Effect of Dramatic Elements on Engagement in an Augmented Reality Experience for a Museum

André T. J. Kristensen, Casper Høgalmen, Leon Müller, Martin Kraus(✉), and Mohanad M. Zeitoun

Aalborg University, 9000 Aalborg, Denmark
martin@create.aau.dk

Abstract. The long-term goal of this work is to improve the experience of museum visitors with the help of Augmented Reality (AR) experiences for informing the visitors about the displayed exhibits through other means than museum labels, as these either lacked information or were placed inconveniently at the exhibitions. This paper proposes the integration of Brenda Laurel's concept of Dramatic Interaction by utilizing different non-playable characters (NPC) that serve as the users' companions during their visits to various exhibits. This approach aims to enhance the user's level of engagement which would positively influence their overall experience at the museum. The proposed solution was evaluated by comparing two versions of the developed system, one makes use of different unique NPCs in its design, while the other instead uses a single narrator. A statistical test was performed on the collected data and the results indicated that all but one aspect of engagement did not appear significantly different between the two evaluated conditions. Telling the story of artifacts through different NPCs seemed to provide a more fun experience to users, when compared to the stories being told by one narrator.

Keywords: Dramatic Interaction · Non-playable characters · Augmented Reality · Museum

1 Introduction

For more than 60 years, museums and cultural institutions have been utilizing handheld electronic technologies [30], ranging from audio or multimedia guides that usually provide additional information about exhibits [2], to digital edutainment games which can be used for children to play and have fun while learning [17]. While a museum setting can attract demographics of a wide age range, one key factor for visiting a museum is seeking new knowledge, whether it is in the field of history, culture, science, technology, art, etc. According to Packer and Ballantyne [24], one of the motivating factors for visiting a museum is to

A. Brooks et al. (Eds.): ArtsIT 2020, LNICST 367, pp. 100–118, 2021.
https://doi.org/10.1007/978-3-030-73426-8_6

expand a person's knowledge and acquire new information on various subjects. Thus, museums present an opportunity for learning which can ultimately lead to a satisfactory experience.

Lately, the aforementioned institutions have also started to incorporate the emerging technologies of Augmented Reality (AR) and Virtual Reality (VR), as these allow visitors to experience exhibitions through new ways of engagement and interaction [13]. Additionally, some museums have adapted these technologies to use virtual characters that communicate stories and information about exhibits, as well as guiding them to other exhibits that might be of interest [3,15,31]. These characters usually take the role of a traditional human guide, helping museums convey the exhibit information directly to each visitor in a more interesting way, compared to just reading about them via stands or labels placed at the exhibits. This is further enhanced by the ability to overlay supporting visuals onto the exhibits through AR.

In collaboration with Vesthimmerlands Museum, six exhibits were chosen as the ones the AR application should focus on, as the curators felt that these were lacking additional information that was not provided by the appertaining labels displayed next to them. Furthermore, some of the exhibits contained interesting characteristics or visual peculiarities that the AR could aid in identifying for visitors, considering they might not notice these when they are examining the exhibits unguided.

With the specific exhibits and basic functionality of the AR features established, we decided to further support this by focusing on achieving an engaging experience through the application. This was deemed relevant, since research suggests that designing for engaging interaction encourages and facilitates learning as well as enhances the user experience [12,32,33]. Additionally, incorporating a form a narrative into an application that attempts to facilitate learning or information, has proved to reduce the cognitive load involved when introduced to a mass of information [10,22] and offer cognitive and imaginative engagement [4,22].

In this paper, we present the results of a comparative study of user engagement where two versions of an AR application, both designed for the context of Vesthimmerlands Museum, are evaluated through a within-subjects experimental design. One version utilizes elements from the concept of Dramatic Interaction (DI), whereas the other version more closely resembles an interactive audio guide. The following section reviews work relating to the current paper and provides an overview of what entails DI. Section 3 describes the materials and methods used in the evaluation. Section 4 presents the results and a discussion of the findings, and Section 5 concludes the paper.

2 Related Work

To provide a better overview of the topics covered in this section, it was divided into three subsections. The first of these investigates general AR applications deployed in various museum contexts, the second delves deeper into the concept of DI and how it relates to engagement, and the third follows up on this by

providing examples of studies that have utilized certain elements relating to the concept.

2.1 AR in Museum Contexts

There are numerous recent examples of AR applications in museums that contribute to the visitor's overall experience.

One study, carried out by Lando, E. [16], compared two different types of AR visualization systems on the learning experience in a museum. These two types were referred to as On-screen space and In-world space. The former renders the perceived virtual content from an exhibition on the screen of the mobile device, while the latter renders the virtual content directly onto the physical exhibition. Through an empirical evaluation, it was discovered that the In-world space system resulted in a more enjoyable and engaging experience when exploring the different exhibitions at the museum. Similar results were concluded by both Miyashita et al. [23], who carried out a study on AR guidance in the Louvre Museum in Paris, and Leue et al.[20], who conducted their research at Manchester Art Gallery. Lando notes that these aspects can positively influence learning, however, a larger experiment with more participants would be required to obtain more conclusive results, seeing that the experiment was only conducted with 12 participants. However, the research by Leue et al. reached the same conclusion that creating an enjoyable and engaging AR application can contribute to the user's learning outcome from a museum visit.

Chang et al. [6] developed an AR auxiliary tool for painting appreciation, and the learning performance of three groups of participants was explored. The three groups were divided into people exploring a museum with the proposed AR solution, people carrying an audio-guide, and non-guided people. When measuring the learning performance of each group, Chang et al. focused on factors such as learning effectiveness, the amount of time spent focusing on the painting, behavioral patterns, and attitude of using the guide system. Results indicated that each of the aforementioned categories showed an alleviated level for the group equipped with the AR guide, compared to the audio-guided group and the unguided group. To further underline the results of the study, Chang et al. emphasize that both teachers and students felt that AR not only promoted participation and motivation, but also created a realistic and novel environment when the real world is combined with the virtual world.

However, even though the coupling between the virtual and the physical scenes is a key element of a mobile AR-guide [14], Sparacino [28] emphasizes that such a system might cause visitors to place too much attention on the information in the guide device, thus lowering the appreciation of the physical artifact at hand. It is therefore important to find the right balance between the virtual guide information and the physical artwork when creating an AR system, to retain the advantages of a system that brings both realms together within the user's range of vision.

This is also emphasized by Marques and Costello [21] in their paper, where they investigated the different concerns and challenges that are prone to appear

when AR applications are placed in the context of museums. They also argue that for AR to not just be adopted by museums because it is an eye-catching technology, it "needs to be a solution to the visitor experience by effectively weaving the virtual with the physical into the narrative, and ensuring that the interface becomes an integral layer, a storytelling tool.". Something worth considering when trying to accomplish this, is the fact that the exhibits themselves should be used to trigger the AR features rather than location based AR, as this helps avoid visitors becoming detracted from the physical museum. However, that method suffers from its own issues such as lighting conditions and line of sight. Usually proper AR activation requires higher light conditions, and some museums tend to have sections that are dimly lit due to conservation concerns. Furthermore, if a museum has peak hours with heavy crowds, there is a higher chance of a visitor losing line of sight to the exhibit with their camera, due to people being forced to walk in front of exhibits. Another unfavorable side effect of crowded spaces is the noise level. Traditional audio guides can be held against the ear, but visitors standing at objects offering augmentation, hold the device in front of their bodies to experience the visuals. This typically results in the sound being more lost to the surroundings.

These are all important matters to consider when designing an AR application to be used in the context of a museum. Sung, Chang, Hou, and Chen [29] additionally point out that a mobile guide should aim to incorporate the context of the promoted environment as much as possible, thereby including the guiding environment, their companions, the exhibits and their cultural and social implications. The interaction between the visitors and these aforementioned aspects should be fully supported when designing a mobile guide to form what Chang et al. calls "human-computer-context" . Again, the planning of the guide system is crucial to not over-exaggerate the digital part of the system, i.e. the "human-computer" guide system, whilst not neglecting the "human-situation" , i.e. the real environment at hand. Failing to balance these two factors will result in a system that is incapable of inducing a satisfying human-computer-context.

2.2 Dramatic Interaction

A concept that has been utilized to achieve a higher level of engagement is Dramatic Interaction, a term introduced by Brenda Laurel [18]. It is the concept of interactors creating their own narratives using a given system to create their own stories with personal significance. Examples of freedom within boundaries include, but are not limited to, allowing interactors to choose the order of interactions or allowing them to choose their own goals within the system. Based on the idea that things humans interact with naturally become more important to them, Laurel analyzes the relationship between player and designer as a collaborative effort towards creating an enjoyable experience. Furthermore, Laurel hypothesizes that experiences tend to spark fascination and engagement when they appear to transcend artificial structure. She therefore points to theatrics as an art form dedicated to the illusion of reality, suggesting that the use of

theatrical techniques in interaction design, such as characters and dramaturgical models, can increase engagement.

These characters are usually implemented as non-playable characters (NPCs) which are defined as characters in the virtual worlds not controlled by a user. While most commonly used in video games, other applications have historically made use of NPCs for a variety of purposes. Applications including narratives often utilize NPCs as plot devices, having them help further a storyline. They can also be applied to aid the user either through helpful information or other means depending on the application type. Lastly, NPCs can be made into personified game functions such as user-interfaces or save points.

2.3 Non-playable Characters in Augmented and Virtual Reality

While the studies described in Section 2.1 indicate learning and enjoyment benefits from introducing AR in a museum context, the results from a study conducted by Jessel et al. [11] did not show a significant improvement in terms of learning. The study focused on enhancing the visitor's learning experience in museums by developing a handheld AR application called M.A.R.T.S (Mobile Augmented Reality Touring System). They compared M.A.R.T.S to two traditional and widely used museum systems: labels and audio guides. In their study, the authors implemented a virtual human guide to aid with conveying information about the works of art displayed at the Bayonne's Basque Museum. The virtual guide was utilized to provide instructions to the visitors regarding which areas of an exhibit can be explored to acquire further knowledge about it by pointing at the respective area. This approach, which was referred to as "Selection" , limited potential confusion that could be caused due to lack of obvious visual reference indicating the works of art or areas of an exhibit that are augmented by the application. The results of the study revealed that M.A.R.T.S did not perform significantly better than the labels and audio-guide in terms of learning more about the exhibits, however, on average participants scored higher using M.A.R.T.S when they had to answer questions about the works of art they had explored. Moreover, presenting information in both textual and auditory manners seemed to be redundant with participants favoring audio only. As a positive note, the "Selection" process proved advantageous when trying to identify parts of a work of art that is simultaneously referenced by the virtual guide. This means that this method can be utilized in AR to direct the visitor's attention to a particular area of an exhibit in a museum.

In 1994, Laurel, Strickland and Tow [19] attempted to utilize VR for the sake of entertainment, something that was quite a novel idea at the time, by conducting a research project they named "Placeholder". They explored new ways of narration in VR, where the environments were inspired by real life locations and users could take the form of four animated spirit critters that inhabited these environments. Doing so, would allow the user to experience the specific critter's "unique visual perception, its way of moving about, and its voice" . To enrich the DI within the virtual world, a character named "The Goddess" was added. The special trait of this character was that its personality was improvised live,

usually by either Laurel herself or one of her coworkers, though occasionally, other actors would perform the role as well, including men. This resulted in the character's personality changing depending on the actor portraying it but also in relation to the participants encountering it. If the participants were children, the character would behave as a friendly helper, if they were couples it would tease them, and if the participants were acting insulting or unfriendly, so would the character. By doing this, the character would react to interactions with it, but it would also react to how it was interacted with. This is an example of how interactions with a character can be dramatized through live performance from actors that theatrically react to the specific participants that encounter it. This is not feasible in the context of the museum application; however, a scripted "performance" of an NPC could be based on such performances.

Lastly, Christopoulos et al. [7] and their paper on their VR application, 'Battle of Thermopylae', should be mentioned, as they incorporate some ideas from DI in their interactive guide approach. Based on studies confirming the importance of an actual human guide, they created a game wherein the player would walk around in historical battle camps and ask residents three questions about their relations to the world around them. According to Slater [27], this bi-directional flow of information strengthens the users' suspension of disbelief and allows them to engage more vividly with the characters. To make sure that users went through this bi-directional flow, players would have to cooperate with the characters in order to progress further into the application. With this, the group aimed for triggering Kolb's experiential learning cycle [25], wherein players would have to observe as well as directly participate in the interactive experience. During their evaluation, the research group focused on engagement and learning outcomes and found that the children evaluated had a high accuracy on the post-experience test with which they were presented. On the other hand, surveys from VR guided tours without interactive elements, showed that few individuals could answer correctly in post-experience tests. The research group concludes that interactive elements, combined with new practices in the area of games, could help facilitate learning and produce more effective learning environments.

3 Materials and Methods

One long-term aim of this study was to develop a digital platform for Vesthimmerlands Museum which they could utilize to further inform the visitors about the displayed exhibits. An AR medium offers a more engaging experience than reading text from exhibit labels which, in the case of Vesthimmerlands Museum, also suffered from a lack of information or inconvenient placements. In order to potentially further enhance the visitor experience, we wanted to measure the effects of Laurel's concept of DI on users' level of engagement.

3.1 Apparatus

The evaluated system is an AR application developed in Unity intended to detect image markers and provide information to the user in response to their inputs.

The specific exhibits that the museum saw benefit from being included in the app can be seen in Figure 1. For easier readability, the two most discussed of these exhibits, namely "Gundestrupkarret"and "Ryttergraven fra Næsby", will be referred to as the cauldron and the sword, respectively. Due to the COVID-19 pandemic, the system was not tested in the museum, but an altered version of the system was created which included 3D models of the two exhibits. This version allowed users to scan the image targets and have interactions as per the original intent.

Fig. 1. From left to right: Aars-egnens Første Bank, Petreas Skilt, Hedegårdmanden, Tatoveringsnåle fra Bronzealderen, Ryttergraven fra Næsby, Gundestrupkarret.

In order to evaluate the effects of dramatic elements on a user's experience and their level of engagement, it was necessary to create a version of the application that was devoid of any aspects of DI. Other key features of interaction were present in both variants. This allowed us to discern the difference between the two experiences by isolating that one key distinction and comparing them directly. The first variant was reminiscent of standard audio guides with a single omniscient second-person narrator where users could choose to get information about specific artifact attributes (see Figure 2). The second variant on the other hand, included a character for each supported exhibit which inferred their personalities into the interactions users would have with the system (see Figure 3).

While Laurel describes a multitude of different methods for dramatizing interaction, we chose to focus on the mix of mediated collaboration, NPCs, and a variety of theatrical writing methods. This would link to the aim of the museum of providing information about the artifacts by having NPCs historically connected to them tell stories about their origins and uses. Visitors would be able to choose questions of relevance to their own interests, thus creating mediated collaboration. Furthermore, with inspiration taken from "The Goddess" of the "Placeholder" project mentioned in Section 2.3, the characters would react to users interacting with the play, pause, or skip button. Each character would have three pre-recorded voice lines for each interaction which would be played at random when the corresponding button was pressed. Unfortunately, regarding the audio for the different characters, with no budget for hiring a professional voice actor, all the voice lines were recorded by two of the authors of this paper who were both amateurs in this field. This undoubtedly affected the quality, however, with research suggesting that audio is preferred over text, and the fact

Fig. 2. Example of exhibit interaction from the application with no dramatic interaction elements

Fig. 3. Example of exhibit interaction from the application with dramatic interaction elements

that personifying a character through text, rather than audio, is more difficult, we decided to proceed with the decision.

These NPCs should also help provide some of the context incorporation that Sung, Chang, Hou, and Chen (see Section 2.1) argue for, due to the decision of making them historically connected to each specific exhibit. The interface of

the application was also designed with the context of a museum in mind, while simultaneously attempting to minimize distractions and visual overload [8] by hiding most of it behind a single expandable button (see top right of Figure 2 or 3). When pressed, a panel will expand, revealing six buttons containing icons that depict each exhibit (see top right of Figure 4). These icons are semi-transparent and will remain so until an exhibit has been visited, resulting in the respective icon turning opaque (see third icon from the left in Figure 4). When any of these six buttons is pressed, a journal will show on screen that functions both as a form of guidance system, by displaying images of the exhibits supported by the application, and as a place for users to review their interactions with exhibits they have visited (see Figure 4).

Fig. 4. Example of a page in the application journal

The augmentation provided by the application was implemented in the form of highlights positioned at points of interest on the exhibit (see Figure 2 or Figure 3). These would aid the users in examining interesting or specific regions of the physical exhibit, while being informed about them by the appertaining character. Furthermore, this should prevent the users from focusing more on their device rather than the exhibit, which was emphasized by multiple authors referenced in Section 2.1.

When starting the application, a user would first be greeted with an introductory message containing instructions on how to trigger interactions. They would also be told of the journal and its functionality. Once the user had gone through the introduction, they would have to scan one of the provided image targets to activate an interaction. Here, users would be presented with three main questions each with two sub-questions, the answers to all of which, were

narrated. If any of these questions contained a piece of dialogue about a specific point of interest on the exhibit, the narrator or character would instruct the user to examine the exhibit with their phone, as highlights that point towards these would be visible. Once all supported exhibits had been interacted with, a final monologue was triggered informing the user that the experience was over.

3.2 Procedure

The experiment with the altered version was conducted as a within-subject design where a participant would experience the application with and without DI. To avoid various order effects, we created four different procedures and opted for an equal number of participants going through each procedure, as a participant was instructed to follow only one of them. Two applications were created for the version containing elements from DI and likewise for the version without. These applications differed in the order of presented exhibits, as one would start with the cauldron followed by the sword and vice versa for the other. Further included in these procedures were two engagement questionnaires (see Section 3.3 Table 2), a comparative questionnaire (see Section 3.3 Table 3) and a consent form. Table 1 shows condensed examples of the four procedures, slightly altered to allow for easier readability.

Due to the restricting circumstances during the COVID-19 pandemic, testing at Vesthimmerlands Museum or setting up a dedicated space for participants to partake in the evaluation was not possible. Instead, folders containing Android Package Kits (APK) and images of the two exhibits to be used as image targets were created. Four folders were created in total, as each of them would contain different APKs depending on which category the corresponding participant was assigned (A, B, C, or D). A participant would receive an instructions file containing a link to one of these folders, a guide on how to prepare their Android phone for downloading and installing the APKs, one of the four procedures seen in Table 1, and links to the digital consent form and questionnaires.

A total of 16 participants (10 males and 6 females), aged between 22 and 32 years old, with different educational and occupational backgrounds were recruited online from various websites and platforms such as Reddit, Facebook, Microsoft Teams, and Discord. Unfortunately, it proved difficult to find participants within the limited time frame this study was carried out under due to requiring an Android phone and an advanced understanding of spoken Danish. Furthermore, the technical complexity of succeeding in installing and running the APK resulted in younger and older age groups being excluded from the evaluation due to a lack of tech familiarity. The implications of this decision are discussed in Section 4.

Table 1. Examples of participant procedures.

Participant A	Participant B	Participant C	Participant D
1. Fill out the consent form	1. Fill out the consent form	1. Fill out the consent form	1. Fill out the consent form
2. Start the application with DI containing the order of cauldron then sword	2. Start the application without DI containing the order of cauldron then sword	2. Start the application with DI containing the order of sword then cauldron	2. Start the application without DI containing the order of sword then cauldron
3. Once finished with both exhibits, answer the questionnaire (Table 2)	3. Once finished with both exhibits, answer the questionnaire (Table 2)	3. Once finished with both exhibits, answer the questionnaire (Table 2)	3. Once finished with both exhibits, answer the questionnaire (Table 2)
4. Start the application without DI containing the order of cauldron then sword	4. Start the application with DI containing the order of cauldron then sword	4. Start the application without DI containing the order of sword then cauldron	4. Start the application with DI containing the order of sword then cauldron
5. Once finished with both exhibits, answer the questionnaire (Table 2)	5. Once finished with both exhibits, answer the questionnaire (Table 2)	5. Once finished with both exhibits, answer the questionnaire (Table 2)	5. Once finished with both exhibits, answer the questionnaire (Table 2)
6. Answer the questionnaire (Table 3)	6. Answer the questionnaire (Table 3)	6. Answer the questionnaire (Table 3)	6. Answer the questionnaire (Table 3)

3.3 Hypothesis and Data Collection

Despite the fact that the developed system could not be evaluated in its intended context at Vesthimmerlands Museum, we deemed it possible to obtain similar results on the aspects of users' engagement through an evaluation of the substitute prototype.

The overall expectation of the experimental design was that the evaluated condition during which the system makes use of DI elements would achieve higher levels of engagement compared to the other condition which lacks DI. Thus, the following hypothesis was defined:

- ***H1:*** Introducing interactive non-playable characters in an AR museum application increases the user's feeling of engagement.

The effects of both conditions on the users' level of engagement were determined via a questionnaire that was adapted from a selection of well-established engagement questionnaires [1,5,9,26]. The questionnaire, seen in Table 2, consisted of

a total of 13 questions, which were selected to measure aspects that constitute the perception of engagement or have an influence on it. These aspects form the following constructs: authenticity, curiosity, focused immersion, enjoyment and interactivity. Furthermore, a few of the questions aimed to assess the concept of narrative engagement due to its importance in the context of DI. The participants could answer the questionnaire through a 5-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree".

Table 2. Engagement questionnaire

Number	Question
1	I did not find it artificial to receive information about the exhibits through the application. (You can compare it to having a human guide with you through a museum.) [9]
2	Discovering the history of the exhibits through the application gave me an authentic feeling. [9]
3	My experience with the application aroused my interest in the history of the exhibits. [1]
4	While using the application I was absorbed in what I was doing. [1]
5	While using the application I was not able to block out most distractions. [1]
6	The actual process of using the application was unpleasant. [1]
7	I find using the application to be enjoyable. [1]
8	I had fun using the application. [1]
9	Using the application provided me with an interactive experience. [26]
10	I felt I had control over my interaction with the application. [26]
11	I had a hard time recognizing the thread of the story. [5]
12	I felt like I was part of the story. [5]
13	I listened to what was said. [5]

In addition to the engagement questionnaire, a comparative questionnaire, seen in Table 3, was set up to be answered at the end of the experiment. This decision was made for the purpose of gaining more insight into the participants' favored experience, since at that point they would have tested both versions of the application, hence enabling them to make a direct comparison. A preferred procedure would have been that while the participants are answering their second questionnaire, they can go back and review their ratings from the first questionnaire, with the option of changing the ratings if they see fit. Due to the fact that the evaluation was conducted at the participants' households without the presence of a facilitator to oversee the entire process, the method of letting the participants change their ratings by themselves was deemed to be a cumbersome process that they would have to keep track of. Therefore, considering

the circumstances, the comparative questionnaire was regarded as a suitable alternative to shed some light on which system version was preferred.

The comparative questionnaire had the first three questions asking the participants to choose on a Likert scale from 1 to 5, which of the two versions they perceived as more engaging, more fun and easier to use. On the scale, 1 refers to the version that does not employ the different NPCs, while 5 refers to the one that does. The last part of the comparative questionnaire had the participants explicitly state their overall preferred version during the evaluation, followed by a question encouraging them to explain the reasoning behind their choice.

Table 3. Comparative questionnaire

Number	Question
1	Which application did you find more engaging?
2	Which application provided a more fun experience?
3	Which application was easier to use?
4	Which version of the application did you prefer?
5	What is the reason for preferring one version over the other?

Questions 1, 2, and 4 in the comparative questionnaire shared the same expectation of NPCs contributing to an experience that is significantly more engaging and more fun, thereby leading participants to favoring that version of the application over the one that is guided by one narrator. As for question 3, considering that none of the functionalities that facilitate the usage of the application are any different in the two versions of the application, it was expected that there would be no perceived difference in terms of ease of use. This led to the formulation of the following hypothesis:

- ***H2:*** The version of the application using non-playable characters is more engaging and fun, and overall preferred, when directly compared to the version using one narrator.

3.4 Statistical Analysis

All the questions in the engagement questionnaire shared the same expected outcome as the one mentioned in Section 3.3 for the complete experience, thereby leading to the decision of treating each question as if it were an individual experiment performed independently from the rest. This is advantageous because it eliminates the necessity to carry out a correction to counteract the problem of multiple comparisons, which is a problem that has a negative impact on the reliability of the outcome. All the questions are, however, evaluated in the same questionnaire in order to save time on what would otherwise have been a lengthy process, which was not possible under the time constraints this study was carried out under.

The experimental design was a repeated measures design due the measurements of a user's level of engagement being collected under two conditions for each participant. The statistical data is ordinal, thus making it suitable to perform a non-parametric test. Moreover, the comparison process between the two conditions of the experiment is a pairwise comparison, which fits the methodology of the Wilcoxon-signed rank test, as it evaluates the significance difference between the dependent samples. The statistical analysis of the collected data was performed with a significance level of 5% ($\alpha = 0.05$), and it was done in the IBM SPSS software.

The data of each of the first three questions in the comparative questionnaire was not a pairwise measurement between two dependent variables. Therefore a set of dummy data was introduced in place of the second variable to allow for a computation of differences, thereby making it eligible for a Wilcoxon-signed rank test. The dummy data consisted of 16 entries of a rating of 3 for each of the three questions, as this rating can be considered a central point, hence a neutral response to the provided Likert scale.

The fourth question in the comparative questionnaire instead required a binomial distribution test to determine whether the participants are biased towards selecting the version with NPCs as their preferred version.

4 Results and Discussion

The p-values for the comparisons in H1 are presented in Table 4. The results for the comparisons in H2, inferred from the comparisons in the comparative questionnaire, are shown in Table 5.

Table 4. The p-values for the comparisons of questions 1 to 13 in the engagement questionnaire. Question 8 investigated the aspect of fun.

Hypothesis	Question Number	p-value	Null hypothesis
H1	1	0.998	Retained
H1	2	0.476	Retained
H1	3	0.122	Retained
H1	4	0.220	Retained
H1	5	0.056	Retained
H1	6	0.388	Retained
H1	7	0.095	Retained
H1	**8**	**0.015**	**Rejected**
H1	9	0.168	Retained
H1	10	0.340	Retained
H1	11	0.083	Retained
H1	12	0.073	Retained
H1	13	0.066	Retained

Table 5. The p-values for comparisons of question 1, 2 and 4 in the comparative questionnaire. Question 2 investigated which of the two versions was more fun.

Hypothesis	Question Number	p-value	Null hypothesis
H2	1	0.07	Retained
H2	**2**	**0.002**	**Rejected**
H2	4	0.227	Retained

The results from the Wilcoxon signed-rank test performed on H1 and H2 indicated a significant difference for 2 of the 16 comparisons. For question 4 in the comparative questionnaire, the results of the binomial distribution test did not yield a statistically significant difference, however a tendency favoring the version with NPCs could be observed.

When asked to directly compare the two versions of the application, participants did not find the version including dramatic elements in the form of NPCs significantly more engaging than the one including a single non-embodied narrator. However, comments provided during the last part of the comparative questionnaire indicated a slight bias towards the former version. Albeit, this trend did not consolidate itself statistically, it can be observed in the answers of question 4 in the comparative questionnaire.

Due to the remote nature of this evaluation, participants were gathered and approached via a multitude of online and social media platforms. Consequently, the motivation for participants to partake in this evaluation did not arise from a natural curiosity, which can normally be observed when present at the museum. This could have affected the general engagement with the application, as participants were not particularly interested in neither the exhibits nor what the narrator or characters had to say about them.

In addition to these findings, some participants negatively commented on the lack of system flexibility regarding the absence of an option to skip an exhibit mid-interaction. It is believed that the main reason for wanting to skip the interaction with an exhibit is due to a low level of interest or curiosity about the subject. Being unable to perform this action, while already experiencing a low level of curiosity, could lead to a further decrease in interest. It can therefore be concluded that the reported lack of system flexibility is a factor that can negatively impact the participant's level of engagement.

Overall, failing to reject the null hypothesis for 12 out of the 13 questions in the engagement questionnaire, indicates that deploying elements of DI does not increase engagement in this specific context. However, question 8 of the engagement questionnaire, which investigated the aspect of fun, did indicate a significant difference. This is further supported by the rejected null hypothesis H2 for question 2 in the comparative questionnaire, wherein participants were asked to directly compare which version provided a more fun experience. Furthermore, comments from the evaluation indicated that a variety of the deployed design elements can have a positive effect on AR applications seeking to induce

an engaging experience. Among those elements are the highlights, which can subtly prompt users to inspect or interact with an object. Also, splitting a longer narrative into smaller parts, whilst implementing UI elements that allow users to control the flow of the narration, prevents participants from having to listen to an entire narration in one sitting. Additionally, hiding unnecessary UI elements from the user's field of view accommodates for the need of creating a balance between the physical and the virtual realms, thus preventing distractions and visual overload.

The perceived ease of use in question 3 of the comparative questionnaire (see Table 3) was evaluated through a statistical analysis to determine whether the expected outcome was realized. The test showed that our expectation of the comparison was correct, and there was no difference between the two versions of the application. Hence, it can be inferred that the deployed UI renders the same performance in both use cases. Although participants reported an equal performance regarding the ease of use of both versions, the precise level for each individual version is not recorded due to the question only being included in the comparative questionnaire.

The results of the comparison in question 4 of the comparative questionnaire (see Table 3) did not report a statistical significance regarding whether participants preferred DI over the interaction with a single narrator. It is to be noted that the authors did not have the required resources nor the artistical skills to allow for the deployment of an iterative design methodology for the development of the NPCs. Considering that NPCs reflect an integral part of the system in the form of dramatic elements, it would have been preferable to evaluate each character individually, prior to implementing it in the final system, to receive feedback regarding its conceptions. The sentiment of amateur character design in both visual and auditory appearance was voiced by some participants, thus further validating the potentially positive impact of more professionally designed NPCs. Other participants commented on the spirit-like appearance of NPCs, while wishing for a more humanoid depiction.

The deployment of an iterative design methodology could have allowed us to discover design flaws that were implemented in the final version of the prototype used for the evaluation. One of these flaws was that participants commented on not knowing when they could expect changes in what was augmented onto the exhibit. As such, some participants might have missed when highlights were shown or changed, or they would perhaps continuously have pointed their phone camera towards the image targets, even though this is not necessary. A more user-friendly design would have included indicators for the inclusion of AR content next to the relevant buttons in an exhibit interaction, thereby allowing users to aim their phone at an AR marker, only at moments when relevant content is about to be displayed.

In addition to not being in the desired context, the remote evaluation did not allow for the possibility of observing the participants' interactions and the subtleties that might have emanated from their body language. These subtleties would have had the potential to reveal further insight regarding the level of

engagement. Asking participants to report these subtleties would most likely be noninformative, as participants themselves can be unaware of said actions.

Although no statistical significance can be reported, trends in the results concerning preference, and comments made by participants, facilitate the conclusion that people inherently interested in the subject conveyed by the exhibits, found the detached narrator better as facts were communicated more clearly. Participants not sharing this sentiment seemed to prefer the version including dramatic elements, as it provided a livelier approach to the potentially uninteresting subject.

Based on the conducted evaluation, it can be concluded that dramatic elements in the form of different NPCs do not necessarily evoke a greater sense of engagement when compared to a version without these elements. However, a more fun experience was reported with a statistically significant difference for the version including DI. Furthermore, data trends, though statistically insignificant, as well as comments support the claim that elements of DI have the potential to induce a greater sense of engagement for people with a low level of curiosity for the presented subject. Based on the shortcomings of the apparatus such as the inability to conduct the evaluation in its intended use-case and having a low sample size, it would be interesting to conduct a refined evaluation, which accommodates for the aforementioned flaws to conclusively report any potential enhancement of engagement in a museum context when using elements of DI.

5 Conclusion

This study was concerned with researching how engagement is affected by introducing elements of Dramatic Interaction (DI) in the form of non-playable characters (NPC) in an Augmented Reality (AR) guide application for Vesthimmerlands Museum in Aars. A within-subjects experiment was conducted, and participants were presented with two AR applications comprised of varying engaging elements. The ongoing COVID-19 pandemic, at the time of conducting the evaluation, posed the need to conduct it remotely.

Based on the deployed experimental method of this paper, it cannot be concluded that the utilized elements of DI in an AR application enhance the level of engagement. However, results indicated that a more fun inducing AR experience is created when utilizing the deployed elements of DI, when compared to the absence of these features.

The revealed increase in fun that elements of DI can have in an AR application could be considered by other AR applications such as games or edutainment software. Furthermore, while designing these applications, developers could consider some of the deployed design principles in this paper, such as the use of highlights and a visually subtle user-interface. Future work could investigate how engagement is affected by conducting the deployed evaluation of this paper in situ in its intended context, seeing that imposed restrictions prevented this. The option to improve the production quality regarding the artistical aspects of this system is worth exploring as well. Moreover, a greater sample size should

be obtained to increase the robustness and meaningfulness of the reported statistical results.

References

1. Agarwal, R., Karahanna, E.: Time flies when you're having fun: cognitive absorption and beliefs about information technology usage. MIS Q.: Manag. Inf. Syst. **24**(4), 665–694 (2000)
2. Alexandri, E., Tzanavara, A.: New technologies in the service of museum education. World Trans. Eng. Technol. Educ. **12**, 2 (2014)
3. Bickmore, T.W., Vardoulakis, L.M., Schulman, D.: Tinker: a relational agent museum guide. Autonomous Agents Multi-Agent Syst. **27**(2), , 254–276 (2013)
4. Bielenberg, D.R., Carpenter-Smith, T.: Efficacy of story in multimedia training. J. Netw. Comput. Appl. **20**(2), 151–159 (1997)
5. Busselle, R., Bilandzic, H.: Measuring Narrative Engagement. Media Psychol. **12**(4), 321–347 (2009)
6. Chang, K.E., Chang, C.T., Hou, H.T., Sung, Y.T., Chao, H.L., Lee, C.M.: Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum. Comput. Educ. **71**, 185–197 (2014)
7. Christopoulos, D., Mavridis, P., Andreadis, A., Karigiannis, J.: Using, environments, virtual, to tell the story: the battle of thermopylae. In: Proceedings of 3rd International Conference on Games and Virtual Worlds for Serious Applications. VS-Games, pp. 84–91 (2011)
8. Endsley, T.C., Sprehn, K.A., Brill, R.M., Ryan, K.J., Vincent, E.C., Martin, J.M.: Augmented reality design heuristics: designing for dynamic interactions. In: Proceedings of the Human Factors and Ergonomics Society (2017), vol. 2017-October, pp. 2100–2104. Human Factors an Ergonomics Society Inc. (2017)
9. Featherman, M. S., Valacich, J. S., and Wells, J. D. Is that authentic or artificial? Understanding consumer perceptions of risk in e-service encounters. Information Systems Journal 16, 2 (apr 2006), 107–134
10. Forrester, M.A.: Can narratology facilitate successful communication in hypermedia environments? Intell. Tutoring Media **7**(1), 11–20 (1996)
11. Ghouaiel, N., Garbaya, S., Cieutat, J.-M., Jessel, J.-P.: Mobile augmented reality in museums : towards enhancing visitor's learning experience. Int. J. Virtual Reality **17**(1), 21–31 (2017)
12. Jacques, R., Preece, J., Carey, T.: Engagement as a design concept for multimedia. Can. J. Learn. Technol./La revue canadienne de l'apprentissage et de la technologie **24**, 1 (1995)
13. Jung, T., tom Dieck, M.C., Lee, H., Chung, N.: Effects of virtual reality and augmented reality on visitor experiences in museum. In: Inversini, A., Schegg, R. (eds.) Information and Communication Technologies in Tourism 2016, pp. 621–635. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-28231-2_45
14. Klopfer, E., Squire, K.: Environmental detectives-the development of an augmented reality platform for environmental simulations. Educ. Technol. Res. Dev. **56**(2), 203–228 (2008)
15. Kopp, S., Gesellensetter, L., Krämer, N.C., Wachsmuth, I.: A conversational agent as museum guide – design and evaluation of a real-world application. In: Panayiotopoulos, T., Gratch, J., Aylett, R., Ballin, D., Olivier, P., Rist, T. (eds.)

IVA 2005. LNCS (LNAI), vol. 3661, pp. 329–343. Springer, Heidelberg (2005). https://doi.org/10.1007/11550617_28
16. Lando, E.: How augmented reality affects the learning experience in a museum. Technical report (2017)
17. Larsen, J., Svabo, C.: The tourist gaze and "Family Treasure Trails" in museums. Tourist Stud. **14**(2), 105–125 (2014)
18. Laurel, B.: Computers as Theatre. Addison-Wesley, Boston (1993)
19. Laurel, B., Strickland, R., Tow, R.: Placeholder: landscape and narrative in virtual environments. ACM Comput. Graph. **28**, 2 (1994)
20. Leue, M.C., Jung, T., Tom Dieck, D.: Google glass augmented reality: generic learning outcomes for art galleries. In: Tussyadiah, I., Inversini, A. (eds.) Information and Communication Technologies in Tourism 2015, pp. 463–476. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-14343-9_34
21. Marques, D., Costello, R.: Concerns and Challenges Developing Mobile Augmented Reality Experiences for Museum Exhibitions. Curator 61, 4 (oct 2018), 541–558
22. McLellan, H.: Hypertextual tales; story models for hypertext design. J. Educ. Multimedia Hypermedia **2**, 3 (1993)
23. Miyashita, T., et al.: An augmented reality museum guide. In: Proceedings - 7th IEEE International Symposium on Mixed and Augmented Reality 2008, ISMAR 2008, pp. 103–106 (2008)
24. Packer, J., Ballantyne, R.: Motivational factors and the visitor experience: a comparison of three sites. Curator: Museum J. **45**(3), 183–198 (2002)
25. Sims, R.R.: Kolb's experiential learning theory: a framework for assessing person-job interaction. Acad. Manag. Rev. **8**(3), 501–508 (1983)
26. Skadberg, Y.X., Kimmel, J.R.: Visitors' flow experience while browsing a Web site: its measurement, contributing factors and consequences. Comput. Hum. Behav. **20**(3), 403–422 (2004)
27. Slater, M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos. Trans. Roy. Soc. London Ser. B Biol. Sci. **364**(1535), 3549–3557 (2009)
28. Sparacino, F.: Accepted for publication in Proceedings The Museum Wearable: real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences. Technical report (2002)
29. Sung, Y.T., Chang, K.E., Hou, H.T., Chen, P.F.: Designing an electronic guidebook for learning engagement in a museum of history. Comput. Hum. Behav. **26**(1), 74–83 (2010)
30. Tallon, L., Walker, K.: Digital Technologies and the Museum Experience: Handheld Guides and Other Media. AltaMira Press (2008)
31. Traum, D., et al.: Ada and grace: direct interaction with museum visitors. In: Nakano, Y., Neff, M., Paiva, A., Walker, M. (eds.) IVA 2012. LNCS (LNAI), vol. 7502, pp. 245–251. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33197-8_25
32. Webster, J., Hackley, P.: Teaching effectiveness in technology-mediated distance learning. Acad. Manag. J. **40**(6), 1282–1309 (1997)
33. Webster, J., Ho, H.: Audience engagement in multimedia presentations. Data Base Adv. Inf. Syst. **28**(2), 63–77 (1997)