

How Real Should Virtual Characters Be?

Daniel Wagner

Graz University of Technology
Institute for Computer Graphics
and Vision, Inffeldgasse 16
Graz, 8010 Austria

Mark Billingham

University of Canterbury
HIT Lab New Zealand
Private Bag 4800
Christchurch 8004 NZ

Dieter Schmalstieg

Graz University of Technology
Institute for Computer Graphics
and Vision, Inffeldgasse 16
Graz, 8010 Austria

Contact email address: daniel@icg.tu-graz.ac.at

ABSTRACT

In recent years 3D virtual characters have become more common in desktop interfaces, particularly in gaming and entertainment applications. In this paper we describe how augmented reality (AR) technology can be used to bring virtual characters into the real world and compare AR characters to other types of virtual characters. We have developed a handheld AR educational application in which a virtual character teaches users about art history. We present results from a user study that explores how realistic the character needs to be for it to be an effective and engaging educational tool and if augmented reality offers benefits for this type of application.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems - *Animations, Artificial, Augmented, and Virtual Realities*. K.3.1 [Computing Milieux]: Computer Uses in Education - *Collaborative learning*.

General Terms

Design, Experimentation, Human Factors, Verification

Keywords

Virtual Characters, Handheld Augmented Reality, User Study

1. INTRODUCTION

One of the goals of human computer interface research is to make interaction with the computer as seamless as possible. Recently researchers have explored how virtual characters can be used to make human computer interaction more like intuitive human to human collaboration.

Although such characters are becoming common on desktop interfaces they are relatively rare in Augmented Reality (AR) applications. AR interfaces involve technology that allows real time computer graphics to be overlaid on the real world and so for the first time the virtual characters could appear to inhabit the same physical space as the user.

In this paper we describe one of the first implementations of a virtual character in a handheld AR application. We have developed a novel educational AR application in which a virtual character assists a user in learning art history (see Figure 1). Unlike previous work with AR characters, we present results from a user study that shows the effect that the character representation has on the user's engagement and performance. The key research question we are exploring is how realistic does a virtual character need to be for the user to feel engaged with it and enjoy the application, and what benefit can be derived from using AR characters. The result of this study has implications for other researchers who are developing virtual characters for AR and non-AR interfaces.

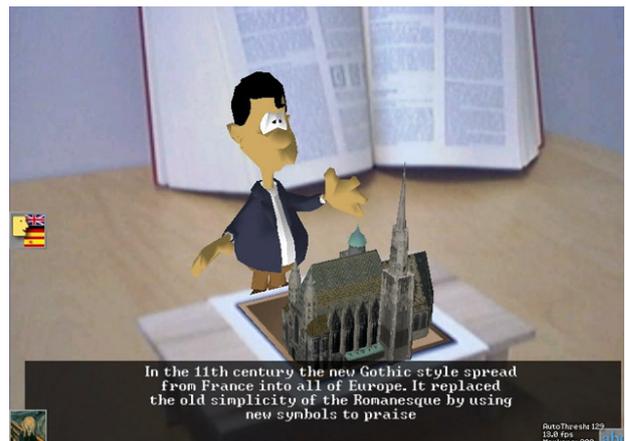


Figure 1: Mr. Virtuoso illustrating a historical art object. Character and church are virtual; book and table are real.

In the rest of the paper we first review related work in the field of virtual characters and augmented reality. Next we describe the handheld AR character we have developed and the educational application it was used in. Then we present an experiment which evaluated user's responses to different types of character. Finally, we outline some design recommendations for character interfaces and directions for future research.

2. RELATED WORK

Our work draws on a previous research in virtual characters, usability studies, AR characters and mobile augmented reality.

2.1. Virtual Characters

Probably the first 'talking head' mimicking human like interaction was DECface [22] developed at Digital Equipment Cooperation.

Although not an entire embodied virtual character, DECface was one of the first 3D virtual faces based on a realistic muscle model.

Following DECface, full bodied virtual characters began to appear, such as the screen based assistant Olga [17]. Olga is a cartoon-like anthropomorphic character that is animated in real-time for supporting information kiosks. It integrates spoken dialogue, 3d animated facial expressions, gestures, speech synthesis and a direct manipulation interface [6].

In the educational domain the INTERNET ADVISOR is an interactive, screen based learning environment in which a virtual agent Cosmo [12] helps the user to accomplish a learning task. Cosmo is represented as a three-dimensional virtual character and is present in the scene constantly, observing the action of the user and providing hints, explanations, and help. The agent can answer questions, can jump in if the user does not know any further, and can point out false steps in a solution attempt.

Similar characters have also been used in fully immersive virtual training environments. The Soar Training Expert for Virtual Environments ('Steve') is an embodied agent for teaching users how to accomplish manual tasks [15]. The user and the agent are immersed together in a virtual reality interface, where the user can perform actions via multi-dimensional input devices and can see their effect on the environment.

Recent research has also explored how to give virtual characters understanding of the user's speech and gesture interaction in the real world. One of the first virtual characters with multimodal input and output behavior output was Gandalf, developed by Thorisson [18]. Gandalf was embodied as a virtual hand and a face which appeared on a monitor or alternatively on a big screen. Interaction was possible using natural speech and gestures, although the user had to wear motion tracking hardware. This work was followed by Rea [7], a character with a full virtual body embedded in a 3D virtual environment. Computer vision techniques allowed Rea to respond to user body motion without the need for encumbering technology. Unlike earlier systems, Rea was designed with a conversational model based on research findings for human face-to-face interaction.

2.2. Studies with Virtual Characters

Although virtual characters have progressed from disembodied heads to full bodied figures that can recognize natural speech and gesture, there have been relatively few studies on the effect of character representation on user experience.

Parise et. al. [13] found that commitment and cooperation with anthropomorphic agents were significantly greater than with non-human like agents characters. However, participants rated the non-human like agents as most 'cute' and trustworthy. They also reported that human-like characters did not need to be photo realistic for users to want to cooperate with them. However, studies show the greater the realism of the character the greater the perceived expectation of intelligence [10],[11]. In studies with the DECface [22] users gave less personal information to the talking face than through a text interface. On the other hand, they attributed more positive personality characteristics to the more pleasant looking agent as well. Thus creating a more visually realistic character may not necessarily create greater engagement with the character, even if users do think the character is more intelligent.

2.3. AR Characters

In our work we are interested in character representation and particularly in the presentation of virtual characters in an augmented reality setting. For the first time, AR technology allows virtual characters to exist in the same real space as the user. Despite this, the use of AR characters has also not been well studied.

In the Welbo system [1] an animated virtual robot assists an HMD equipped user in setting up virtual furniture. Welbo has speech synthesis capabilities and can understand simple instructions. It is aware of the user's actions and movement and reacts to commands by moving furniture or pointing to objects.

In one of the first examples of an AR character, Balcisoy et. al. [3] created a virtual agent that could play checkers with a person in the real world. The agent didn't have any conversational ability, nor was it able to respond to a real user's speech and gesture commands. Simply by appearing in the same space as a real user, the authors say it creates a strong sense of presence. However, Balcisoy also did not report on any formal user studies exploring how the user's felt about the agent.

In the AR Puppet project Barakonyi studied how animated characters improve the man-machine communication in AR applications. In his work he focuses on the interaction of virtual characters with their virtual as well as physical environment. In the AR game MonkeyBridge [5] two players have to help their autonomous agents in form of monster-like characters to cross a river. The characters are not scripted but intelligently decide which virtual and physical objects to use in order to accomplish the task. The AR Lego [4] application employs two agents: a physical robot and a virtual animated repairman to assist an untrained user in assembling and maintaining a LEGO Mindstorms robot. Neither of these interfaces was evaluated in formal user studies.

These interfaces show that AR characters have an interesting potential for entertainment applications. However there has been no formal evaluation comparing an AR representation to other types of characters. In our work we are interested in the effect of character representation, especially on a handheld platform.

2.4. Handheld AR

As significant computing and graphics power became available on the handheld platform, researchers have naturally begun to explore the use of personal digital assistants (PDA's) for AR applications. First work such as the AR-PDA project [8] and BatPortal [9] use the PDA as a thin client for showing AR content generated on a remote server. This was necessary as the early PDAs did not have enough capability for stand-alone AR applications. Then in 2003 the first author of this paper ported ARToolKit [2] to the PocketPC and developed the first fully self-contained PDA AR application [21]. Since that time, Studierstube [14] has been ported to the handheld platform and the first stand alone collaborative PDA AR applications have been developed.

The PDA platform proved to be more cost-effective and less error-prone than HMD-based mobile setups and ideal for lightweight social interactions. There are many possible entertainment and educational applications. For example, in the Invisible Train game [19] two players control their virtual trains on a real wooden miniature railroad track. The virtual trains are only visible to the players through the PDA's video see-through

display. In the AR-Kanji game [20] players learn Kanji symbols. The game uses a small set of cards, each having a Kanji symbol on one side and an ARToolkit marker on the other side. The PDA asks the player to find the Kanji symbol for a specific object. As the player turns a card, either the correct virtual object or a question mark is shown.

Experiences with previous projects make us believe that the PDA is an ideal platform for Augmented Reality especially for untrained users. While HMDs are expensive, fragile and difficult to handle, PDAs have low per-unit costs, a compact form factor and their low weight allows comfortable single-handed usage even over longer periods. Its touch screen enables the creation of intuitive user interfaces.

2.5. Our Work

As can be seen we have drawn on a wide range of related work for this project. However our research is different in several ways:

- It is one of the first uses of a virtual character in a handheld AR application
- It is the first time that the effect of visual realism has been formally evaluated with an AR character
- It is the first use of an AR character in an educational application.

In the remainder of the paper we describe in more detail our handheld AR platform and the virtual character we have developed for the PDA. We also give results from a study that compares engagement with an AR character to other types of character representations.

3. THE HANDHELD AR PLATFORM

We have created a component-based software architecture for the development of PDA-based handheld AR applications. Our approach was to rely on existing software tools if possible, while providing high performance and suitability for a small mobile platform. Programming tools for PDA-development are error-prone and debugging remains a slow and cumbersome task – even on device emulators. Consequently, developers attempt to write and test most of their code natively on a workstation. All components of our framework are available for PDAs running Windows CE as well as PC-based workstations running Windows 2000/XP. Figure 2 shows a simplified chart of the software design of our framework.

While handheld platforms nowadays often include support for OpenGL ES, most existing software still relies on Open GL. To overcome this gap we developed a library called Klimt¹ that wraps an underlying OpenGL ES implementation and adds the most important missing OpenGL and WGL features such as floating-point data types, missing primitive types, etc. Using Klimt, many existing OpenGL applications require only minimal modifications in order to run hardware-accelerated on a PDA. Our AR framework Studierstube [16] supports multiple concurrent applications. It does not directly interface with OpenGL, but instead builds on top of the open source scene graph library Coin². We ported the Coin rendering library, which implements the Open Inventor API, to Windows CE, running on top of OpenGL/Klimt.

¹ Klimt: open source 3D graphics library for mobile devices, <http://studierstube.org/klimt>

² Coin3D: open source scene graph renderer, <http://www.coin3d.org>

The existence of an either built-in or attachable camera lends itself to use computer vision as the primary tracking method on a PDA-based platform. We ported the well-known marker tracking library ARToolkit [2] to Windows CE, improved its performance and outfitted it with many new extensions crucial to handheld devices. Our library called ARToolkitPlus³ can track at a top speed of 5.0 ms per image on a current PDA, so self-contained vision tracking is no longer a major bottleneck for embedded devices. We implemented a new marker system that can handle up to 4096 markers and real-time dynamic thresholding to compensate for the ever-changing lighting conditions of mobile applications.

For massive multi-user support, we developed our own real-time database called Muddleware, which is implemented as a memory-mapped XML database on the server side that can easily handle thousands of requests per second. Furthermore being a database, Muddleware implicitly provides persistence, which we found crucial for a setup with an arbitrary number of light-weight clients that can enter and leave at any time. On the client side Muddleware provides a simple API (C++, Java and XML script) to address data associatively via XML XPath on the server. To improve robustness and portability Muddleware works on top of ACE⁴.

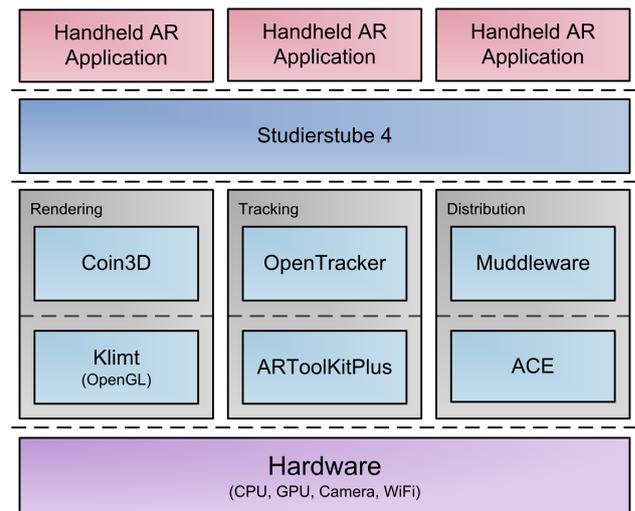


Figure 2: Software design of the developed framework

4. MR VIRTUOSO AR CHARACTER

Using our handheld AR platform, we recently implemented a virtual character called Mr. Virtuoso who is a fully animated and textured 3D figure. As shown in Figure 1 Mr. Virtuoso can appear to exist superimposed over the real world.

The lack of processing and graphics power on mobile devices prevents the use of traditional character animation techniques. For example most existing animation libraries make heavy use of floating-point math which is not available in hardware on today's PDAs. We therefore developed an animation package called FPK that makes best use of the particular restrictions of mobile

³ ARToolkitPlus: http://studierstube.org/handheld_ar

⁴ ACE: The ADAPTIVE Communication Environment, <http://www.cs.wustl.edu/~schmidt/ACE.html>

devices. FPK uses pure fixed-point math for highest play-back performance and stores all geometrical data as 16-bit values (coordinates) and lookup indices (normals) to effectively reduce memory load. Animations are stored as keyframes and linear per-vertex interpolation (“vertex-tweening”) is used for play back. As a result, playing back a 3D character on the PDA using linear tweening causes almost no noticeable overall performance drop. Animations can be imported from MD2⁵ and Cal3D⁶ files.

FPK is split into two separate layers: a lower layer providing basic animation services, and a higher layer exposing a scripting interface via a simple XML dialect. Using this XML dialect authors can create complex sequences of animations including precise timings for audio dubbing and subtitle rendering.

5. ART HISTORY APPLICATION

We tested the different virtual character representations in our recently implemented art history educational game. The players’ objective is to sort a collection of artworks according to their date of creation along a timeline drawn on a wall-mounted billboard, see Figure 3.



Figure 3: The AR History Game

Every marker on the timeline carries one of the artworks, which are only visible through the player’s AR PDA. Initially the artworks are in random order. A player can pick up any artwork with his or her PDA, by clicking on the artwork on the display and drop it on a free position by clicking on an empty marker on the display. While an item is located on the PDA, the player can access explanations about it by placing the artwork on Mr. Virtuoso’s desk (see Figure 1). Mr. Virtuoso will then be prompted to provide his expertise on the subject through the use of text, audio playback and animation. After an artwork is placed onto its correct position, it cannot be moved again.

Besides the hints from Mr. Virtuoso, the game engine can provide even more help in several ways: It can show arrows pointing “left” and “right” next to the artwork if it should be placed earlier or later on the timeline (see Figure 4). Furthermore the game engine can display an item’s date of creation when the item is placed at its correct location on the timeline. If the timeline is very long (more than 10 items) players can easily lose oversight. To prevent that, the game can display the timeline as a series of icons

on the bottom of the PDAs’ screens. The game master can enable and disable any of these options at any time during the game.

The art history application features an overall selection of 20 artworks from which the game master can select a subset for play. The game features textured, animated 3D models, multimedia background material and pre-recorded audio narration in three languages (English, German and Spanish). A graphical user interface for the game master allows runtime configuration of all game features.

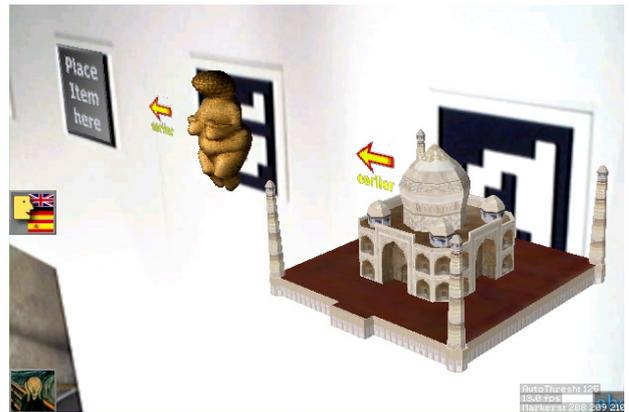


Figure 4: Arrows can help finding an item’s correct location

6. EXPERIMENTAL EVALUATION

The original art history game can be played collaboratively by up to four players. Since we wanted the subjects to focus on the learning part and the virtual character rather than on collaboration with other players, we created a modified version that was played by a single player. All helping options such as arrows pointing into the direction of an item’s correct spot on the timeline were turned off.

In order to explore the effect of the virtual character representation we conducted an experimental evaluation with five experimental conditions:

- A** Text only: The virtual character is just represented by text windows appearing on the screen (Figure 5a).
- B** Text and Audio: As in condition A, but in addition an audio voice over was played (Figure 5a).
- C** 2D Image: As in condition B, but in addition a 2D image representing the character was shown on the screen (Figure b)
- D** 3D Character: As in condition B, but in addition a 3D animated virtual character was shown on the screen. The character was fixed to the screen as a TV moderator (Figure 5c)
- E** AR Character: As in condition D, but the 3D virtual animated character appears fixed in space in the real world (Figure 5d).

Condition A includes an absolute minimum presence of the virtual character. Conditions add progressively more and more realistic cues, while condition E (right-most picture in Figure 5) is the only case where the virtual character is seen as part of the user’s physical environment, and so it is the only true AR condition. It should also be noted that the virtual head shown in condition C is not animated, unlike conditions D and E.

⁵ MD2: file format for animated 3D models in Quake2

⁶ Cal3D: open source character animation library

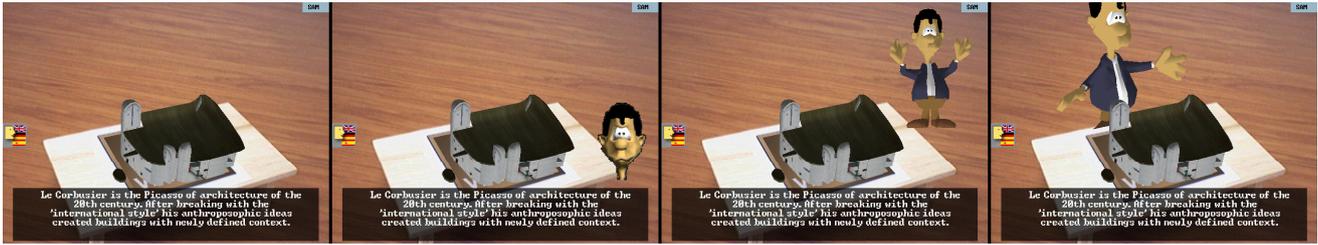


Figure 5: The five cases of Mr.Virtuoso: left-most picture: text and/or audio only; picture 2: 2D image of Mr.Virtuoso’s head; picture 3: screen aligned and 3D animated; right-most picture: fully 3D registered

For the experiment, users played the game one in each condition, trying to correctly arrange four artworks each time. They were told to use Mr. Virtuoso to learn as much as possible about the art works as they would be tested on their knowledge after each condition. The emphasis was on learning about the art rather than correctly arranging the artworks in the quickest amount of time. The order of presentation of the conditions and the artworks for each condition were changed to prevent order effects in the results.

For each condition we measured the time taken to complete the task, and we asked the users a number of questions relating to how much they enjoyed playing the game, how real they thought the character was and how much they learnt. Subjects were also asked a number of multi-choice questions about the artworks, asked to rearrange pictures of the artworks and interviewed about the experience.

The key questions we were looking to answer include:

- Is there a relationship between the character representation and perceived realism?
- Is there a relationship between character representation and enjoyment of the experience?
- Is there a relationship between character representation and how much people felt they learnt?

Knowing the answers to these questions may help developers create more effective virtual character based entertainment experiences in the future and also better understand how AR technology can be used to develop new types of characters.

7. RESULTS

There were 13 participants of which 9 were male and 4 were female, aged 20 to 33 years. Most of the participants were native English speakers. None of them knew the game before. The experiment lasted about 40 minutes per subject including a short finishing discussion. Data analysis was performed SPSS version 13. The main effect was tested with repeated ANOVA. If a main effect was found, pair-wise post-hoc comparisons using the Bonferroni adjustment for multiple comparisons were performed.

Table 1: Average Number of Times Mr. Virtuoso used

Cond.	A	B	C	D	E
Asked #	3.3	3.6	3.5	3.6	3.9

Subjects used Mr. Virtuoso heavily, Table 1 showing the average number of times used per condition. There were four artworks

shown per condition, so the agent was used almost once per artwork to discover more information about art. ANOVA shows no significant difference ($F(4,60) = 0.26, P = 0.90$) between conditions in the time taken to arrange the art works in the correct order with.

Table 2: Time to perform task

Cond.	A	B	C	D	E
Time (s)	248.5	230.0	236.2	237.3	252.7

Table 3 shows the average results that users got right on the four multi-choice test questions on the artwork after each condition. No significant differences were found ($F(4,45) = 1.01, P = 0.41$).

Table 3: Average Number of Questions Correct

Cond.	A	B	C	D	E
Score	3.0	2.8	2.5	2.3	2.6

Subjects were asked to mark on a Likert scale of 1 to 7 how much they agreed or disagreed with a number of statements, where 1 = Strongly Disagree and 7 = Strongly Agree.

There were a number of questions about the game and ease of use of the interface, including:

Q1: I enjoyed playing the art history game

Q2: The PDA interface was easy to use

Q3: The task was easy to solve

Q4: I felt I learned new facts about art items from the game

Table 4 shows the average results for each of these questions. As can be seen there is little difference between conditions for these results. ANOVA tests found no significant differences for the user survey scores for these questions.

Table 4: Subjective Survey Scores

Condition	A	B	C	D	E
Q1	5.46	5.85	5.92	5.85	5.85
Q2	6.08	5.92	5.92	6.08	6.15
Q3	5.69	5.69	5.84	6.00	5.92
Q4	5.39	5.85	5.92	5.84	6.15

A second set of questions related to the virtual character:

- Q5: *Mr. Virtuoso seemed real to me*
- Q6: *Mr. Virtuoso was helpful for completing the task*
- Q7: *Mr. Virtuoso improved the overall experience*
- Q8: *I found Mr. Virtuoso to be friendly*
- Q9: *Mr. Virtuoso seemed to be part of the real world*

There was a significant difference between the results for all of these questions. Figure 6 shows the average results for Q5: *Mr. Virtuoso seemed real to me*. As the virtual character exhibits more visual and audio cues the subjects felt that it was real. An analysis of variance was conducted with type of virtual character (A – E) as the within-subjects factor. Doing this we found a significant difference between conditions ($F(4,48) = 11.18, P < 0.001$). Post-hoc found that Mr. Virtuoso in conditions E ($P < 0.001$) and D ($P < 0.001$) was rated as significantly more real than in condition A. Condition D was also rated significantly higher than condition A ($P < 0.01$). There were no other significant differences between conditions.

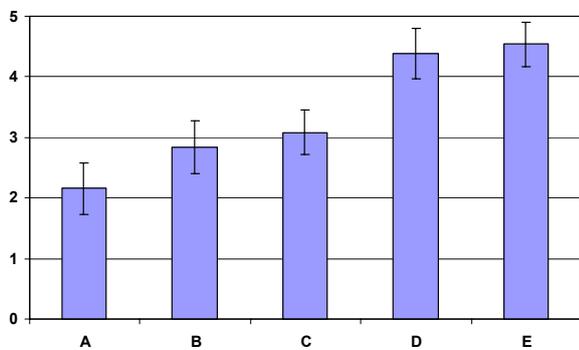


Figure 6: How real Mr. Virtuoso was (Q5)

Table 5 shows the average results for Q6, Q7 and Q8. An ANOVA on Q6: *Mr. Virtuoso was helpful for completing the task* showed a significant difference across conditions ($F(4,48) = 8.186, P < 0.001$). Post-hoc comparisons between conditions found a significant difference between conditions D and E and condition A ($P < 0.05$). Similarly, an ANOVA on Q7: *Mr. Virtuoso improved the overall experience* produced a significant difference across conditions ($F(4,48) = 5.22, P < 0.01$). Finally an ANOVA on Q8: *I found Mr. Virtuoso to be friendly* produced a significant difference across conditions ($F(4,36) = 12.322, P < 0.001$). Post-hoc comparisons found a significant difference between conditions D and E and condition A ($P < 0.01$), and conditions B and A ($P < 0.05$). In all cases the score of the condition without audio (condition A) was lower than the other conditions, while the two conditions with 3D graphics (conditions D and E) were the highest.

Table 5: Helpfulness, Experience, and Friendliness

Cond.	A	B	C	D	E
Q6	3.46	4.92	4.85	5.69	5.85
Q7	3.23	4.69	4.31	5.15	5.39
Q8	2.10	4.50	3.90	5.20	5.20

As the quality of the character representation increased it also seemed to be more part of the real world. Figure 7 shows a graph

of the average response to question 9: *Mr. Virtuoso seemed to be part of the real world*. An ANOVA showed a significant difference across conditions ($F(4,36) = 6.46, P < 0.001$). Post-hoc comparisons found a significant difference between conditions C and A ($P < 0.05$), conditions D and A ($P < 0.05$), and conditions E and A ($P < 0.05$). As before the 3D virtual characters (D, E) are significantly different from the text-only condition (A).

In addition to providing subjective survey responses, subjects were also asked to rank each of the conditions in order according to the following criteria. For each criteria 1 = lowest, 5 = highest.

- R1: *How real Mr. Virtuoso seemed*
- R2: *How much fun it was*
- R3: *How much you learnt*
- R4: *How helpful was Mr. Virtuoso*

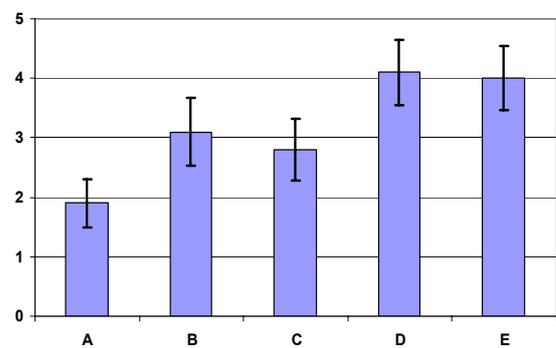


Figure 7: How much Mr. Virtuoso seemed part of the real world.

The rankings for R1 and R2 are significantly different across conditions. An ANOVA for R1 finds ($F(4,44) = 67.42, P < 0.001$). Post-hoc comparisons show significant difference ($P < 0.02$) between all conditions except B and C. In fact, all of the subjects except one ranked the AR condition as most real. Similarly, an ANOVA for R2 found $F(4,48) = 30.25, P < 0.001$. All but two of the subjects ranked the AR condition either highest or second on how fun it was. Figure 8 shows the results for rankings R1 and R2. In this case, when users where forced to chose, as the virtual character had more realistic characteristics they thought it was more real and correspondingly more fun.

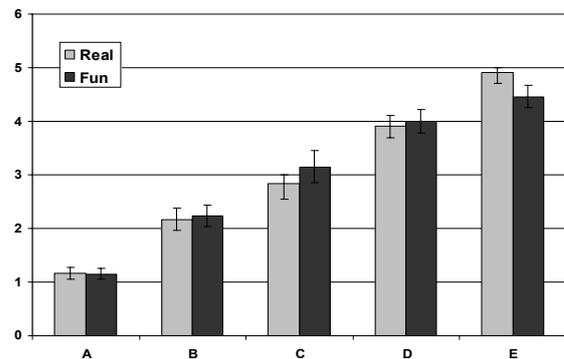


Figure 8: Average ranking on realism and fun

However there was no significant difference between rankings on R3: *How much you learnt*. An ANOVA finds ($F(4,32) = 0.80, P = 0.53$). This is consistent with the survey results for Q4 and the multi-choice question results. Table 6 shows the average rankings.

For ranking R4: *How helpful Mr. Virtuoso was*, an ANOVA produced a significant difference across conditions ($F(4,36) = 3.78, P < 0.01$). Post-hoc comparisons confirm that this is because of the difference between the condition with no audio (Condition A) and the other conditions.

Table 6: Ranking of Learning and Helpfulness

Cond.	A	B	C	D	E
R3	2.54	3.67	3.19	2.55	2.91
R4	1.93	3.15	3.50	3.23	3.23

8. INTERVIEWS

In the interviews with the subjects several consistent themes emerged. Although we emphasized before and during the game that time was of no importance, 4 out of 12 users complained afterwards that they felt slowed down by Mr. Virtuoso, particularly in condition E where they needed to wait for his walking animation to finish. The main reason for this that Mr. Virtuoso presented information in his own speed and could not be interrupted. Since this was observed negatively by many players, we plan to remove this issue for a next version of the game.

Most subjects pointed out that they were very aware of the fact that the more feature-rich versions of Mr. Virtuoso did not provide more information than the other versions. Still they usually liked the AR version more because they felt it looked more natural and realistic. However many subjects pointed out the importance of Mr. Virtuoso's voice feature as being critical.

Three subjects complained that Mr. Virtuoso did not make eye contact with them while speaking which made them feel uncomfortable or even offended them. Condition D (screen-aligned) did not have this problem since the screen-aligned character implicitly looked in the player's direction. Those three subjects generally preferred condition D over E.

One subject did not like the AR version because he had the feeling that he was interrupting Mr. Virtuoso from his other actions when asking him for assistance. This was not an issue with the screen-aligned version because Mr. Virtuoso would walk into the screen from the right side and would therefore not be visible before.

Finally, two players felt that the animated character distracted them from reading the text and listening to the voice.

9. DISCUSSION

The main focus of our study has been to explore the effect of different virtual character representations on user engagement, enjoyment and educational outcomes in a learning task. The objective measures of time to complete the task and number of correct test questions did not vary significantly across conditions, showing that the various character representations did not have any effect on the educational outcomes. Similarly when the users were asked their subjective opinion of the different conditions there was no difference in how much they enjoyed each condition or how difficult they found the task and how much they felt they learnt. Although when forced to rank the conditions in order, as the virtual character became more realistic the users felt the condition was more fun (Figure 8).

This is not too surprising since users were engaged in a relatively short task and as the virtual character representation becomes increasingly realistic it did not provide additional educational content, such as gesturing to specific parts of the virtual artwork. User comments indicate that if the task had a greater spatial learning element (such as learning how an engine is taken apart) the more realistic characters could have an impact on the learning outcomes. Whether or not this is true, would be subject of another user study.

The increasing enjoyment ranking when forced to make a choice may be due to the novelty factor; AR characters are more novel than a disembodied voice over and so relatively more fun.

The subjective results in response to the character did show significant differences across conditions. When asked how real Mr. Virtuoso seemed and how much he seemed to be part of the real world, the main differences was between 3D and non-3D representations of the character and audio vs. non-audio. In all cases the non-audio condition rated lowest while the 3D character was the highest, but there was no difference between AR and screen aligned characters.

It is interesting to note the effect that adding audio can have on the user's perception, causing a large jump in average scores between conditions A and B, while there was no additional benefit of adding a 2D representation to the audio. Although if the 2D character had been animated this may have had a greater effect. Players did not rate the AR version more realistic than the screen-aligned 3D virtual version and it was not perceived to be more helpful or friendly than the 3D virtual version. Only when forced to choose did subjects rank the AR version more realistic and more fun.

One of the reasons for this could be that the AR character did not exhibit any more communication cues than the screen-aligned 3D virtual character that would make it seem more real. Although he would walk around the real table, the AR version of Mr. Virtuoso did not give any spatially related information on the objects such as pointing to specific spots on items as a real person would do while explaining. Furthermore a realistic person is expected to behave politely and to look at the people he is talking too.

10. CONCLUSION AND FUTURE WORK

In this paper we have presented the first study that compares user response to an Augmented Reality virtual character to other types of character representations. Our study shows that an Augmented Reality character is not automatically superior to a 3D screen-aligned virtual counterpart, but in general 3D virtual characters (AR and non-AR) are preferred over other representations. However, regardless of the character representation there was no effect on the learning outcome for this task.

There are several lessons that can be learned from this study that could be used to inform the design of virtual character interfaces:

1. Audio is extremely important and can improve a virtual characters perceived friendliness, and how real it seems.
2. Animated 3D virtual characters provide a significant improvement over non-animated 2D and text only characters in terms of how real they are perceived to be.
3. Users feel that AR characters and screen aligned 3D virtual characters are equally part of the real world.

4. Unless AR characters are shown doing a task interacting with real world or virtual content there may be little benefit using them in an educational setting compared to screen-aligned 3D virtual characters

Clearly further work needs to be done to understand the types of applications where an AR character may provide significant benefit over non-AR character representations, especially those educational tasks that provide a significant spatial learning component. It seems an AR character does not only have to look more realistic than a 2D or 3D screen-aligned agent, but also behave more realistically, such as being able to gaze at the user while talking to them.

As a next step we intend to improve our virtual character to overcome the identified weaknesses and use it in further applications. To do this we will use the presented framework in a project that brings mobile Augmented Reality into museums.

11. ACKNOWLEDGEMENTS

We would like to thank our colleagues at Graz University of Technology for helping in creating the application, the colleagues at HIT Lab New Zealand for helping with the study and the staff of Imagination Computer Services for helping in creating the 3D content. This project was funded in by part by Austrian Science Fund *FWF* under contract No. L32-N04 and Y193.

12. REFERENCES

- [1] Anabuki, M., Kakuta, H., Yamamoto, H., Tamura, H., Welbo: An Embodied Conversational Agent Living in Mixed Reality Space, *CHI 2000*, Extended Abstracts, 2000, pp. 10-11.
- [2] ARToolkit website: www.hitl.washington.edu/artoolkit/.
- [3] Balcisoy, S., Torre, R., Ponder, M., Fua, P., Thalmann, D., Augmented Reality for Real and Virtual Humans. *IEEE Computer Graphics International*, pp. 303-308, 2000
- [4] Barakonyi, I., Psik, T., Schmalstieg, D.: Agents That Talk And Hit Back: Animated Agents in Augmented Reality, *IEEE and ACM International Symposium on Mixed and Augmented Reality 2004 (ISMAR'04)*, Nov. 2-5, 2004, Arlington, VA, USA, pp. 141-150
- [5] Barakonyi, I., Weilguny, M., Psik, T., Schmalstieg, D., MonkeyBridge: Autonomous Agents in Augmented Reality Games, *ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'05)*, Valencia, Spain, June 15-17, 2005
- [6] Beskow, J., Elenius, K., McGlashan, S.: Olga - A dialogue system with an animated talking agent. *In Proceedings of Fonetik 97*, Umeå, Sweden, May 1997
- [7] Cassell, J., Bickmore, T. W., Billinghurst, M., Campbell, L., Chang, K., Vilhjálmsson, H. T., Yan, H.: Embodiment in Conversational Interfaces: Rea. *Proceedings of the conference on Human Factors in Computing Systems (CHI 99)*, 1999 pp. 520-527
- [8] Gausemeier, J., Fruend, J., Matyszczok, C., Bruederlin, B., and Beier, D., Development of a real time image based object recognition method for mobile AR-devices. *Proc 2nd international conference on Computer graphics, Virtual Reality, visualisation and interaction in Africa*, pp. 133-139, 2003
- [9] Ingram, D., Newman, J., Augmented Reality in a WideArea Sentient Environment. *Proc. of the 2nd IEEE and ACM International Symposium on Augmented Reality (ISAR 2001)*, October 2001, New York.
- [10] Keeling, K., S. Beatty, et al. (2004). Face Value? Customer Views of Appropriate Formats for Embodied Conversational Agents (ECAs), *Online Retailing. 37th Hawaii International Conference on System Sciences*, 2004..
- [11] Koda, T. and Maes, P., Agents with faces: the effect of personification. *5th IEEE International Workshop on Robot and Human Communication*, London, 1996
- [12] Lester, J., Voerman, J., Towns, S., Callaway, C.: Cosmo: A Life-like Animated Pedagogical Agent with Deictic Believability, *Working Notes of the IJCAI '97 Workshop on Animated Interface Agents: Making Them Intelligent*, pp. 61-69, Nagoya, Japan, August 1997
- [13] Parise, M., S. Kiessler, et al., My partner is a Real Dog: Cooperation with Social Agents. *Computer Supported Cooperative Work '96*, Cambridge, MA, USA, ACM Press, New York, NY, USA. 1996
- [14] Reitmayr, G., Schmalstieg, D., Mobile Collaborative Augmented Reality. *Proc. ISAR 2001*, New York, USA, Oct. 2001.
- [15] Rickel, J., Johnson, W.L.: Animated agents for procedural training in virtual reality: perception, cognition, and motor control, *Applied Artificial Intelligence, Vol. 13*, pp. 343-382, 1999.
- [16] Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavári, Z., Encarnação, L.M., Gervautz, M., Purgathofer, W.: The studierstube augmented reality project. *Presence: Teleoperators and Virtual Environments 11* (2002) 33–54.
- [17] Sunblad, O., Y. Sundblad. OLGA - a Multimodal Interactive Information Assistant. *Proceedings of the conference on Human Factors in Computing Systems (CHI 98)*. 1998
- [18] Thórisson, K. R., Gandalf: An Embodied Humanoid Capable of Real-Time Multimodal Dialogue with People. *First ACM International Conference on Autonomous Agents*, California, February 5-8, 536-7.
- [19] Wagner D., Pintaric T., Ledermann F., Schmalstieg D. Towards Massively Multi-User Augmented Reality on Handheld Devices. *Proc. of the Third International Conference on Pervasive Computing (Pervasive 2005)*, Munich, Germany, 2005.
- [20] Wagner, D., Barakonyi, I.: Augmented reality kanji learning. *Proc. of the 2003 IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2003)*, Tokyo, Japan, IEEE Computer Society (2003) 335–336.
- [21] Wagner, D., Schmalstieg, D.: First steps towards handheld augmented reality. *Proc. of the 7th International Symposium on Wearable Computers (ISWC2003)*, White Plains, NY, USA, IEEE Computer Society (2003) 127–137.
- [22] Waters, K., Levergood, T. M., DECface: A system for synthetic face applications. *Multimedia Tools and Applications*, 1 (1995), 1-16.