.21
User satisfaction	α = .80	3.95	.63	3.89	.58	.39
Enjoyment	r = .47**	4.29	.53	4.18	.63	.13

Note: * significant difference at $p < .05$, ** significant difference at $p < .01$. Reliabilities of scales with only two items were assessed using Pearson's r correlations.

5   Conclusion

Our study aimed to investigate whether users of an interactive surround video prefer auditory feedback when they have successfully interacted with the narrative, or if they like it better when there is no feedback. Since users enjoy the sense of autonomy that interactive media give them, we had expected the experimental version in which a short sound was played when hot-spots in the movie were activated, to score higher on the user experience measures that we had included in our questionnaire. However, it turned out that the version without feedback was enjoyed more than the version with feedback. We learned from our results that presence, autonomy and usability drive the enjoyment of the immersive interactive application *A Perfect Party*. We expected auditory feedback to increase perceived effectance on story development but found that it did not increase effectance, but reduced the perception of being in a non-mediated environment (presence) and thus might lower suspension of disbelief. In other words: the small sound effects that were implemented in the movie lowered the sense of 'being in the story world' and might have indeed, as theoretically predicted by Genette [6] been perceived as being slightly disruptive.

Another more general but meaningful conclusion may be drawn from some of the background characteristics of our group of respondents. We observed that neither age nor gender had any significant effect on the overall experience of enjoyment that the surround video generated. Also, whether a respondent had a lot of experience playing video games or using other VR applications or not, did not make any difference in their appreciation of the movie and the system it was displayed with. Basically, all people of all ages scored high on system usability, which meant that they enjoyed wearing the Oculus Rift. User satisfaction was also quite high for all groups: most people were delighted to participate in the interactive movie that was displayed around them. This means that a shift towards *lean in* entertainment media is made just as easily

by *lean back* audiences as it is for people who are familiar with *lean forward* media. For producers of interactive 360° content, this finding may point towards promising possibilities of developing successful surround video projects for all kinds of audiences within a general public, without having to feel limited to the traditional (male) video game players' preferences considering themes and settings.

6   Discussion

However, both the experimental setting and the content of the application have confronted us with some limitations, which may have had influence on the final outcome of our research project. To start with, the story structure of our application itself does not allow for much autonomy, as most users mainly do what they are suggested to do (stop the unwanted guest, find the ring), either by the characters in the video, or by the voice-over that formulates the thoughts of the main character. The overall enjoyment of the application may therefore have been limited by this perceived lack of autonomy. Roth et al. [16] have found that replaying an IS application had positive effects on effectance. We might therefore expect that the perceived effectance in the current interactive movie would have been rated higher as well, when respondents would have been invited to watch it two or more times, and then to deliberately choose different alternatives. Also, the nature of the feedback sound itself may have influenced the perception, as people may have perceived the little musical sound as annoying or inappropriate. Therefore generalization towards all kinds of sounds or even all kinds of feedback may be premature at this point.

7   Future Research

With our current study, we have shown that among viewers of all ages, participation in an interactive surround video without feedback sounds was found more enjoyable than with feedback sounds. In our future research, we therefore intend to pursue our investigations into the ways that viewers can be guided through the narrative without disturbing their sense of presence. It will also be useful to investigate the effect of other kinds of feedback on interactive points, such as visual cues or different sound effects. As [17] suggested, the sense of participation in a scene may be just as important as the actual power to influence the outcome. In an ideal setting, a viewer would be given the possibility to linger in the scene and look around at people and objects that would attract his attention, while at the same time being gently invited by elements in the same scene to move forward in the narrative. We intend to focus on 'string of pearl' kind of story structures, instead of branched stories. The agency over the temporality of a scene that 'string of pearls' allows for is a drastic shift compared to traditional cinema, but is familiar to video gaming and it seems to fit the affordances of surround video in a way that exceeds the second point of critique towards interactive narratives as stated by Genette [6].

Another focus of future investigation will be the use of voice-over, and the possibility of heightening the users' sense of presence when they are guided through a narrative by a voice that is talking from a second person perspective, as if one's own

inner thoughts are being verbalized. This aspect is also quite new to cinema but well known to many gamers. Both topics promise to evoke intriguing new challenges to writers, producers and researchers of *lean in* experiences of interactive storytelling.

References

1. Nielsen, F.: Surround video: a multihead camera approach. Vis. Comput. **21**, 92–103 (2005)
2. Vosmeer, M., Schouten, B.: Interactive cinema: engagement and interaction. In: Mitchell, A., Fernandez-Vara, C., Thue, D. (eds.) ICIDS 2014. LNCS, vol. 8832, pp. 140–147. Springer, Heidelberg (2014)
3. Katz, H.: The media handbook: A Complete Guide to Advertising Media Selection, Planning, Research, and Buying. Routledge, New York (2010)
4. Stern, A.: Embracing the combinatorial explosion: a brief prescription for interactive story R&D. In: Spierling, U., Szilas, N. (eds.) ICIDS 2008. LNCS, vol. 5334, pp. 1–5. Springer, Heidelberg (2008)
5. Murtaugh, M.: The Automatist Storytelling System. Masters Thesis, MIT Media Lab (1996). http://xenia.media.mit.edu/∼murtaugh/thesis/
6. Genette, G.: Narrative Discourse. Cornell University Press, New York (1980)
7. Fitzpatrick, M.: Story Structure. Patterns in Narrative Structure for Interactive Media (2003). http://staff.jccc.net/mfitzpat/storystructures/index.htm
8. Fendt, M.W., Harrison, B., Ware, S.G., Cardona-Rivera, R.E., Roberts, D.L.: Achieving the illusion of agency. In: Oyarzun, D., Peinado, F., Young, R., Elizalde, A., Méndez, G. (eds.) ICIDS 2012. LNCS, vol. 7648, pp. 114–125. Springer, Heidelberg (2012)
9. Roth, C., Klimmt, C., Vermeulen, I.E., Vorderer, P.: The experience of interactive storytelling: comparing "Fahrenheit" with "Façade". In: Anacleto, J.C., Fels, S., Graham, N., Kapralos, B., Saif El-Nasr, M., Stanley, K. (eds.) ICEC 2011. LNCS, vol. 6972, pp. 13–21. Springer, Heidelberg (2011)
10. Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T.: Measuring and defining the experience of immersion in games. Int. J. Hum Comput Stud. **66**(9), 641–661 (2008)
11. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper and Row, New York (1990)
12. Vorderer, P., Klimmt, C., Ritterfeld, U.: Enjoyment: at the heart of media entertainment. Commun. Theory **14**(4), 388–408 (2004)
13. Verdugo, R., Nussbaum, M., Corro, P., Nunnez, P., Navarrete, P.: Interactive films and coconstruction. ACM Trans. Multimedia Comput. Commun. Appl. **7 (4)**, 39:1–39:24 (2011)
14. Roth, C., Vermeulen, I.: Real story interaction: the role of global agency in interactive storytelling. In: Herrlich, M., Malaka, R., Masuch, M. (eds.) ICEC 2012. LNCS, vol. 7522, pp. 425–428. Springer, Heidelberg (2012)
15. Medialab Amsterdam (2015). http://medialab.hva.nl/interactivecinema/
16. Roth, C., Vermeulen, I., Vorderer, P., Klimmt, C.: Exploring replay value: shifts and continuities in user experiences between First and Second Exposure to an Interactive Story. Cyberpsychology, Behav. Soc. Netw. **15**(7), 378–381 (2012)
17. Tanenbaum, J.: Being in the story: readerly pleasure, acting theory, and performing a role. In: Si, M., Thue, D., André, E., Lester, J.C., Tanenbaum, J., Zammitto, V. (eds.) ICIDS 2011. LNCS, vol. 7069, pp. 55–66. Springer, Heidelberg (2011)