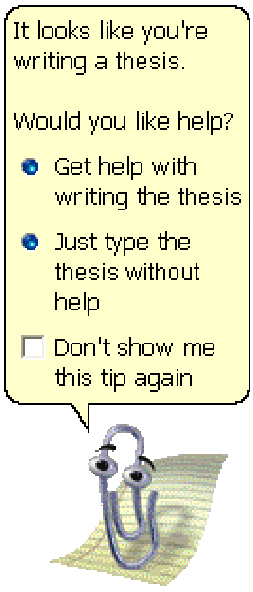
A man’s best friend?

A study into subjective user experience and task performance with a human guide and an embodied agent.



by

Renate ten Ham

31-08-2005

A man's best friend?

A study into subjective user experience and task performance with a human guide and an embodied agent.

R. H. ten Ham
Communication Studies
University of Twente
31-08-2005

First Supervisor: Ard Heuvelman
Second Supervisor: Mariët Theune

Samenvatting

Binnen dit onderzoek zijn een tweetal experimenten met een virtueel personage uitgevoerd. Het doel was onderzoeken in hoeverre een virtueel persoon en een echt persoon die een routebeschrijving geven inwisselbaar zijn en welke rol gebaren hierbij spelen. De resultaten van het eerste experiment onder studenten wijzen uit dat gebaren voor de subjectieve ervaring van gebruikers zeer nuttig zijn. De proefpersonen die de routebeschrijving met gebaren hebben gezien, waren veel enthousiaster over de virtuele persoon dan degenen die de routebeschrijving zonder gebaren hebben gezien. De virtuele persoon met gebaren werd als zeer goed ervaren en de resultaten wijzen erop dat bij jongeren een virtueel persoon zeer goed ingezet kan worden voor het geven van informatie. Met het tweede experiment werd gekeken of leeftijd een beïnvloedende factor is. Uit de resultaten bleek dat oudere gebruikers subjectief gezien minder goed met de virtuele persoon overweg kunnen dan de jongere gebruikers. Opvallend resultaat is echter dat er geen significant verschil is in de hoeveelheid informatie die ouderen zich konden herinneren van de virtuele persoon ten opzichte van de echte persoon. De resultaten wijzen erop dat een virtuele persoon voor zowel jongeren als ouderen goed te gebruiken is, hoewel ouderen iets weerstand tegen werken met een virtueel persoon lijken te hebben.

Abstract

Within this project, two experiments were conducted with a virtual character. The goal was to investigate to what extent a virtual character can replace a real human in the field route descriptions and what role gestures play during interaction with a virtual character. The outcomes of the first experiment among students give an indication that gestures are very useful for the subjective experience from users. The participants who saw the route description with gestures, were much more enthusiastic that those who saw the route description without gestures. The virtual character received very good reviews. The results indicate that among younger people, a virtual character can very well be used to provide information. The second experiment was used to investigate if age was an influencing factor. The results indicated that older users didn't like the virtual character as much as the younger participants did. However, there is no significant difference in the amount of information that was recalled between older participants who saw the virtual character and those who saw the real human. The results indicate that both younger and older people can interact with a virtual character, although the older people seem to have some resistance towards working with a virtual character.

Table of Contents

Preface

Last Friday, I went to a graduation party. Somebody asked me about my graduation subject and I started talking about virtual characters and the experiments I conducted. The person who asked me about my subject sighed and said, "yeah, when I just started on my project, I was probably as enthusiastic as you are right now". He seemed sincerely surprised when I told him I was almost finished and would graduate within a week.

It has been a great experience to round off my Communication Studies with a subject I really like. However, without the good advice, help and pep talks of my supervisors Ard Heuvelman and Mariët Theune this thesis would probably never been created.

The company and advice of computer science students such as Ronald, Mattijs and Erik have made my stay at this faculty a lot more pleasant than expected.

Furthermore I would like to thank Zsofia Ruttkay and Jan van Dijk for their advice and useful hints. Also I would like to thank Benoit Morel of Cantoche for the use of their virtual characters.

Probably all my friends have a smaller or bigger part in moral support, practical work or were otherwise helpful, but special thanks are for: Laura, Gwendy, PP, Lizanne, Florian, Marit and Hélène.

Last but certainly not least I would like to thank my boyfriend, Sander. He has this funny habit of trying to tidy op my desk/bag/computer or all other places I used to spread around my articles, books and other stuff I needed. This usually resulted in me running around and telling him I can't find anything anymore. However, without him the process of graduation would have taken much, much longer.

Renate ten Ham

1 Introduction

1.1 Background

At the Computer Science Department of the University of Twente, the Human Media Interaction group works on several projects. One of them is Angelica: A Natural-language Generator for Embodied, Lifelike Conversational Agents. The aim of this project is the combined generation of language and nonverbal signals in information presentation by embodied agents. The main question of this project is which modality (verbal or nonverbal), or combination of modalities, to use for expressing a given piece of information. The focus will be on pointing and iconic gestures, which are used to identify objects and to express concepts. The application domain for this research is the generation of route descriptions. Although several students in computer science already graduated within this project, some students in communication are now invited to write a thesis on this subject, in association with the Human Media Interaction group. This way the project will be regarded from a new perspective; technology and communication combined.

1.2 Relevance of Research

Probably most people are familiar with the somewhat annoying Windows paperclip " Clippy", or the dog that will help you search your computer. These are some of the first attempts to make the program interface more friendly, by giving a level of personal assistance while working with the computer. Nowadays, more developed characters are to be found on more and more places on the Internet. You can find them in roles such as presenters, teachers or salespersons. They will help you fill in forms, point out the relevant content on a website or just be there to give the website a more lively look. These characters are being equipped with human-looking virtual characters that can use natural language and display nonverbal behaviours, to make human-computer interaction similar to face-to-face communication between humans. These characters are referred to using different terms, including 'synthetic personae' (McBreen, Shade, Jack & Wyard, 2000), 'embodied conversational agents' (ECA) (Cassell, Sullivan, Prevost, and Churchill, 2000), and 'animated interface agents' (Dehn & Mulken, 2000). For brevity, in this thesis they will be called 'embodied agents' or simply 'agents'. As embodied agents grow more intelligent, the amount of useful applications grows as well.

Increasingly, agents are used for tasks that are traditionally performed by humans, such as providing information, explaining or answering questions as an instructor or a teacher. Allbeck and Badler (2002) argue: "virtual humans can represent other people or function as autonomous helpers, teammates, or tutors enabling novel interactive educational and training applications". Lester, Zettlemoyer, Gregoire and Bares (1999) state "… because of their strong visual presence and clarity of communication, explanatory lifelike agents offer significant potential for playing a central role in next-generation learning environments." The question is how people experience working with an embodied agent. Will their subjective experience be different if they receive information from an embodied agent instead of a human presenter? Most research into agents and the user response to agents contains a comparison between an agent and a text or speech only condition. A comparison between a full bodied human guide and an embodied agent might be a useful addition to answer the question to what extent an agent can replace a human presenter. Because the application domain of Angelica, the framework within this thesis is written, a route description is chosen as the presentation task the agent and human presenter will perform.

1.3 Project Objectives

The main goal of this study is to investigate to what extent an agent can be used instead of a human in a direction-giving situation.

In order to reach the above-mentioned main goal, an answer has to be given to the following main question:

Is a lifelike agent comparable to a real person in a direction-giving situation by measuring subjective user experience and task performance?

Five sub questions are formulated to help answering this main question:

- Do gestures have an impact on subjective experience with either a human guide or an embodied agent giving a route description?
- Do gestures have an impact on task performance with either a human guide or an embodied agent giving a route description?
- What are the differences in subjective user experience between the participants who saw the human guide and those who saw the embodied agent?
- What are the differences in task performance between the participants who saw the human guide and those who saw the embodied agent?
- Is age an influencing factor in how people judge their subjective experience with an agent or a human guide when given a route description?

Initially, an experiment with students was conducted. The main goal was to compare the response of participants confronted with an embodied agent explaining a route to the response of those who saw a human guide. The second goal was to investigate the influence of gestures. After the first experiment among students, the question arose if age could be an influencing factor. At that point, the decision to conduct a second experiment was made. Because of this chronological course, this thesis is divided into four parts.

The structure of this thesis is as follows. In the second chapter, an overview of the most relevant literature will be given. This literature research was performed to find relevant theories and to find a suitable embodied agent for the experiment. Chapter three contains the method, chapter four the results and the conclusions of the first experiment among students can be found in chapter five.

Again, a literature study was necessary to discover if any experiments were done with embodied agents among seniors. This is described in chapter six. This literature study resulted in another experiment. This experiment has the same structure as the first experiment, containing method, results and conclusions in chapters seven, eight and nine. The final conclusions can be found in chapter ten.

2 Literature overview Human-Computer Interaction

When a user is interacting with an embodied agent, this is called Human-Computer-Interaction. On one side is the computer, with its technological aspects. With the current developments, high-tech gadgets and more advanced technology become readily available for users at home. Designers work on computers and interfaces that are able to communicate with users in a different way than people ever expected from their computer. On the other side of the Human-Computer Interaction is the user, with his human nature of communicating and human way of response. These are two different perspectives of the same Human-Computer Interaction. Therefore, in this chapter the computer perspective and the human perspective will be separately discussed in section 2.1 and 2.2. After that, the combination of both in section 2.3 concerning Human-Computer Interaction will be discussed. The last section contains the conclusions of the literature overview.

2.1 The Computer Perspective

2.1.1 Embodied Agents

The development and research into embodied agents is growing. Cassell et al. (2002) observe that "users' behaviours appeared natural, as though they were interacting with another person" when using MACK (Media lab Autonomous Conversational Kiosk), an embodied agent answering questions about and giving directions to the MIT Media Lab's research groups, projects and people. King and Ohya (1996) carried out an experiment with stimuli varying from simple geometric shapes to lifelike human forms, which were rated on agency and intelligence. One of their conclusions is that a human-like appearance and 'subtle behavioural displays' – such as eye blinking –have a great effect on the user's appraisal of these capabilities. Many researchers are now developing their own human-like embodied agents; some of the agents that are used in research will be introduced here.

Figure 1. Rea interacting with a user

REA (see figure 1) is a Real Estate Agent, developed by Cassell, Viljhálmsson and Bickmore (2001). REA is a life-size embodied conversational agent that can interact with users with appropriate speech, animated hand gestures, body movements, and facial expressions. With appropriate gestures she can emphasise the most important parts of her utterances. She can respond to the verbal and non-verbal behaviour of the user and knows when a user wants to talk. REA will allow the user to interrupt her, and will recognise when it is her turn to talk again. She can also give feedback to the user, like nodding her head when the user is talking and asking questions when she does not understand what the user is saying.

Figure 2. Max

The University of Bielefeld developed an embodied conversational agent called Max (see figure 2). Kopp, Gesellensetter, Krämer and Wachsmuth (2005) decided applications with an agent should be tested in a real-world situation. They conducted a study with Max at real human size in a museum. Max can spot visitors walking by and can attract their attention in order to have a conversation with them. The outcomes are suggesting, "…people are likely to use human-like communication strategies (greeting, farewell, small talk elements, insults), are cooperative in answering his questions, and try to fasten down the degree of Max's human-likeness and intelligence".

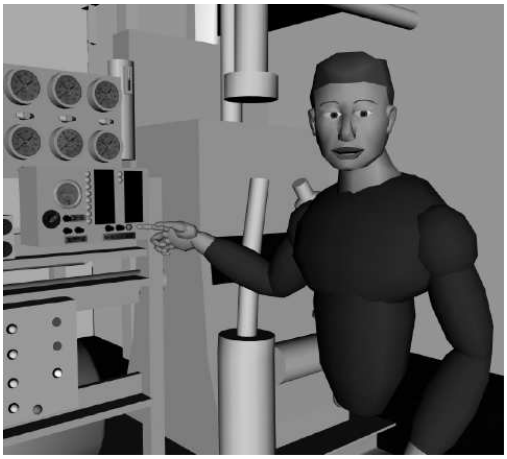


Figure 3. Steve describing a power light

Steve (Soar Training Expert for Virtual Environments) is developed by Rickel and Johnson (1997) and lives in a virtual environment. He instructs and assists students with several procedural tasks, by showing them how something is done and answering questions afterwards. He can point at objects to draw the students' attention towards these objects, so the students can ask him to explain how things work. He will not perform an action when the students cannot see it, or do not look at him, he will simply adapt his presentation.

2.1.2 Gestures

McNeill (1992) argues that gestures are an integral part of language as much as words, phrases and sentences – gesture and language are one system. Gestures are seen as part of natural communication (Noot & Ruttkay, 2005). Kendon (1994) found evidence that recipients do pay attention to gestures and that they take them in account while interpreting an utterance. Theune, Heylen and Nijholt (2005) point out: "speech is the main carrier of information, but nonverbal signals such as gestures and facial expressions also play an important role...". Bickmore and Cassell (2001) define Embodied Conversational Agents as "anthropomorphic interface agents which engage a user in real-time dialogue, using speech, gesture, gaze, and other verbal and non-verbal channels to

emulate the experience of human face-to-face interaction". They argue that the nonverbal channels can provide cues such as attentiveness, positive affect, and liking and attraction. The amount of realism might influence liking and attraction and therefore also credibility.
This means an embodied agent should be able to communicate in several ways. Cassell et al. (2001) see the use of several conversational modalities, such as speech, hand gestures and facial expression, as part of the four conversational functions, which are proposed as key to the design of embodied conversational agents. When having a conversation, people apply use several fundamental communication protocols. According to Cassell et al., the four most common protocols are:

- Content elaboration and emphasis
- Initiation and termination
- Turn-taking
- Feedback

The effect of gestures in a direction-giving situation is still under research. Some researchers believe that the use of gestures will draw attention to the more important parts of a route description. Kendon (1994) argues "gestures can make a difference in how recipients understand and retain what is conveyed in an utterance". Other researchers found no differences and even question the use of direction-giving gestures. Cohen (1980) tested the influence of gestures on task performance in a direction-giving situation, but found no differences between the participants who received the route description with route illustrating gestures and those who saw the route description without illustrating gestures.

Most route descriptions in real life contain landmarks. Sorrows and Hirtle (1999) state "landmarks are significant in one's formation of a cognitive map of both physical environments and electronic information spaces. Landmarks are defined in physical space as having key characteristics that make them recognizable and memorable in the environment". Participants will use the landmarks to confirm they are still on the right way and to identify the intersections where they have to take a turn (Lovelace, Hegarty & Montello, 1999 and Look, Kottahachchi, Laddaga & Shrobe, 2005).

Stocky (2002) asked subjects to give a route description to three distinct locations. 82% of gestures in a direction-giving situation were relative to the direction-giver's perspective. Participants used gestures to emphasize the direction, and when, for example, they said, "Turn right," they gestured to their own right rather than gesturing to the listener's right. This somewhat contradicts their speech, however, in that 95% of the directions were given in the second person narrative ("you go") rather than first person ("I go")". Most participants receiving a route description do not seem to notice this contradiction.

2.1.3 Previous Experiments on Embodiment

Users have been shown to like embodied agents and find them engaging (Takeuchi & Naito, 1995 and Koda & Maes, 1996). Most embodied agent evaluations have focused on comparing interfaces with or without an embodied agent, and on comparing agents with different visual appearances. Embodiment has proven to be very effective (Koda & Maes ,1996 and Beun, de Vos & Witteman, 2003). McBreen et al. (2000) compared the following agent embodiments: a photo of a real person with or without lip movement, a 3D talking head, and a video of a real person. They also compared a disembodied condition, where the agent was represented by a voice only. The same (human) speech soundtrack was used in all cases. Overall, the videos were rated best for likeability (friendliness, competence, naturalness) and several other aspects. It is generally assumed that for an agent to be optimally engaging and effective, it

has to be as lifelike as possible. As argued above, several studies showed that when an embodied agent seems more human in its appearance and behaviour, more human qualities are accredited to it. As mentioned before, a comparison between a full-bodied human guide and an embodied agent might be a useful addition.

2.2 The Human Perspective

2.2.1 Communication and credibility

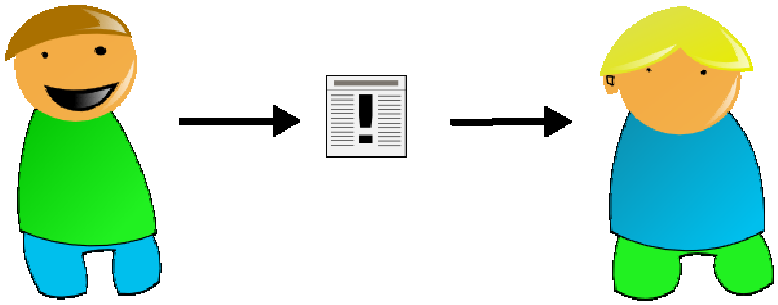


Figure. 4. A simple communication model from www.wikipedia.com

Kepplinger (1991) describes how each (verbal) message is influenced by the message itself, the sender and the receiver of the message. The characteristics of the receiver, his or her sensitivity for non-verbal behaviour and the way something is presented are an indication of how a message will be received. Kepplinger argues that the interpretation of a message is related to how the receiver judges the credibility of both the message and the sender. The way the message is presented is also an influencing factor on credibility. O'Keefe (1990) states that competence and trustworthiness of the message and the sender are important ways in which the opinion of a receiver can be influenced.

Reardon (1991) sees expertise as an important part of credibility. Ruttkay, Doorman and Noot (2002) see trustworthiness of the sender as an important part of engagement. Some authors use the word likeability (McBreen et al., 2000; Koda & Maes, 1996) but it seems that credibility, trustworthiness, competence and liking all influence the way receiver of a message judges it.

The last factor, which is found in literature that may influence competence, and thus credibility, is dominance. Reeves and Nass (1996) see dominance as the most important personality trait, which is linked with sympathy. Bickmore, Caruso and Clough-Gorr (2005) conducted a study with an agent as a personal trainer who would ask users about their exercise plans and if people actually did exercise. The intention of the system is that people change their health behaviour and start exercising. This system with integrated embodied agent is called Fit Track. The relationship-building behaviours of the embodied agent included a warm facial expression and a relaxed body posture. In the condition with these relationship-building behaviours disabled, participants had lower scores on measurements of liking and desire to continue working with the agent. The health and exercise behaviour itself did not differ significantly.

2.2.2 Information Processing

There are two basic views concerning information processing in a multi-modal environment. It is possible that people can get distracted when information is given by something they don't know: an embodied agent. This is especially true

when that agent performs gestures and changes posture. There is some research that suggests these actions may actually distract people (Walker, Sproull & Subramani, 1994). On the other hand, authors such as Mayer and Moreno (1999) suggest that presenting information both visually and verbally may stimulate the cognitive capabilities.

Lester, Kahler, Barlow, Stone and Bhogal (1997) argue that information provided by agents instead of text-only is more actively processed. They call this the persona effect, which has been tested by several other researchers. A study by Mulken, Andre and Müller (1998) suggest: " …the presence of Persona neither has a positive nor a negative effect on comprehension and recall performance, and that the type of information does not seem to play a role in this. However, Persona does have a positive effect on the subject's impression of the presentation: even its mere presence causes presentations to be experienced as less difficult and more entertaining. In addition, tests following presentations by Persona are experienced as less difficult".

2.2.3 Subjective Experience

Dehn and Mulken (2000) describe several dimensions, which are used to measure the users' attitude towards a certain interface. Again, believability and likeability are mentioned. Entertainment, comfortability and usability are some of the other dimensions. The more people enjoy working with a certain program, the better the results probably are. Depending on the type of experiment and task, dimensions regarding subjective experience can be made operational. Ruttkay et al. (2002) see satisfaction with the agent and the preference to use the agent instead of traditional material as one of the aspects, which are important to evaluate a character.

2.3 Human Computer Interaction

Nass, Steuer and Hendriksen (1994) presented a new experimental paradigm for the study of human-computer interaction. With five experiments they provided evidence that individuals' interactions with computers are fundamentally social. The outcomes were that people treated the computer as polite as another human being. Participants were amendable to flattery and reacted the same way as if a real human was flattering them. Also gender stereotypes were found in how participants treated the computer. Striking is the suggestion that the participants did not show this social behaviour because they thought computers were human-like or they were thinking about the computer programmer, but these reactions came naturally. As Nass, Steuer and Tauber state: "These social responses are not a function of deficiency, or of sociological or psychological dysfunction, but rather are natural responses to social situations. Furthermore, these social responses are easy to generate, commonplace, and incurable". This is very important for the development of embodied agents. The outcomes of these experiments suggest that people who are interacting with embodied agents might very easily show social behaviour towards the embodied agent, especially if the agent is showing social behaviour itself.

The Computers-are-social-actors studies (CASA) find more and more proof for these assumptions. For example, a study conducted by Lee and Nass (1998) suggested that people's social responses to media affect their feelings of social presence. Reeves and Nass (1996) have shown that people respond to computers and other media like they respond to people, treating them as social actors and attributing them with personality. Computers are ever less viewed as tools and ever more as partners or assistants to whom tasks may be delegated (Rist, Andre & Baldes, 2003).

3 First Experiment

3.1 Introduction

Previous research is still ambiguous about the effect that an embodied agent in an interface has. Several studies with an embodied agent compared to a real face, a cartoon face or no face at all, have as of yet not been able to make clear what can be expected. This experiment will compare an embodied agent with a mediated real person and measure user subjective experience and task performance. As mentioned before the design will be partly exploratory, because the intended dependent variables have not been used before to compare an embodied agent with a real person. The effect of gestures will be investigated as well, by presenting the same information in the same way with or without gestures and measuring subjective user experience and task performance. Because of the expected differences between the condition with and without gestures, this part is not exploratory. Therefore, only hypotheses considering the use of gestures will be formulated.

In this chapter the objectives and hypothesis can be found in section 3.2 and 3.3. After this, the different design aspects and methodology are discussed in section 3.4 till 3.9.

3.2 Objectives Student Experiment

The following four questions are to be answered in this experiment.

- Do gestures have an impact on subjective experience with either a human guide or an embodied agent giving a route description?
- Do gestures have an impact on task performance with either a human guide or an embodied agent giving a route description?
- What are the differences in subjective user experience between the participants who saw the human guide and those who saw the embodied agent?
- What are the differences in task performance between the participants who saw the human guide and those who saw the embodied agent?

3.3 Hypotheses

The hypotheses concerning the differences between the agent with or without gestures are as follows:

H1: Participants who saw the route description given by the agent with gestures will trust the guide more than participants who saw the route description without gestures.
H2: Participants who saw the route description given by the agent with gestures will evaluate the presentation style more positive than participants who saw the route description by the agent without gestures.
H3: Participants who saw the route description given by the agent with gestures will evaluate the quality of the route description more positive than participants who saw the route description without gestures.
H4: Participants who saw the route description given by the agent with gestures will remember more of the route description than participants who saw the route description without gestures.

There are no hypotheses formulated concerning the comparison between the human guide and the embodied agent, because this part of the experiment was exploratory. The reason for this lies in the lack of earlier studies' comparing a full-body embodied agent with a human guide. Especially in the domain of route descriptions, there is not enough evidence to state well-founded hypotheses.

3.4 Design

There are four conditions in this experiment. Subjects were initially presented with a route description with gestures given by a human guide, recorded on video (condition 1) or by an embodied agent (condition 2), or were presented with a route description without gestures given by a human guide, recorded on video (condition 3) or by an embodied agent (condition 4). Methodological standards were met by making the human guide and the agent guide as similar to each other as possible, only varying the dimension under investigation: i.e., the synthetic versus human appearance of the guide. How this is achieved is discussed in section 3.7. For both versions of the guide we used the name Laura: the actual name of the human guide. After the participants had watched the route description by the human or the agent guide, they were asked several series of questions, measuring among other things the guide trustworthiness and presentation style. Then they were shown a movie with the same route description, but this time presented by the version of the guide they had not seen yet. After this second movie, when the participants had seen both agent and human guide, they were asked to indicate which version of the guide they preferred.

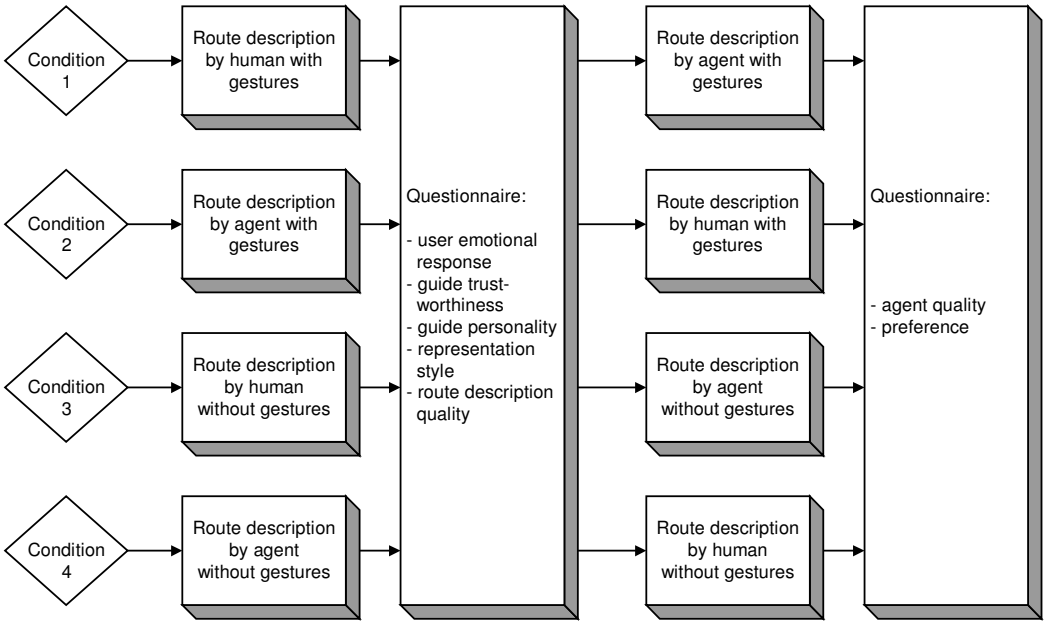


Figure. 5. Graphical representation of first experiment.

The outcomes of an experiment by Stocky (2002, see section 2.1.2) is the reason the guides in our experiment will make gestures from their own perspective. For example, the guides will point towards the viewers' right when explaining a left turn.

The use of landmarks (see section 2.1.2) in this experiment will be dual, on one hand they will be of use for the participants, since they might use them to remember the route. On the other hand the landmarks will be part of the method to measure how and how much people remembered the route description.

To make sure people will have to make some effort to remember everything, but are able to remember the whole route description, the amount of turns and

landmarks has to be considered carefully. According to Miller (1956), who discovered people can handle about 7, plus or minus two, chunks of information in working memory, seven or eight turns should be the right amount. The reason for this is that the participants will be all higher educated people and one might expect them to be trained in remembering bigger amounts of information. After a small pre-test among 3 males and 3 females the final version of the route description was prepared, with six turns, without two "straight on" indications (see appendix A).

Figure. 6. The human guide (left) and the agent (right)

3.5 Dependent Variables

After having seen the route description given by either the agent or the human guide (see figure 6), with or without gestures, the participants in the experiment answered several questions. In this section is explained how these questions were grouped. All questions were measured on a nine-point scale, except the question about preference. The experimental design and task, a route description, have an exploratory character; therefore new dimensions have to be made operational.

3.5.1 Guide Trustworthiness

In the literature research several factors concerning credibility, likeability and trustworthiness are found (see chapter 2). Because there is not yet general accepted term for it, in this thesis the term trustworthiness as a general name will be used. The literature indicates several important influencing factors. Because of the exploratory character of this experiment, the effect of grouping of all these factors in one scale has to prove itself. Based on the literature research, guide trustworthiness was measured in terms of seven items: expertise, believability, realism, reliability, friendliness, sympathy, and dominance. This is a moderate reliable index, alpha = .66.

3.5.2 Presentation Style

The presentation style, the way the guide presents the route, should contain multiple questions where participants can give their opinion about user

experience. Presentation style regards the way the guide presented the route. Twelve nine-point scale items formed this index: good-bad, pleasant-unpleasant, polite-impolite, natural-artificial, flowing-clumsy, relaxed-tense, energetic-lethargic, dynamic-static, accurate-inaccurate, calm-excited, exuberant-apathetic, and interested-disinterested. This is a very reliable index, alpha = .78.

3.5.3 Route Description Quality

Route description quality should measure the way participants feel about the message itself and should contain questions about how easy or structured the route description was. Therefore, this index is comprised of eight nine-point scale items: concise-tedious, simple-complex, easy-difficult, interesting-boring, structured-unstructured, useful-useless, clear-unclear, and comprehensible-incomprehensible. This is a very reliable index as well, alpha = .82.

3.5.4 Task Performance

This was measured by asking the participants to write down the route they just heard in their own words, naming as many landmarks and turns as they could. The participants would receive points for the amount of landmarks they could remember and the amount of turns they remembered correctly. The author and two independent people awarded each answer with the amount of points they judged as appropriate. Any discrepancies were considered and argued before the final marks were granted. These final marks are used for data analysis. Goal of this variable was not to determine merely if people remembered the exact route, but also how much information participants could recall overall. This is the reason the landmarks did not count as part of the route description, but more as an overall test of memory. Due to the fact that SPSS cannot calculate with zero, each participant received 1 point extra on "landmarks" and 1 point extra on "turns". The maximum thus became seven on landmarks and also seven on turns, the overall maximum was 14.

3.5.5 Preference

Satisfaction will be measured by asking preference for the agent or the human guide for this task. People may choose the agent for this task if they were sufficient satisfied with the way the information was presented. Preference was determined using one simple question: "Which of the two do you prefer: virtual person (agent) or real person (video)?"

Besides these five above-mentioned topics, questions about the personality of the guide and agent quality were asked. Participants were also asked for a further explanation in their own words when they finished a group of questions. The whole list of questions is placed in appendix B. The outcomes of these remaining topics were not relevant to answer the main questions of this thesis, and therefore were placed in appendix C.

3.6 Participants

Subjects in the first experiment were 146 undergraduate students from different faculties such as Computer Science and Psychology. They were all following a course in Media Psychology and were rewarded with bonus points to participate. Subjects were randomly assigned to one of the conditions, with age and gender approximately balanced across conditions. The average age of the participants was 21, between 18 and 27; 60 % of the participants were female.

3.7 Procedure

The experiment was performed in a Web environment. After a short instruction, the participants started the questionnaire on their computer. The short films, about a minute each, with the route presentations were integrated into the

questionnaire. The participants could not see the films twice, nor could they go back to see or change their previous answers.
Depending on the group they were assigned to, participants would start with watching a film with either the agent or the human guide presenting the route. Both films started with the guide introducing herself: "Hi, I'm Laura." She would then thank them for their cooperation and explain she was going to give them a route description. This way the participants could get used to the voice and the appearance of the guide.

3.8 Material

For the agent the Living Actor™ technology from Cantoche[1] was used. To make the agent as human-like as possible, an agent was selected that looked realistic rather than cartoon-like and had a large repertoire of gestures. The agent that best met these requirements happened to be female, the Cantoche character 'Julie'. To reduce the differences between the agent and the human guide as much as possible, someone who looked like the agent was asked to play the role of the human guide, dressed in exactly the same clothes as the agent.

3.9 Technical details

3.9.1 The Embodiment

The agent is a realistic, 3D, full body female model. Her body had the right proportions and moved in a human-like way. When she spoke, her body would lean slightly forward, and when making a gesture, her whole upper body moved a bit to that side, instead of just her arm stretching (see figure 7).

Figure 7. The phases in the arm movement

Recorded audio was used, so that both versions had exactly the same sound. This way, the speech was spontaneous instead of sterile and avoided possible negative effects of a synthetic voice (Sproull, Subramani, Kiesler, Walker & Waters, 1996). The face was kept neutral, although facial expressions were possible. The reason for this is that the human guide had a neutral expression and an effort was made to make both conditions as similar as possible. Full lip synchronisation was not completely possible; this is not a part of the software. The lips were moving during each utterance, but not exactly in the right way. This was not considered a big problem because the agent was full bodied. Her eyes blinked and her head moved in a natural way, and her body posture changed every now and then. A full overview of all used gestures can be found in appendix D. Cantoche designed the agent for a presentation task. She was chosen for this experiment because she provided the impression of a calm and friendly person.

[1] http://www.cantoche.com

3.9.2 Creating the Films

The films of the route presentations were created as follows. First, we made a video recording of the human guide as she spontaneously described the route. Then we scripted the agent to simulate the gestures that had been made by the human guide as closely as possible, e.g., pointing left and right. Because of limitations in the gesture repertoire of the agent, this simulation deviated in a few respects from the original recording. Therefore we made a final recording of the human guide as she was describing the route, this time mimicking the agent. The human actor was not asked to imitate the agent in every behavioural detail, only at the more global level of gestures. The use of different gestures would have made the presentations of the guides too dissimilar to allow for a reliable comparison, but we considered the smaller unconscious movements such as blinking and head movements as part of what made the human guide appear human and the agent guide synthetic.

Finally, we added the speech of the human guide to the agent, synchronized the agent's gestures and lip movements with the speech, and created a white background for both movies. Overall, they acted and looked similar, the main difference being that one guide was human and the other an embodied agent.

4 Results First Experiment

4.1 Introduction

The differences between an agent with and without gestures are the most relevant in order to evaluate the hypotheses. For the exploratory research the comparison between the human guide and agent are the most important. Therefore, the results in this chapter are divided into three sections. First the results of the participants judging the agent with and without gestures will be given in section 4.2. After this, the results of the comparison between the human guide and agent will be described in section 4.3, divided in a subsection with gestures and without gestures. Section 4.4 will contain the assessment of the hypotheses. The full results, also containing the comparison between the human guide with and without gestures can be found in appendix C.

With the exception of the question about preference, where the participants had to indicate whether they preferred the human or the agent guide and task performance, a nine-point scale was used for all questions. In the results given below, the high end of each scale is given. The SPSS program is used (using T tests) to compare the mean of the scores on all dependent variables as described in section 3.4. This test compares the mean of each item or index for both conditions. The t-value indicates the difference between the two conditions. Differences where $p < .05$ will be treated as significant. The last column represents the mean difference between the two conditions (MD).

The agent with gestures is expected to score better on subjective user experience and task performance, therefore a one-sided test will be performed to compare the agent with and without gestures. The second part of the experiment, a comparison between the human guide and the embodied agent is exploratory. Therefore, all differences are tested two tailed: no expectations or hypotheses if the agent or the human guide would perform better were formulated.

Because of the high reliability of the indexes, the main effects were tested first. This means the items that formed an index were joined together and measured as one item, which is called a main effect. The dependent variables will each be discussed in a new sub section where first the results about the main effect will be revealed. After that the separate items are shown in a table.

4.2 Agent With and Without Gestures Compared

4.2.1 Guide Trustworthiness

Overall, there was a significant main effect with regard to the guide's trustworthiness ($t = .58$, $p < 0.01$), the participants found the agent with gestures trustworthier. When the separate items were compared it showed that the agent with gestures scored higher on most items, except that the agent with gestures was found slightly, not significantly more reliable. All separate items are shown in table 1.

Table 1. Separate items for guide trustworthiness

	With	Without	MD
Competent	6.03	5.16	0.87 **
Convincing	6.45	6.18	0.27
Realistic	5.42	4.34	1.08 **
Reliable	6.05	6.11	0.06
Friendly	6.71	5.75	0.96 **
Likeable	6.21	5.75	0.46
Dominant	5.47	4.95	0.52 *

* p<.05, ** p<.01

4.2.2 Presentation Style

There was no significant main effect for this index; however the agent with gestures scored higher on almost every item that measured presentation style, as shown in table 2.

Table 2. Separate items for presentation style

	With	Without	MD
Good	4.79	4.30	0.49
Pleasant	4.92	4.02	0.90 **
Polite	6.39	6.52	0.13
Natural	5.47	4.82	0.66
Flowing	5.82	4.77	1.04 **
Relaxed	6.05	5.86	0.19
Energetic	5.29	4.50	0.79
Dynamic	4.47	3.23	1.25 **
Accurate	6.42	6.75	0.33 **
Exuberant	4.26	4.14	0.13
Calm	6.87	6.98	0.14
Interested	5.53	4.64	0.89 **

** p<.01

4.2.3 Route Description Quality

There was no significant main effect for this index either. Table 3 shows that the agent with gestures scored higher on seven out of eight items regarding the route description quality. None of the items scored significantly different.

Table 3. Separate items for route description quality

	With	Without	MD
Concise	4.05	3.80	0.26
Simple	3.82	3.57	0.25
Easy	3.97	3.66	0.31
Interesting	3.95	3.43	0.52
Structured	5.92	5.64	0.28
Useful	4.45	4.55	0.09
Clear	5.34	4.95	0.39
Comprehensible	5.63	5.30	0.34

4.2.4 Task Performance

The amount of information participants could remember did not differ significantly between both conditions, but surprisingly, the participants who saw the agent without gestures scored slightly better.

Table 4. The amount of turns, landmarks and total recalled

	With	Without	MD
Landmarks	3.52	3.70	0.18
Turns	4.05	4.50	0.45
Total	7.78	8.20	0.42

4.2.5 Preference

There is no significant difference in preference[2], although 30 % of participants who saw the agent with gestures preferred the agent, against 20 % who saw the agent without gestures.

Table 5. Preference for human guide or agent

	With	Without
Preference agent	30%	20%
Preference human guide	70%	80%

4.3 Agent and Human Guide With Gestures Compared

4.3.1 Guide Trustworthiness

There was no significant main effect for this index. The participants felt that the agent was more competent than the human guide ($t = 0.98$, $p<0.05$). The scores on the other items concerning trustworthiness did not differ significantly between the human guide and the agent. Reliability of the guide was rated exactly the same for both guides. Striking is the notion that agent was found only slightly less realistic than the human guide.

Table 6. Separate items for guide trustworthiness

	Agent	Human	MD
Competent	6.03	5.05	0.98 *
Convincing	4.45	6.60	0.15
Realistic	5.42	5.98	0.55
Reliable	6.05	6.05	0.02
Friendly	6.71	6.43	0.29
Likeable	6.21	6.18	0.03
Dominant	5.47	5.53	0.05

* p<.05

[2] As described in section 3.4, the participants saw either the agent or the human guide and answered the questions concerning subjective user experience and task performance. After this, they saw the other guide presenting the same route and were asked about their preference.

4.3.2 Presentation Style

Overall, there was a significant main effect with regard to the presentation style index ($t = 0.39$, $p<0.05$), such that participants found the presentation style of the agent better than the style of the human guide. Table 7 shows all the separate items from this index. The presentation style of the agent was seen as significantly more relaxed, dynamic and interested than the presentation style of the human guide. A few of the remarks were: " very much like a human" and " neutral, but very accurate and polite". The real person was found " too boring" and "pretended".

Table 7. Separate items for presentation style

	Agent	Human	MD
Good	4.79	4.70	0.09
Pleasant	4.92	4.70	0.22
Polite	6.39	6.23	0.17
Natural	5.47	4.88	0.60
Flowing	5.82	5.28	0.54
Relaxed	6.05	5.35	0.70 *
Energetic	5.29	4.75	0.54
Dynamic	4.47	3.36	0.85 *
Accurate	6.42	3.38	0.04
Exuberant	4.26	4.03	0.24
Calm	3.16	3.13	0.03
Interested	5.53	4.83	0.70 *

* p<.05

4.3.3 Route Description Quality

There was a significant main effect with regard to route description quality ($t = 0.50$, $p<0.05$); the participants found the route description better when the agent presented it. The agent scored higher on every single item, although only one item is significant: the route description given by the agent was considered significantly less boring than the description given by the human guide.

Table 8. Separate items for route description quality

	Agent	Human	MD
Concise	4.05	3.30	0.75
Simple	3.82	3.33	0.49
Easy	3.97	3.58	0.40
Interesting	3.95	3.00	0.95 **
Structured	5.92	5.55	0.37
Useful	4.45	4.08	0.37
Clear	5.34	5.00	0.34
Comprehensible	5.63	5.28	0.36

** p<.01

4.3.4 Task Performance

There were no significant differences in recall between the participants who saw the agent with gestures and those who saw the human guide with gestures explaining the route.

Table 9. The amount of turns, landmarks and total recalled

	Agent	Human	MD
Landmarks	4.05	4.22	0.17
Turns	3.52	3.87	0.34
Total	7.57	8.10	0.52

4.3.5 Preference

Participants who saw the human guide first and filled in the questionnaire concerning user experience about the human guide, have a significantly higher preference for the agent.

Table 10. Preference for human guide or agent

	Agent first	Human first
Preference agent	68%	38% **
Preference human guide	32%	62% **

** p<.01

4.4 Agent and Human Guide Without Gestures Compared

There were not many significant differences between the two groups of participants who saw the route description without gestures. Therefore, an overall table (see table 11) containing the items that differed significantly can be found below. The human guide without gestures scored overall slightly better than the agent without gestures. The fact that the agent without gestures scored much lower on realism than the human guide without gestures is remarkable, since the agent with gestures did not score significantly lower than the human guide with gestures. A big difference can be found in task performance (see table 12): the participants who saw the agent without gestures remembered more than those who saw the human guide without gestures. Also notable are the results for preference. Participants who saw the human guide first and filled in the questionnaire concerning user experience about the human guide, have a significantly higher preference for the agent.

Table 11. An overall table with all significantly different items

	Agent	Human	Mean difference
Guide trustworthiness			
Realistic	4.34	6.43	2.09 **
Friendly	5.75	6.64	0.89 *
Dominant	4.95	5.57	0.63 **
Presentation style			
Good	4.30	5.26	0.97 *
Interested	4.64	5.43	0.79 *
Route description quality			
No significant differences			
Task performance			
Turns	4.50	3.76 *	0.74 *
Total (Turns + landmarks)	8.20	7.00 *	1.20 *

* p <.05, ** p<.01

Table 12. Preference for human or agent

	Agent first	Human first
Preference agent	20%	40% **
Preference human guide	80%	60% **

** p<.01

It is very remarkable that although participants felt the human guide without gestures was trustworthier and gave a better route description, the participants who saw the agent without gestures remembered more. The gap between subjective experience and task performance seems big.

4.5 Evaluation of Hypotheses

H1: Participants who saw the route description given by the agent with gestures will trust the guide more than participants who saw the route description without gestures.

There is statistical proof to support this hypothesis; there was a significant main effect, such that participants found the agent with gestures trustworthier. This hypothesis is accepted.

H2: Participants who saw the route description given by the agent with gestures will evaluate the presentation style more positive than participants who saw a route description by the agent without gestures.

This hypothesis is confirmed; the agent with gestures scored higher on every item, most of them significant.

H3: Participants who saw the route description given by the agent with gestures will evaluate the quality of the route description more positive than participants who saw a route description by the agent without gestures.

There is no statistical proof to support this hypothesis; none of the items scored significantly different.

H4: Participants who saw the route description given by the agent with gestures will remember more of the route description than participants who saw a route description without gestures.

There is no support for this hypothesis; the participants who saw the agent with gestures did not remember more landmarks or turns than the participants who saw the same agent without gestures.

5 Discussion First Experiment

The four main questions for this experiment were:

- Do gestures have an impact on subjective experience with either a human guide or an embodied agent giving a route description?
- Do gestures have an impact on task performance with either a human guide or an embodied agent giving a route description?
- What are the differences in subjective user experience between the participants who saw the human guide and those who saw the embodied agent?
- What are the differences in task performance between the participants who saw the human guide and those who saw the embodied agent?

The first two questions will be discussed in section 5.1, the last two questions will be discussed in the second section of this chapter. After this, some options for further research will be given in section 5.3.

5.1 With or Without Gestures

As the data suggests, gestures do have an impact on subjective user experience. The agent with gestures scored significantly higher on trustworthiness and presentation style, and was found to be much more realistic. Remarkable is the fact that the agent without gestures scored significantly higher on accuracy. This might be an indication that the gestures that were made from the direction-giving perspective were a bit confusing. Most items on route description quality were rated higher by participants who saw the agent with gestures, however, this was not significant. The agent with gestures scored also higher on preference. So, looking at the subjective experience, participants enjoyed the agent with gestures more than the agent without gestures.

However, it seems that gestures were not particular helpful in task performance. On the contrary, the participants, who saw the route description without gestures, scored slightly higher. This is quite interesting, because this leads to the assumption that, depending on what designers want to achieve, a choice between task performance and trust has to be made. It is possible that participants found the agent to such extent boring; they stopped looking and just listened. In that case, there is no need to put an embodied agent in an interface. The fact that participants might have noticed the gestures contradicted speech because of the perspective might have influenced the results.

Overall, users did prefer the agent with gestures to the agent without, and did trust the agent with gestures significantly more. Also presentation style and route description were rated higher. This outcome can be compared to the outcomes of a study by Mulken et al. (1998, see section 3.2.3). The subjective experience in this experiment was rated higher in the condition with agent, but there was no difference in recall with the condition without agent. So, it is possible that an agent without gestures is somewhat comparable to a condition without any agent at all. The fact that there were so many significant differences in the subjective experience of users is very encouraging for designers that agents with gestures are indeed judged more positive.

5.2 Human Guide or Embodied Agent

In this section, the focus will be on the comparison of a human guide and an embodied agent with gestures. The first reason for this is that the agent with gestures scored higher on subjective user experience than the agent without gestures. The agent with gestures was also preferred to the agent without

gestures. The second reason is the fact that the differences between the human guide and agent without gestures were very small, whereas the differences between the human guide and agent with gestures are much bigger.

Overall, the embodied agent with gestures received more positive ratings than the video recording of the human guide. There are several factors that may help explain these results. First of all, the agent was of good quality (as confirmed by the participants' ratings), having a realistic appearance, a natural human voice, and quite natural movements that included not only gestures but also more subtle behaviours such as blinking, head movements and posture shifts. All in all, despite being an animation, the agent appeared fairly realistic and this may have led to more positive judgments than were found in previous experiments.

At the same time, there are also some factors that may have negatively influenced the scores of the human guide. One of these is the fact that in the final version of our recording, she was acting instead of behaving spontaneously. She had to recreate her earlier spontaneous description, this time keeping in mind which gestures the agent could and could not make. For this reason she may have come across as less self-assured and less relaxed, and thus also as less competent.

The fact that the agent was trusted significantly more and scored higher on all subjective experience scales is very promising. This might support the suggestion that agents have strong potential as a guide, tutor or advisor. A striking result was that the comparison in presentation style turned out in favour of the agent rather than the human guide. The quality of the route description given by the agent was also perceived more positively on every dimension. Even though these results may have been partially influenced by the design of the experiment, this is encouraging news for developers of interface agents. The limitations and possible effects of the experimental design will be further discussed in chapter 10.

The fact that there is no significant difference in task performance between the participants who saw the human guide and those who saw the agent may be interpreted as good news as well. This means people can obtain information from either an embodied agent or a human guide and will be able to remember the same amount of information.

5.3 Further Research

The main factor that may have caused a preference for the agent is that the participants in this experiment were young people, who are generally open to new technology and may appreciate a novelty, like a virtual character, more than a well-known phenomenon like a real person. With an older age group, the outcomes could well have been different. A comparison between an older group of participants and the original group of students might make clear if age is of influence on the results.

6 Literature Overview: Older adults and New Technology

6.1 Introduction

Because the outcomes of the first experiment might be caused in part by the fact that the participants were all students, a second experiment with older participants was conducted. In this chapter a literature overview concerning older adults and new technology is given, to investigate how the opinion of other researchers about this subject is.

6.2 Older adults and New Technology

Probably most younger people will have the experience of their parents or grandparents asking them questions about their computer. This might be the reason that a stereotypical view that most older adults are unwilling to interact with new technologies exists. Research into older adults and new technologies argues against this view (Rogers, 2000 and Williamson, Bow & Wale, 1997). Nowadays, not many people see chronological age as a predictor of new technology competency and use (Hazzlewood, 2002), there are many older users who work with computers and know how to deal with problems. However, recent research discovered that older people, when asked to perform real-world computer tasks such as data entry or database inquiry, perceived less comfort, efficacy, and control over computers than did the other participants (Czaja & Lee, 2003). This means that the stereotypical view of older people struggling with their computer may be true after all. An explanation of these different outcomes might be found in the time people spend with a computer and the software. Multiple researchers found that with training and time, the initial negative attitude towards computers changes (Kelley & Charness, 1995 and Morris, 1994). Learning computer applications and their interfaces is possible but takes significantly longer and is harder for older people (Baldi, 1997). Williamson et al. (1997) state: "we live in a society that devalues the ageing experience. The expression 'You can't teach an old dog new tricks' sums up our attitudes towards older adults, many of whom have absorbed this myth into their everyday lives and firmly believe that they are too old to learn new things". Changing this attitude into a more positive view of their own capabilities older people may very well have no problems with computers at all. This positive view might be reached by giving older adults a different training in computers, because they learn in a different way than younger people.

6.3 Older Adults and Memory

The effects of age become noticeable from the mid forties onward, normal aging (excluding pathological conditions such as Alzheimer's disease) produces different degrees of impairment on the different forms of memory. Age gives a slight decline in the number of items which can be held in short term memory (Hawthorn, 2000). Memory performance, that requires the formation of new connections, such as new facts, is relatively weak in older age. Older adults exhibit major declines in episodic memory, performing more poorly on laboratory tests that involve episodic recall or recognition of virtually any stimuli, for example, single words or prose passages, spatial locations, pictures, faces, and activities (Burke & MacKay, 1997). Hawthorn (2000) states: "in the light of the material on difficulty in speech planning, word retrieval and conforming to a precise vocabulary, use of a verbal command language is likely to be more difficult for older users".

6.4 Older Adults and Agents

As mentioned before, there is not much research on embodied agents and older people. Bickmore et al. (2005) state that one of the few experiments in this field is conducted by Smith (2000). Smith (2000) conducted an experiment with an embodied agent and 15 grandmothers between 55 and 65 years old. The agent looked like a 6-year old child. The system, called GrandChair, made an attempt to let older users talk more and longer with the agent. The agent changed postures and would encourage the user on the appropriate moments to tell more. The agent was found to elicit significantly more and longer stories than a text-prompt control condition. Bickmore et al. (2005) conducted an experiment among 21 older adults, using the same relational agent as in the FitTrack experiment among students (see section 2.2.1). The 10 participants in the experimental group used the system during 60 days. The end result was, that most participants in the experimental group found the agent to be very friendly and reported they would like to continue using the system. "Results indicate the agent was accepted and liked, and was significantly more efficacious at increasing physical activity (daily steps walked) than the control. The fact that participants in this experiment could use the system for 60 days might have a big influence on the results. As mentioned earlier, older people tend to have a different attitude towards computers and new technology than younger people, but these differences disappear when people get appropriate training and can spend more time with the computer.

7 Second Experiment

7.1 Introduction

As the conclusions of the first experiment indicate, an experiment amongst senior participants may be a very useful supplement to the first experiment. As seen in the literature study in chapter six, older adults may respond very differently to an embodied agent, especially because the time they spend with the agent is very short. This experiment may help to give the main question of this thesis a broader perspective, because many experiments are done with students, seldom with older adults.

The literature research made clear some different outcomes might be expected. For example, an overall lower score on task performance can be expected when executing the same experiment among older participants. Older adults can be expected to respond differently compared to students, especially in this experiment: they see the agent only for a minute. Senior participants can also be expected to remember less of the route description, which may influence their mood during the experiment. But it will not be possible to show them the route description twice; this will influence the validity of the design.

7.2 Objectives Senior Experiment

In order to answer the main question, several sub questions were formulated. The questions about the use of gestures were answered in the first experiment. In this experiment, the possible influence of age is under investigation. Therefore, the following five questions are to be answered in this experiment.

- What are the differences in subjective user experience between the older and younger participants who received a route description by an embodied agent?
- What are the differences in task performance between the older and younger participants who received a route description by an embodied agent?
- What are the differences in preference between the older and younger participants who received a route description by an embodied agent?
- Will senior participants who saw a route description given by the human guide judge their subjective experience different than the senior participants who saw a route description given by the agent?
- Will senior participants who saw a route description given by the human guide remember more than the senior participants who saw a route description given by the agent?

7.3 Hypotheses

Based on the above-mentioned literature the following hypotheses were formed to be tested in the second experiment.

The hypotheses concerning the differences between the senior participants and the student participants:

H1: Senior participants who saw a route description given by the agent will trust the guide less than student participants who saw a route description given by the agent.
H2: Senior participants who saw a route description given by the agent will evaluate the presentational style less positively than student participants who saw a route description given by the agent.

H3: Senior participants who saw a route description given by the agent will evaluate the quality of the route description less positively than student participants who saw a route description given by the agent.
H4: A higher percentage of senior participants than younger participants will prefer the real person instead of the agent for this specific direction-giving task.

The hypotheses concerning the differences between the agent and the human guide, perceived by the senior participants:

H5: Senior participants who saw a route description given by the human guide will trust the guide more than the senior participants who saw a route description given by the agent.
H6: Senior participants who saw a route description given by the human guide will evaluate the presentational style more positively than the senior participants who saw a route description given by the agent
H7: Senior participants who saw a route description given by the human guide will evaluate the quality of the route description more positively than the senior participants who saw a route description given by the agent.
H8: Senior participants who saw a route description given by the human guide will remember more of the route description than the senior participants who saw a route description given by the agent.

7.4 Design

The two experiments have to be very similar in order to give well-grounded answer on the questions above. Therefore the experimental design is on most facets, such as films and questionnaire, exactly the same. Changes to the initial experiment will be outlined below.

Based on the results of the first experiment (see section 4.2) the second experiment had only two conditions. The two conditions without gestures were removed, because the agents with gestures scored higher on trustworthiness and the other dependent variables. The goal of the second experiment was to investigate if age is an influencing factor. Subjects in this experiment were initially presented with a route description with gestures given by a human guide, recorded on video (condition 1) or by an embodied agent (condition 2).

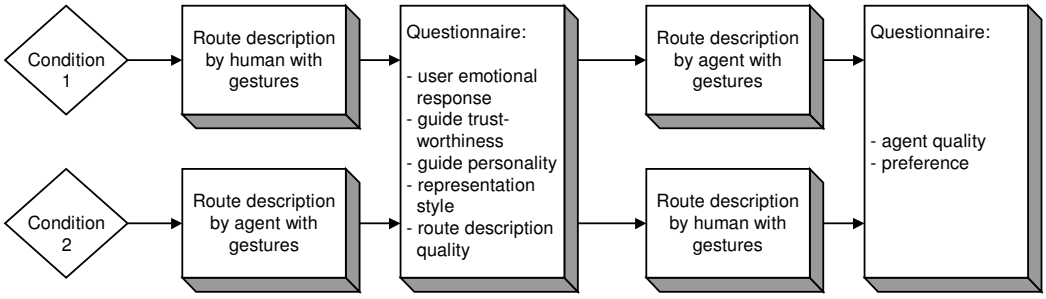


Figure. 8. Graphical representation of second experiment.

7.5 Participants

Subjects in the second experiment were partly teachers from a technical college in Arnhem, the Netherlands, partly older people asked at random for cooperation in a public library. Because the participants from the first experiment were undergraduate students, participants could only join if they were above 40, and had at least one year higher education. These requirements were set so there would be at least an age difference of one generation between both groups of participants and possible differences could not result from differences in

educational level. Subjects were randomly assigned to one of the conditions, with age and gender approximately balanced across conditions. The average age of the participants was 51, between 40 and 64; 60% of the participants were male (N = 49).

To make the second experiment comparable to the first experiment, a weighting variable was used. Females were weighed heavier than males and because the participants were not perfectly spread across conditions some males weighed heavier as well. After weighting, the amount of cases was exactly identical to the amount of cases in the first experiment. In table 13 the weighting values can be found.

Table 13. The weighting values used in second experiment

	Males	Females
Condition 1	1.00	2.10
Condition 2	1.30	2.60

7.6 Dependent Variables

7.6.1 Guide Trustworthiness

Guide trustworthiness was measured the same way as with the students. In this experiment it turned out to be a moderately reliable index, alpha = .69.

7.6.2 Presentation Style

Presentation style was measured in the same way as with the students. In this experiment it turned out to be a very reliable index as well, with an alpha of .84.

7.6.3 Route Description Quality

Route description quality was measured in the same way as with the student experiment. Again, in this experiment it turned out to be a very reliable index, alpha = .81.

8 Results Second Experiment

8.1 Introduction

To test the hypotheses about possible differences between the group of students and the elder group of participants, the two data files were merged. Again, T tests were performed to compare the means on the separate items. When the results were analysed, there were many differences between the younger group of participants and the older group. First the differences between the students and seniors who saw the agent will be discussed in section 8.2. Section 8.3 contains the differences between the senior participants who saw the agent, and the senior participants who saw the human guide. The full results can be found in appendix E. The hypotheses will be discussed in the third and final section of this chapter.

8.2 Students and Seniors who Saw the Agent

Despite the high reliability of the indexes, no significant main effects are found. Therefore, in this section, all separate items will be discussed.

8.2.1 Guide Trustworthiness

The students scored higher on most items. There were significant differences on three out of seven items.

Table 14. Separate items for guide trustworthiness

	Seniors	Students	MD
Competent	5.11	6.03	0.91
Convincing	5.12	6.45	1.33 **
Realistic	4.82	5.42	0.61
Reliable	5.90	6.05	0.15
Friendly	5.01	6.71	1.70 **
Likeable	4.81	6.21	1.41 **
Dominant	5.69	5.47	0.21

**p<0.01

8.2.2 Presentation Style

Some big differences between younger and older people regarding the presentation style of the agent are found. The seniors judged the presentation from the agent to be a lot more polite (t= 2.89, p<0.01), more relaxed (t= 1.73, p<0.01), a lot more accurate (t= 3.19, p<0.01), more energetic (t= .82, p<0.01) and much more calm (t= 3.81, p<0.01). A few of the remarks senior participants made after these questions were " guide gave good indications" " as it should be". The students on the other hand found the agent better (t= 1.12, p<0.05) and more exuberant (t= 1.98, p<0.01).

Table 15. Separate items for presentation style

	Seniors	Students	MD
Good	4.09	5.21	1.12 *
Pleasant	4.38	4.92	0.54
Polite	6.49	3.61	2.89 **
Natural	5.04	4.53	0.52
Flowing	5.60	5.82	0.22
Relaxed	5.68	3.95	1.73 **
Energetic	5.53	4.7	0.82 **
Dynamic	3.86	4.47	0.61
Accurate	6.77	3.58	3.19 **
Exuberant	3.75	5.74	1.98 **
Calm	6.43	3.16	3.27 **
Interested	5.00	4.47	0.52

* p<.05, **p<0.01

8.2.3 Route Description Quality

As expected, the seniors found the route description more difficult (t= 1.90, p<0.01), a lot less interesting (t= 2.59, p<0.01) and less useful (t= 1.20, p<0.01). Striking was that despite these results, the seniors found the route to be much more structured (t= 2.57, p<0.01) than the students. A lot of seniors commented things like: "too much information" "too long" and " can't remember all of that".

Table 16. Separate items for route description quality

	Seniors	Students	MD
Concise	4.66	4.05	0.60
Simple	3.68	3.82	0.14
Easy	4.12	6.03	1.90 **
Interesting	3.47	6.05	2.59 **
Structured	6.65	4.08	2.57 **
Useful	3.25	4.45	1.20 **
Clear	5.34	4.66	0.68
Comprehensible	5.44	5.63	0.19

**p<0.01

8.2.4 Task Performance

As expected, there was a significant effect on task performance (t =1.67, p<0,01), the senior participants could not remember as much as the students. It is remarkable that the amount of landmarks remembered did not differ much. The overall effect is caused by the fact that students remembered more correct turns (t =1.24, p<0,01).

Table 17. The amount of turns, landmarks and total recalled

	Seniors	Students	MD
Landmarks	3.20	3.63	0.42
Turns	3.25	4.50	1.25 **
Total	6.45	8.13	1.68 **

**p<0.01

8.2.5 Preference

Only 14% of the senior participants who saw the agent first preferred the agent, against 32% of the student participants, a significant difference.

Table 18. Preference for human guide or agent

	Seniors	Students
Preference agent	14%	32% **
Preference human guide	86%	68% **

**p<0.01

Without difference between agent first or human guide first, 48% of all student participants preferred the embodied agent for this task, while of all senior participants only 20% preferred the embodied agent.

Table 19. Overall preference for human guide or agent

	Seniors	Students
Preference agent	20%	48% **
Preference human guide	80%	52% **

**p<0.01

8.3 Comparison Between Agent and Human Guide as Judged by Seniors

8.3.1 Guide Trustworthiness

Overall, there was a significant main effect with regard to the guide trustworthiness index, in that participants found the human guide more trustworthy ($t = .90$, $p<005$). When the separate items were compared it showed that participants who saw the agent judged it as more reliable than participants who saw the human guide, but this wasn't a significant result. In table 20, all the separate items can be found.

Table 20. Separate items for guide trustworthiness

	Agent	Human	MD
Competent	5.11	6.19	1.08 *
Convincing	5.12	6.26	1.14 *
Realistic	4.82	5.94	1.12 *
Reliable	5.90	5.10	0.81
Friendly	5.01	6.90	1.90 **
Likeable	4.81	6.84	2.03 **
Dominant	5.52	5.69	0.17

* p<.05, ** p<.01

8.3.2 Presentation Style

The human guide scored slightly higher on almost every item that measured presentation style. None of the items scored significantly different.

Table 21. Separate items for presentation style

	Agent	Human	MD
Good	4.09	5.10	1.01
Pleasant	4.38	5.10	0.72
Polite	6.49	6.97	0.48
Natural	5.04	5.77	0.73
Flowing	5.60	5.16	0.44
Relaxed	5.68	6.13	0.45
Energetic	5.53	5.58	0.05
Dynamic	3.86	4.06	0.20
Accurate	6.77	7.42	0.65
Exuberant	3.75	3.84	0.09
Calm	6.43	6.94	0.51
Interested	5.00	5.29	0.29

8.3.3 Route Description Quality

As table 22 shows, the agent scored higher on five out of eight items regarding the route description quality. None of the items scored significantly different.

Table 22. Separate items for route description quality

	Agent	Human	MD
Concise	4.66	4.52	0.14
Simple	3.68	3.42	0.26
Easy	4.12	3.81	0.32
Interesting	3.47	3.71	0.24
Structured	6.65	6.58	0.07
Useful	3.25	3.71	0.46
Clear	5.34	6.00	0.66
Comprehensible	5.44	5.32	0.12

8.3.4 Task Performance

The amount of information participants could remember did not differ significantly between the two conditions. However, the senior participants who saw the human guide scored slightly better.

Table 23. The amount of turns, landmarks and total recalled

	Agent	Human	MD
Landmarks	3.20	3.22	-.02
Turns	3.25	3.45	-.20
Total	6.45	6.67	-.22

8.3.5 Preference

Of the senior participants who saw the agent 14% preferred the agent, against 26% of the senior participants who saw the human guide. However, this is not a significant difference.

Table 24. Preference for human guide or agent

	Agent first	Human first
Preference agent	14%	26% *
Preference human guide	68%	74% *

* p<.05

8.4 Evaluation of Hypotheses

H1: Senior participants who saw a route description given by the agent will trust the guide less than student participants who saw a route description given by the agent.

Statistical outcomes suggest positive proof for this hypothesis. Because there was no significant main effect, only part of this hypothesis can be accepted. Data suggests students found the agent more friendly and likable, but not significantly more realistic or competent.

H2: Senior participants who saw a route description given by the agent will evaluate the presentation style less positively than student participants who saw a route description given by the agent.

There is no statistical proof to support this hypothesis. Senior participants judged the guide on several items more positively than the students. Students found the agent to be more polite, relaxed and accurate, which are important items to measure route description. The hypothesis is rejected.

H3: Senior participants who saw a route description given by the agent will evaluate the quality of the route description less positively than student participants who saw a route description given by the agent.

There is some statistical proof to support this hypothesis. Student participants did judge the route description more positively on most items, but not all.

H4: A higher percentage of senior participants than younger participants will prefer the real person instead of the agent for this specific direction giving task.

There is convincing statistical support for this hypothesis, almost 50% of the younger participants preferred the agent, while only 20 % of the seniors preferred the agent.

H5: Senior participants who saw a route description given by the human guide will trust the guide more than the senior participants who saw a route description given by the agent.

There is statistical proof to support this hypothesis. The senior participants found the human guide trustworthier. This hypothesis is accepted.

H6: Senior participants who saw a route description given by the human guide will evaluate the presentational style more positive than the senior participants who saw a route description given by the agent.

There is not much statistical proof to support this hypothesis, the human guide scored slightly higher. The two items that did score significantly differently are very important measurements for this index: the human guide was considered significantly better and more accurate.

H7: Senior participants who saw a route description given by the human guide will evaluate the quality of the route description more positively than the senior participants who saw a route description given by the agent.

There is no statistical proof to support this hypothesis. The human guide scored slightly higher, but none of the items differed significantly. This hypothesis is rejected.

H8: Senior participants who saw a route description given by the human guide will remember more of the route description than the senior participants who saw a route description given by the agent.

This hypothesis had to be rejected; the scores for task performance between senior participants who saw the agent and those who saw the human guide were very similar.

9 Discussion Second Experiment

The main questions for this experiment were:

- What are the differences in subjective user experience between the older and younger participants who received a route description by an embodied agent?
- What are the differences in task performance between the older and younger participants who received a route description by an embodied agent?
- What are the differences in preference between the older and younger participants who received a route description by an embodied agent?
- Will senior participants who saw a route description given by the human guide judge their subjective experience different than the senior participants who saw a route description given by the agent?
- Will senior participants who saw a route description given by the human guide remember more than the senior participants who saw a route description given by the agent?

The first three questions will be discussed in section 9.1, the questions concerning older people judging the human guide and the agent will be discussed in 9.2.

9.1 Students and Seniors Compared

Due to the short amount of time participants spent with the embodied agent, it was expected that the senior participants would score lower on subjective user experience. To a certain extent, this expectation was proven to be true. However, the senior participants judged the presentation from the agent to be more polite, relaxed, more accurate, more energetic and much calmer. This is an interesting outcome. The agent was not trusted as much by the seniors as the by younger participants, but the seniors did judge the presentation style a lot more positive. But where does this effect come from? A look at the data with the comparison of the students and seniors judging the human guide revealed the solution. On each item the agent scored higher, the human guide scored higher among the seniors as well. This means that seniors appreciate qualities such as politeness, accuracy and calmness in higher any case. It is rather strange though, to see that senior participants did not really trust the guide as much as the students, but did give her such high marks on presentational style. Also the outcomes for route description quality are interesting. The senior participants considered the route more difficult, a lot less interesting and less useful but also more structured than the students. It seems that the appreciation for structure is playing a part here as well. A conclusion may be that the comparison between how the younger and older participants judged the agent is not completely fair. The data suggests that older people may attribute qualities as politeness and accuracy to virtually anybody because they appreciate these qualities more.

The experiments showed that within all groups of participants who had judged the human guide, more people preferred the agent. This was the case with younger participants but the older participants also showed this preference, although the preference for the agent among seniors was overall lower than the preference for the agent among students.

9.2 Seniors Judging the Agent and the Human Guide

The senior participants who saw the human guide found her significantly trustworthier than those who saw the agent. Notable is the fact that on one of

the items measuring this (reliability) the participants who saw the agent scored higher. Apparently, the agent did make a reliable impression. On presentational style, the differences were very small. Also the fact that the agent scored higher on five out of eight items regarding the route description quality is a very good result for those who are developing agents that are also aimed at older adults. These results are strongly supported by the fact that the differences in task performance between older adults who saw the human guide and older adults who saw the agent were minimal.

10 Overall Conclusion

The main question, as found in section 1.3 was:

Is a lifelike agent comparable to a real person in a direction-giving situation by measuring subjective user experience and task performance?

In this chapter, first the overall limitations will be discussed. They were mentioned earlier in this thesis, but are considered more extensively here. After this, the second section will contain the overall conclusion and the answer to the main question as posed above.

10.1 Limitations

An important restriction is that to be comparable with a human guide, the agent has to sound natural and display human-like nonverbal behaviours. Especially in fully interactive situations, which go beyond pure information presentation like in this experiment, achieving this still presents an important challenge.

A few limitations of this experimental design are the following. First, arguably the most important property of agents is their ability to interact with users. In this experiment, however, an agent for a non-interactive task was used: presenting route information. This option was chosen, so that all participants would get the same information in the same way. Secondly, a video recording rather than a 'live' person was used to compare with the agent. However, watching a video is not fully comparable to being face-to-face with another person. For example, Burgoon, Stoner, Bonito, and Dunbar (2003) found that mediated interaction (video conferencing) in a decision-making task scored much lower than face-to-face interaction on social judgments such as involvement, trust and sociability. On the other hand, this effect of mediation can be expected to be smaller in situations where there is no actual interaction, as in our experiment. People are used to seeing mediated people presenting information, for example newsreaders on television. And an embodied agent is in any case mediated: people need a computer to interact with it. This means that to keep the experimental conditions as similar as possible, the human guide in our experiment had to be mediated too.

The participants may also have had higher expectations of the human guide than of the virtual guide: when people see a real person explaining a route, they may expect more spontaneous gestures than were actually performed by the actress. This could have caused the participants to judge the route description by the human as relatively static and boring. On the other hand, one of the younger participants remarked: "I can imagine an agent explaining something in a very boring way". An agent will give a steady, always similar performance, and people expect this to happen. This may explain as well why the real person was found to be more static and boring. In addition, the combination of a human voice and appearance with artificial behaviour (as in some sense the human guide was mimicking the artificial agent) may have been perceived as inconsistent. As shown by Isbister and Nass (2000), people tend to dislike inconsistency within agents. However, the agent was also inconsistent in the sense that it coupled an artificial appearance and behaviour with a human voice (i.e., the voice of the human guide). Some participants remarked that they found this unnatural, although this did not lead to a more negative judgement. However, some subjects (in particular those having previous experience with embodied agents) commented that they found the combination of a human voice with a synthetic

agent somewhat unnatural. So, the use of a natural voice might both have been an advantage and a disadvantage.

10.2 Conclusion

The results are suggesting that the performance of a lifelike agent is indeed comparable to the performance of a real person. Especially in task performance, no significant differences to the advantage of either the agent or the real person are found. In the first experiment the results suggest advantages of a route description supported by gestures, especially on user's subjective experience.

Older users do prefer a human guide, but this is not fully supported by the outcomes of presentation style and route description quality. Seniors may have preferred the real person because they are not into novelties such as embodied agents, but this does not mean they cannot work with it. There is no significant difference in task performance between the seniors who saw the description given by the human guide or those who saw the agent. However, the lack of trust is not something to neglect, it is a fact that seniors who saw the agent did not trust her as much as those who saw the human guide.

10.3 Further Research

There are still several options for further research. The age group used in this experiment are mostly people who are still employed. An experiment with users who are retired, may result in different outcomes. But these people may very well be the ones where an embodied agent can be very useful, since many people grow lonely after their partner passes away. An experiment with a male agent and other speech and interaction levels may be very useful as well, since that was not part of the objectives in this experiment.

A more professional actor or a more spontaneous route description might also change the outcomes, since it is possible that the not entirely spontaneous behaviour from the human guide led to a lower score for the human guide.

Another interesting factor is the time people spend with the embodied agent. One might expect that if people get the time to get used to how an agent presents the information, they might be able to focus even more on the message instead of the presenter. In this experiment people were confronted with only very short movies; but if there is extended usage, the differences in perception will probably become clearer. The effects of extended usage might also reduce the amount of differences between the students and the seniors with regards to subjective user experience.

References

Allbeck, J. & Badler, N. (2002, July). Toward Representing Agent Behaviors Modified by Personality and Emotion, paper presented at First Intl. Joint Conf. on Autonomous Agents and Multi-Agent Systems, Bologna, Italy

Baldi, R.A. (1997). Training older adults to use the computer: Issues related to the workplace, attitudes, and training. *Educational Gerontology 23*(5)*,* 453-465.

Beun, R.J., Vos, E. de & Witteman, C. (2003). Embodied conversational agents: effects on memory performance and anthropomorphisation. In Thomas Rist, Ruth Aylett, Daniel Ballin et al. (Eds.), *Intelligent Virtual Agents: 4th International Workshop, IVA 2003*, 315-319.

Bickmore, T., Caruso, L. and Clough-Gorr, K. (2005). Acceptance and usability of a relational agent interface by urban older adults. *Proceedings of CHI'05*, 1212-1215.

Bickmore, T. and Cassell, J. (2001) Relational Agents: A Model and Implementation of Building User Trust. In Beaudouin-Lafon, M., Jacob, R.J.K. (Eds.) *Proceedings of the SIGCHI conference on Human factors in computing systems,* 396-403

Burgoon, J.K., Stoner, G.M., Bonito, J.A., & Dunbar, N.E. (2003). Trust and Deception in Mediated Communication. *Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS'03) - Track1 - Volume 1*, 44.1

Burke, D.M. & MacKay, D.G. (1997). Memory, language and ageing. *Philosophical Transactions of the Royal Society: Biological Sciences*, *352*, 1845-1856.

Cassell, J., Stocky, T., Bickmore, T., Gao, Y., Nakano, Y., Ryokai, K., Tversky, D., Vaucelle, C., and Vilhjálmsson, J. (2002). MACK: Media lab Autonomous Conversational Kiosk. *Proceedings of the 7th international conference on Intelligent user interfaces*, 224-225.

Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (Eds.). (2000). *Embodied Conversational Agents*. Cambridge, USA: MIT Press.

Cassell, J., Viljhálmsson, H. & Bickmore, T. (2001) BEAT: The Behavior Expression Animation Toolkit*. Proceedings of the 28th annual conference on Computer graphics and interactive techniques,* 477-486.

Cohen, A. (1980). The use of hand illustrators in directiongiving situations. In W. von Raffler-Engel (Ed.), *Aspects of Nonverbal Communication*, Lisse: Swets and Zeitlinger BV.

Czaja, S.J. and Lee, C.C. (2003). Designing computer systems for older adults. In Jacko, J.A. and Sears, A. (Eds.), *The Human-Computer-Interaction Handbook* (pp. 413-427). New Jersey: Mahwah.

Dehn, J. & Mulken, S. van. (2000). The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies, 52*, 1-22.

Hawthorn, D. (2000). Possible implications of aging for interface designers. *Interacting with Computers, 12*, 507-528.

Hazzlewood, J. (2002, October). *From crystal set to cyberland: Third Age learners and the Internet.* Paper written for the 6th Global Ageing IFA Conference: Maturity Matters, Perth.

Isbister, K. & Nass, C. (2000). Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies, 53(1)*, 251-267.

Kelley, C.L. & Charness, N. (1995). Issues in training older adults to use computers*. Behaviour & Information Technology 14(2),* 107-120.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction 27, 3*, 175-200.

Kepplinger, H. M. (1991). The impact of presentation Techniques: theoretical aspects and em[erical findings. In Frank Biocca (Ed.), *television and political advertising. Vol 1: Psychological processes* (pp. 173-194). Hilldale: Lawrence Earlbaum Ass.

King, W.J. & Ohya, J. (1996). The representation of agents: Anthropomorphism, agency and intelligence. *Conference companion on Human factors in computing systems: common ground*, 289-290.

Koda, T. & Maes, P. (1996). Agents with faces: the effect of personification. *Proceedings of the Fifth IEEE International Workshop on Robot and Human Communication (RO-MAN '96),* 189-194.

Kopp, S., Gesellensetter, L., Krämer, N.C. & Wachsmuth, I. (2005). A Conversational Agent as Museum Guide – Design and Evaluation of a Real-World Application. *Proceedings of IVA'05.*

Lee, E-J. & Nass, C.I. (1998). Does the ethnicity of a computer agent matter? An experimental comparison of human-computer interaction and computer-mediated communication. *Proceedings of the 1998 Workshop on Embodied Conversational Characters*.

Lester, J., Kahler, S., Barlow, T., Stone, B. & Bhogal.R. (1997). The persona effect: Affective Impact of Animated pedagogical Agents. In S. pemberton (Ed.). *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 359-366). Atlanta.

Lester, J. Zettlemoyer, L, Gregoire, J., & Bares, W. (1999). Explanatory lifelike avatars: performing user centered tasks in 3D learning environments. *Proceedings of the third annual conference on Autonomous Agents* (pp. 24-31). Seattle.

Look, G., Kottahachchi, B., Laddaga, R., Shrobe, H. A Location Representation for Generating Descriptive Walking Directions. *Proceedings of the 10th international conference on Intelligent user interfaces* (pp. 122-129). San Diego.

Lovelace, K.L., Hegarty, M. & Montello D.R. (1999). Elements of Good Route Directions in Familiar and Unfamiliar Environments. *Proceedings of the International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science* (pp. 65-88). Stade, Germany.

Mayer, R.E. & Moreno, R. (2002). Animation as an Aid to Multimedia Learning. *Educational Psychology Review 14-1*, 87-99.

McBreen, H., Shade, P., Jack, M., & Wyard, P. (2000). Experimental assessment of the effectiveness of synthetic personae for multi-modal E-retail applications. *Proceedings of the fourth international conference on Autonomous agents* (pp. 39-45). Barcelona.

McNeill, D. (1992). *Hand and Mind: What Gestures reveal about Thought*. Chicago: University of Chicago Press.

Miller, G. (1956). The magical number seven Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review, vol. 63*, 81-97.

Mulken, S. van, Andre, E. & Müller, J. (1998). The persona effect: how substantial is it? *Proceedings of HCI on People and Computers XIII* (pp. 56-66).

Morris, J. M., (1994). Computer training needs of older adults. *Educational Gerontology, 20*, 541-555.

Nass, C., Steuer, J. & Hendriksen, L. (1994). Machines, Social Attributions, and Ethopoeia: Performance Assessment of Computers Subsequent to "Self-" or "Other-" Evaluations. *International Journal of Human-Computer Studies, 40*(3)*,* 543-559.

Nass, C., Steuer, J. & Tauber, E. (1994). Computers are social actors. *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence* (pp. 72-78). Boston, MA.

Noot, H. & Ruttkay, Zs. (2005). Variations in Gesturing and Speech by GESTYLE. *International Journal of Human-Computer Studies, 62*(2), 211-229.

O'Keefe, D.J. (1990). *Persuasion: Theory and Research*. Newbury Park: Sage Publications, Inc.

Reardon, K. (1991). *Persuasion in Practice.* Newbury Park: Sage Publications, Inc.

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places.* New York: Cambridge University Press.

Rickel, J. & Johnson, W.L. (1997). Intelligent Tutoring in Virtual Reality: A Preliminary Report. *Proceedings of the Eighth World Conference on AI in Education* (pp. 294-301).

Rist, T., Andre, E. & Baldes, S. (2003). A flexible platform for building applications with life-like characters. *Proceedings of the 8th international conference on Intelligent user interfaces* (pp. 158-165). Miami, Florida, USA.

Rogers, W. A. (2000). Attention and aging. In D.C. Park & N. Schwarz (Eds.), *Cognitive Aging: A primer* (pp. 57-73). Philadelphia: Psychology Press.

Ruttkay, Zs., Doorman, C. & Noot, H. (2002). Evaluating ECAs - What and how?. *Proceedings of the AAMAS02 Workshop on Embodied conversational agents - let's specify and evaluate them!.* Bologna, Italy.

Smith, J. (2000). *GrandChair: Conversational Collection of Grandparent's Stories.* Unpublished master's thesis, MIT.

Sorrows, M.E. & Hirtle, S.C. (1999). The nature of landmarks for real and electronic spaces, *Proceedings of the International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science* (pp. 37-50).

Sproull, L., Subramani, R., Kiesler, S., Walker, J., & Waters, K. (1996). When the interface is a face. *Human-Computer Interaction, 11*, 97-124.

Stocky, T. (2002). *Conveying Routes: Multimodal Generation and Spatial Intelligence in Embodied Conversational Agents*. Master's thesis, MIT. Retrieved August 2005 from http://www.media.mit.edu/gnl/pubs/StockyThesis.pdf

Takeuchi, A., & Naito, T. (1995). Situated facial displays: towards social interaction. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 450-455). Denver, Colorado.

Theune, M., Heylen, D. & Nijholt, A. (2005). Generating embodied information presentations. In O. Stock & M. Zancanaro (Eds.), *Multimodal Intelligent Information Presentation.* Kluwer Academic Publishers.

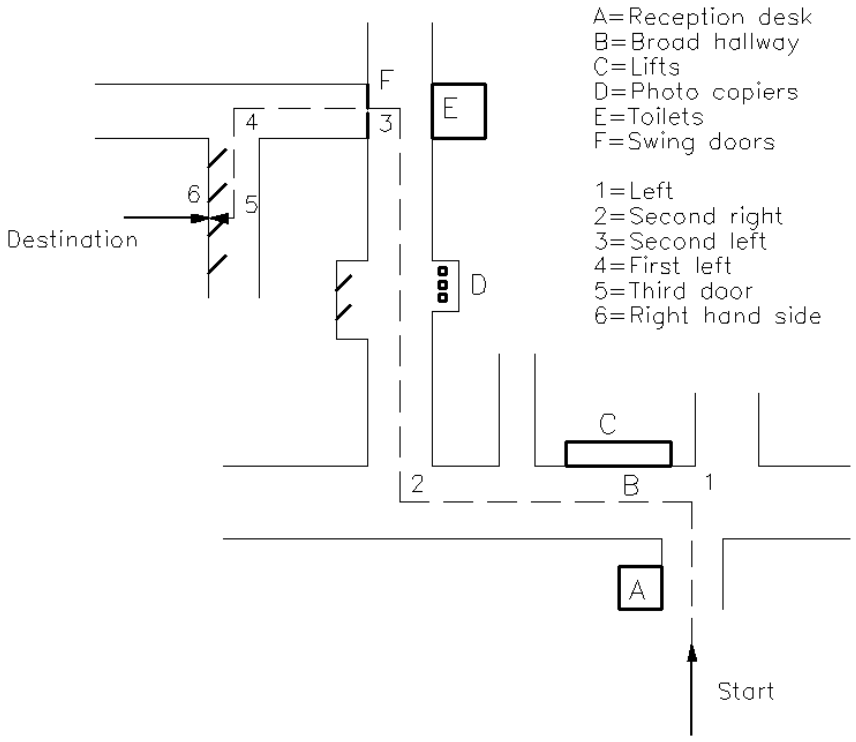
Walker, J. H., Sproull, L. & Subramani, R. (1994). Using a human face in an interface. *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence* (pp. 85-91). Boston, MA.

Williamson, K., Bow, A. & Wale, K. (1997). Older People and the Internet*, Link-up.* 9-12.

Appendix A: Route description

The route verbally
"When you enter the building, walk past the reception desk and take a left turn into a broad hall way. Walk past the lifts and take the second turn to the right. You will pass some photocopier, just go straight on. After the toilets on your right hand side, go left trough the swing doors. After a few meters take the first turn to the left. You are now in the right hallway, take the third door on your right hand side and you have arrived at your destination"

A map of the route

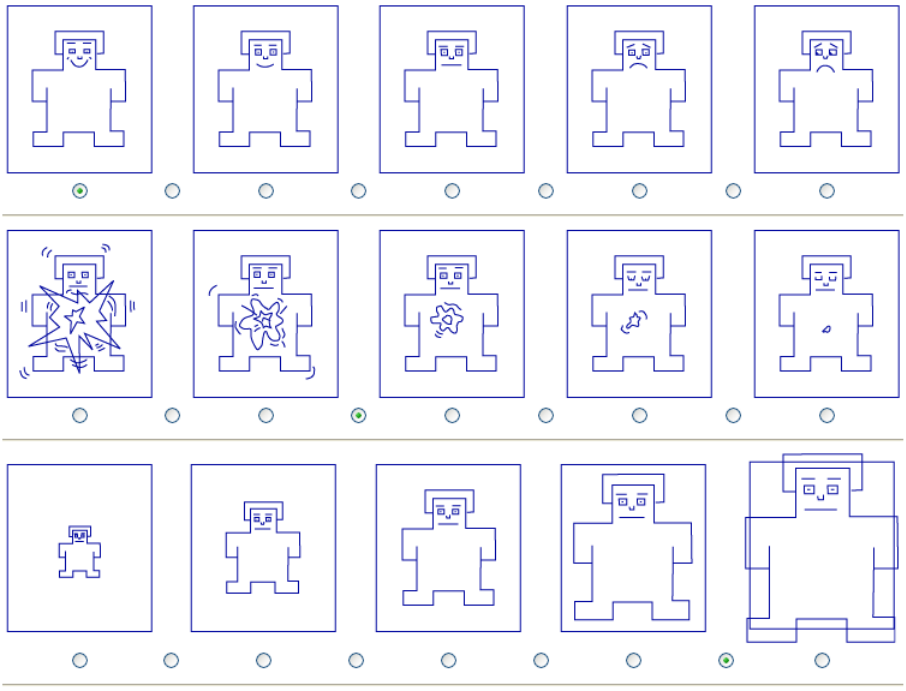


Please note that the numbers 1 till 6 represent the six turns, and the letters A till F represent the six landmarks.

Appendix B: The questionnaire

Eerste reactie

1) Geef voor elk van de drie onderstaande visuele schalen aan welke figuur het meest overeenkomt met het gevoel dat dit prototype bij u oproept.



Vragen over de gids

2) Gebruik de onderstaande schalen om aan te geven hoe deze gids op u *overkomt*. (geef dit voor elk aspect aan s.v.p.)

zeer deskundig	○	◉	○	○	○	○	○	○	○	zeer ondeskundig
zeer ongeloofwaardig	○	◉	○	○	○	○	○	○	○	zeer geloofwaardig
zeer realistisch	○	○	○	○	○	○	◉	○	○	zeer onrealistisch
zeer onbetrouwbaar	○	○	○	○	◉	○	○	○	○	zeer betrouwbaar
zeer vriendelijk	○	◉	○	○	○	○	○	○	○	zeer onvriendelijk
zeer onsympathiek	○	○	○	◉	○	○	○	○	○	zeer sympathiek
zeer onderdanig	○	○	○	○	○	○	◉	○	○	zeer dominant
zeer eigenzinnig	○	◉	○	○	○	○	○	○	○	zeer meegaand
zeer emotioneel	○	○	○	○	○	◉	○	○	○	zeer (emotioneel) stabiel
zeer koel	○	◉	○	○	○	○	○	○	○	zeer warm
zeer ongevoelig	○	○	○	○	◉	○	○	○	○	zeer gevoelig
zeer berekenend	○	◉	○	○	○	○	○	○	○	zeer open
zeer sociaal gericht	○	◉	○	○	○	○	○	○	○	zeer op zichzelf gericht
zeer ontspannen	○	○	○	○	○	○	◉	○	○	zeer gespannen
zeer zelfverzekerd	○	◉	○	○	○	○	○	○	○	zeer onzeker
zeer opportunistisch	○	○	○	◉	○	○	○	○	○	zeer conformistisch
zeer impulsief	○	○	○	○	◉	○	○	○	○	zeer beheerst
zeer serieus	○	○	○	○	○	◉	○	○	○	zeer enthousiast
zeer voorzichtig	○	○	○	◉	○	○	○	○	○	zeer onbevreesd
zeer traditioneel	○	○	○	○	○	◉	○	○	○	zeer innovatief
zeer goed van vertrouwen	○	○	○	○	◉	○	○	○	○	zeer wantrouwend
zeer nuchter	○	○	○	○	○	◉	○	○	○	zeer fantasierijk
zeer concreet denkend	○	○	○	○	○	◉	○	○	○	zeer abstract denkend

3) Wat vindt u van de *manier* waarop de gids de route uitlegt? (geef dit voor elk aspect aan a.u.b.)
N.B. De routebeschrijving zelf komt in de volgende vraag aan de orde.

zeer goed	○	○	◉	○	○	○	○	○	○	zeer slecht
zeer onprettig	○	○	○	○	◉	○	○	○	○	zeer prettig
zeer beleefd	○	○	○	◉	○	○	○	○	○	zeer onbeleefd
zeer natuurlijk	○	○	○	○	○	◉	○	○	○	zeer gekunsteld
zeer houterig	○	○	◉	○	○	○	○	○	○	zeer vloeiend
zeer ontspannen	○	○	○	○	◉	○	○	○	○	zeer gespannen
zeer energiek	○	○	○	○	○	◉	○	○	○	zeer sloom
zeer statisch	○	○	○	◉	○	○	○	○	○	zeer dynamisch
zeer zorgvuldig	○	◉	○	○	○	○	○	○	○	zeer onzorgvuldig
zeer uitbundig	○	○	○	○	◉	○	○	○	○	zeer ingehouden
zeer kalm	○	○	◉	○	○	○	○	○	○	zeer opgewonden
zeer geinteresseerd	○	○	○	○	○	◉	○	○	○	zeer ongeinteresseerd

4) Wat vindt u van de routebeschrijving? (geef dit voor elk aspect aan a.u.b.)

zeer langdradig	○	○	○	◉	○	○	○	○	○	zeer bondig
zeer ingewikkeld	○	○	○	○	◉	○	○	○	○	zeer eenvoudig
zeer makkelijk	○	○	◉	○	○	○	○	○	○	zeer moeilijk
zeer boeiend	○	○	○	○	◉	○	○	○	○	zeer saai
zeer gestructureerd	○	○	○	◉	○	○	○	○	○	zeer ongestructureerd
zeer onbruikbaar	○	○	○	○	○	◉	○	○	○	zeer bruikbaar
zeer duidelijk	○	○	◉	○	○	○	○	○	○	zeer onduidelijk
zeer onbegrijpelijk	○	○	○	○	◉	○	○	○	○	zeer begrijpelijk

5) Beschrijf in uw eigen woorden de route die de gids zojuist heeft uitgelegd.
Probeer zoveel mogelijk details van de route te noemen.

Voorkeur voor een prototype

1) a. U heeft nu twee prototypen van gidsen gezien, zoals die in een computerapplicatie zouden kunnen worden toegepast, te weten een echte persoon (video) en een virtuele persoon.

Welke van de twee heeft uw voorkeur?

- ○ Echte persoon (video)
- ◉ Virtuele persoon

b. Geef hieronder een toelichting op uw antwoord.

ik wordt minder afgeleid

Vragen over gebruik van virtuale personen

Over het gebruik van virtuele personen in computerapplicaties willen we nog een paar vragen stellen.

2) Heeft u in computertoepassingen eerder te maken gehad met dit soort sprekende virtuele personen??

- ◉ Ja
- ○ Nee

Zo ja, in wat voor een toepassing?

computerspelletjes

3) Wat vindt u van de agent die in dit experiment werd gebruikt?
(geef dit voor elk aspect aan s.v.p.)

zeer goed	◉	○	○	○	○	○	○	○	○	zeer slecht
zeer ouderwets	○	○	○	◉	○	○	○	○	○	zeer modern
zeer realistisch	○	○	◉	○	○	○	○	○	○	zeer onrealistisch
zeer geavanceerd	○	○	○	◉	○	○	○	○	○	zeer achterhaald
zeer onbruikbaar	○	○	○	○	◉	○	○	○	○	zeer bruikbaar
zeer traditioneel	○	○	◉	○	○	○	○	○	○	zeer innovatief

Ruimte voor een korte toelichting:

4) a. Indien u zou mogen kiezen voor een vrouwelijk danwel een mannelijk virtueel persoon in een computerapplicatie, waar zou dan uw voorkeur naar uit gaan?

- ○ vrouwelijk virtueel persoon
- ○ mannelijk virtueel persoon
- ○ maakt niet uit
- ◉ dat hangt ervan af

b. Geef hieronder een toelichting op uw antwoord.

van de informatie die gepresenteerd wordt

Appendix C: Full data analysis student experiment

* p<.05
** p<.01

Het tweezijdige significantie niveau is weergegeven.

A Agent – zonder vs met gebaren

I SAM-emoties (het gevoel dat opgeroepen wordt)

	Zonder gebaren	Met gebaren	Mean Difference
Valence	4.95	4.37	.59 *
Arousal	6.39	6.68	-.3
Dominance	5.43	4.97	.46

II Persoonlijkheid - overtuiging

	Zonder gebaren	Met gebaren	Mean Difference
Deskundig	5.16	6.03	-.87*
Geloofwaardig	6.18	6.45	-.27
Realistisch	4.34	5.42	-1.08 **
Betrouwbaar	6.11	6.05	.06
Vriendelijk	5.75	6.71	-.96 *
Sympathiek	5.75	6.21	-.46
Dominant	4.95	5.47	-.52

III Persoonlijkheid (Cattell)

	Zonder gebaren	Met gebaren	Mean Difference
Eigenzinnig	5.23	4.71	.52
Emotioneel	6.50	6.89	-.39
Koel	4.23	4.61	-.38
Gevoelig	3.91	4.55	-.64 *
Open	4.80	4.89	-.09
Sociaal	5.02	4.42	.60
Gespannen	4.11	3.58	.53
Onzeker	3.77	2.97	.80 *
Opportunistisch	4.95	4.66	.30
Impulsief	7.16	7.18	-.02
Enthousiast	3.55	3.39	.15
Voorzichtig	4.66	5.34	-.68 *
Innovatief	4.57	5.55	-.98 **
Wantrouwend	4.00	4.05	-.05
Fantasierijk	3.36	3.29	.07
Abstract denkend	3.61	3.82	-.20

IV Manier waarop route wordt uitgelegd

	Zonder gebaren	Met gebaren	Mean Difference
Goed	4.30	4.79	-.49
Prettig	4.02	4.92	-.90 *
Beleefd	6.52	6.39	.13
Natuurlijk	4.82	5.47	-.66
Houterig	4.77	5.82	-1.04 *
Ontspannen	5.86	6.05	-.19
Energiek	4.50	5.29	-.79
Statisch	3.23	4.47	-1.25 **
Zorgvuldig	6.75	6.42	.33
Uitbundig	4.14	4.26	-.13
Kalm	6.98	6.84	.14
Geïnteresseerd	4.64	5.53	-.89 **

V Mening over routebeschrijving

	Zonder gebaren	Met gebaren	Mean Difference
Langdradig	3.80	4.05	-.26
Ingewikkeld	3.57	3.82	-.25
Makkelijk	3.66	3.97	-.31
Boeiend	3.43	3.95	-.52
Gestructureerd	5.64	5.92	-.28
Bruikbaar	4.55	4.45	.10
Duidelijk	4.95	5.34	-.39
Begrijpelijk	5.30	5.63	-.34

VI Mening over gebruikte agent

	Zonder gebaren	Met gebaren	Mean Difference
Goed	5.07	6.00	-.93 *
Modern	5.16	5.95	-.79 *
Realistisch	5.25	5.92	-.67
Geavanceerd	4.70	5.82	-1.11 **
Bruikbaar	5.16	6.24	-1.08 **
Innovatief	4.66	5.16	-.50

VII Informatie herinnering

	Zonder gebaren	Met gebaren	Mean Difference
Wendingen	4.50	4.05	.45
Oriëntatiepunten	3.70	3.52	.18
Totaal	7.58	8.20	.62

B Mens - zonder vs met gebaren

I SAM-emoties (het gevoel dat opgeroepen wordt)

	Zonder gebaren	Met gebaren	Mean Difference
Valence	4.86	4.83	.03
Arousal	6.24	6.13	.11
Dominance	4.71	5.45	-.74 *

II Persoonlijkheid – overtuiging

	Zonder gebaren	Met gebaren	Mean Difference
Deskundig	5.36	5.05	.31
Geloofwaardig	6.40	6.60	-.20
Realistisch	6.43	5.98	.45
Betrouwbaar	6.21	6.05	.16
Vriendelijk	6.64	6.43	.22
Sympathiek	6.38	6.18	.21
Dominant	5.57	5.53	.04

III Persoonlijkheid (Cattell)

	Zonder gebaren	Met gebaren	Mean Difference
Eigenzinnig	4.83	5.05	-.22
Emotioneel	6.88	6.58	.31
Koel	4.62	4.35	.27
Gevoelig	4.43	4.78	-.35
Open	4.71	4.90	-.19
Sociaal	4.55	4.65	-.10
Gespannen	4.17	4.65	-.48
Onzeker	3.36	3.70	-.34
Opportunistisch	4.81	5.13	-.32
Impulsief	7.05	7.03	.02
Enthousiast	3.12	3.18	-.06
Voorzichtig	4.69	5.10	-.41
Innovatief	3.76	4.15	-.39
Wantrouwend	4.00	3.78	.23
Fantasierijk	2.95	3.13	-.17
Abstract denkend	3.62	4.45	-.83

IV Manier waarop route wordt uitgelegd

	Zonder gebaren	Met gebaren	Mean Difference
Goed	5.26	4.70	.56
Prettig	4.67	4.70	-.03
Beleefd	6.69	6.23	.47
Natuurlijk	5.12	4.88	.24
Houterig	5.36	5.28	.08
Ontspannen	5.69	5.35	.34
Energiek	4.67	4.75	-.08
Statisch	3.64	3.63	.01
Zorgvuldig	6.81	6.38	.43
Uitbundig	4.05	4.03	.02
Kalm	6.83	6.88	-.05
Geïnteresseerd	5.43	4.83	.60

V Mening over routebeschrijving

	Zonder gebaren	Met gebaren	Mean Difference
Langdradig	3.71	3.30	.41
Ingewikkeld	3.98	3.33	.65
Makkelijk	4.07	3.58	.50
Boeiend	3.38	3.00	.38
Gestructureerd	5.69	5.55	.14
Bruikbaar	4.71	4.08	.64
Duidelijk	5.31	5.00	.31
Begrijpelijk	5.76	5.28	.49

VI Mening over gebruikte agent

	Zonder gebaren	Met gebaren	Mean Difference
Goed	4.93	5.88	-.95 *
Modern	5.88	6.00	-.12
Realistisch	4.60	5.00	-.40
Geavanceerd	4.88	4.93	-.05
Bruikbaar	5.52	5.53	-.01
Innovatief	5.05	5.35	-.30

VII Informatie herinnering

	Zonder gebaren	Met gebaren	Mean Difference
Wendingen	3.76	4.22	-.46
Oriëntatiepunten	3.24	3.87	-.63 *
Totaal	7.00	8.10	-1.10 *

C Zonder gebaren – mens vs agent

I SAM-emoties (het gevoel dat opgeroepen wordt)

	Agent	Mens	Mean Difference
Valence	4.95	4.86	.09
Arousal	6.39	6.24	.15
Dominance	5.43	4.71	.72

II Persoonlijkheid – overtuiging

	Agent	Mens	Mean Difference
Deskundig	5.16	5.36	-.20
Geloofwaardig	6.18	6.40	-.22
Realistisch	4.34	6.43	-2.09 **
Betrouwbaar	6.11	6.21	-.10
Vriendelijk	5.75	6.64	-.89 *
Sympathiek	5.75	6.38	-.63
Dominant	4.95	5.57	-.62 **

III Persoonlijkheid (Cattell)

	Agent	Mens	Mean Difference
Eigenzinnig	5.23	4.83	.39
Emotioneel	6.50	6.88	-.38
Koel	4.23	4.62	-.39
Gevoelig	3.91	4.43	-.52
Open	4.80	4.71	.09
Sociaal	5.02	4.55	.48
Gespannen	4.11	4.17	-.06
Onzeker	3.77	3.36	.42
Opportunistisch	4.95	4.81	.15
Impulsief	7.16	7.05	.11
Enthousiast	3.55	3.12	.43
Voorzichtig	4.66	4.69	-.03
Innovatief	4.57	3.76	.81 **
Wantrouwend	4.00	4.00	.00
Fantasierijk	3.36	2.95	.41
Abstract denkend	3.61	3.62	-.01

IV Manier waarop route wordt uitgelegd

	Agent	Mens	Mean Difference
Goed	4.30	5.26	-.97 *
Prettig	4.02	4.67	-.64
Beleefd	6.52	6.69	-.17
Natuurlijk	4.82	5.12	-.30
Houterig	4.77	5.36	-.58
Ontspannen	5.86	5.69	.17
Energiek	4.50	4.67	-.17
Statisch	3.23	3.64	-.42
Zorgvuldig	6.75	6.81	-.06
Uitbundig	4.14	4.05	.09
Kalm	6.98	6.83	.14
Geïnteresseerd	4.64	5.43	-.79

V Mening over routebeschrijving

	Agent	Mens	Mean Difference
Langdradig	3.80	3.71	.09
Ingewikkeld	3.57	3.98	-.41
Makkelijk	3.66	4.07	-.41
Boeiend	3.43	3.38	.05
Gestructureerd	5.64	5.69	-.05
Bruikbaar	4.55	4.71	-.17
Duidelijk	4.95	5.31	-.35
Begrijpelijk	5.30	5.76	-.47

VI Mening over gebruikte agent

	Agent	Mens	Mean Difference
Goed	5.07	4.93	.14
Modern	5.16	5.88	-.72
Realistisch	5.25	4.60	.65
Geavanceerd	4.70	4.88	-.18
Bruikbaar	5.16	5.52	-.36
Innovatief	4.66	5.05	-.39

VII Informatie herinnering

	Agent	Mens	Mean Difference
Wendingen	4.50	3.76	.74 *
Oriëntatiepunten	3.70	3.23	.47
Totaal	8.20	7.00	1.20 *

D Met gebaren - agent vs mens

I SAM-emoties (het gevoel dat opgeroepen wordt)

	Agent	Mens	Mean Difference
Valence	4.37	4.83	-.46
Arousal	6.68	6.13	.56
Dominance	4.97	5.45	-.48

II Persoonlijkheid - overtuiging

	Agent	Mens	Mean Difference
Deskundig	6.03	5.05	.98 *
Geloofwaardig	6.45	6.60	-.15
Realistisch	5.42	5.98	-.55
Betrouwbaar	6.05	6.05	.00
Vriendelijk	6.71	6.43	.29
Sympathiek	6.21	6.18	.03
Dominant	5.47	5.53	-.06

III Persoonlijkheid (Cattell)

	Agent	Mens	Mean Difference
Eigenzinnig	4.71	5.05	-.34
Emotioneel	6.89	6.58	.32
Koel	4.61	4.35	.26
Gevoelig	4.55	4.78	-.22
Open	4.89	4.90	-.01
Sociaal	4.42	4.65	-.23
Gespannen	3.58	4.65	-1.07 **
Onzeker	2.97	3.70	-.73
Opportunistisch	4.66	5.13	-.47
Impulsief	7.18	7.03	.16
Enthousiast	3.39	3.18	.22
Voorzichtig	5.34	5.10	.24
Innovatief	5.55	4.15	1.40 **
Wantrouwend	4.05	3.78	.28
Fantasierijk	3.29	3.13	.16
Abstract denkend	3.82	4.45	-.63

IV Manier waarop route wordt uitgelegd

	Agent	Mens	Mean Difference
Goed	4.79	4.70	.09
Prettig	4.92	4.70	.22
Beleefd	6.39	6.23	.17
Natuurlijk	5.47	4.88	.60
Houterig	5.82	5.28	.54
Ontspannen	6.05	5.35	.70
Energiek	5.29	4.75	.54
Statisch	4.47	3.63	.85 *
Zorgvuldig	6.42	6.38	.04
Uitbundig	4.26	4.03	.24
Kalm	6.84	6.88	-.04
Geïnteresseerd	5.53	4.83	.70 *

V Mening over routebeschrijving

	Agent	Mens	Mean Difference
Langdradig	4.05	3.30	.75
Ingewikkeld	3.82	3.33	.49
Makkelijk	3.97	3.58	.40
Boeiend	3.95	3.00	.95 **
Gestructureerd	5.92	5.55	.37
Bruikbaar	4.45	4.08	.37
Duidelijk	5.34	5.00	.34
Begrijpelijk	5.63	5.28	.36

VI Mening over gebruikte agent

	Agent	Mens	Mean Difference
Goed	6.00	5.88	.13
Modern	5.95	6.00	-.05
Realistisch	5.92	5.00	.92 *
Geavanceerd	5.82	4.93	.89 *
Bruikbaar	6.24	5.53	.71
Innovatief	5.16	5.35	-.19

VII Informatie herinnering

	Agent	Mens	Mean Difference
Wendingen	4.05	4.22	-.17
Oriëntatiepunten	3.52	3.88	-.34
Totaal	7.58	8.10	-.52

Appendix D: Overview of gestures

		
take a left turn	the second turn right	go straight on
		
toilets on right hand side	left trough swing doors	first turn to the left
		
third door	on your right hand side	

Appendix E: Full data analysis senior experiment and combination senior and student experiment.

Ouderen

A Met gebaren – agent vs mens

I SAM-emoties (het gevoel dat opgeroepen wordt)

	Agent	Mens	Mean Difference
Valence	5.98	5.94	.04
Arousal	6.06	4.94	1.13 *
Dominance	4.69	4.45	.24

II Persoonlijkheid - overtuiging

	Agent	Mens	Mean Difference
Deskundig	5.11	6.19	-1.08 *
Geloofwaardig	5.12	6.26	-1.14 *
Realistisch	4.82	5.94	-1.12 *
Betrouwbaar	5.90	5.10	.81
Vriendelijk	5.01	6.90	-1.90 **
Sympathiek	4.81	6.84	-2.03 **
Dominant	5.69	5.52	.17

III Persoonlijkheid (Cattell)

	Agent	Mens	Mean Difference
Eigenzinnig	4.61	4.39	.22
Emotioneel	6.04	6.52	-.47
Koel	3.94	4.03	-.09
Gevoelig	3.31	4.10	-.79 *
Open	4.74	5.06	-.33
Sociaal	5.79	5.00	.79
Gespannen	4.55	4.10	.45
Onzeker	3.58	2.87	.71
Opportunistisch	4.56	5.58	-1.02 **
Impulsief	6.70	7.03	-.33
Enthousiast	3.67	3.77	-.10
Voorzichtig	5.54	5.65	-.11
Innovatief	3.68	4.03	-.35
Wantrouwend	4.33	4.26	.07
Fantasierijk	3.69	3.32	.36
Abstract denkend	3.57	3.71	-.14

IV Manier waarop route wordt uitgelegd

	Agent	Mens	Mean Difference
Goed	4.09	5.10	-1.01
Prettig	4.38	5.10	-.72
Beleefd	6.49	6.97	-.48
Natuurlijk	5.04	5.77	-.73
Houterig	5.60	5.16	.44
Ontspannen	5.68	6.13	-.45
Energiek	5.53	5.58	-.05
Statisch	3.86	4.06	-.20
Zorgvuldig	6.77	7.42	-.65
Uitbundig	3.75	3.84	-.09
Kalm	6.43	6.94	-.51
Geïnteresseerd	5.00	5.29	-.29

V Mening over routebeschrijving

	Agent	Mens	Mean Difference
Langdradig	4.66	4.52	.14
Ingewikkeld	3.68	3.42	.26
Makkelijk	4.12	3.81	.32
Boeiend	3.47	3.71	-.24
Gestructureerd	6.65	6.58	.07
Bruikbaar	3.25	3.71	-.46
Duidelijk	5.34	6.00	-.66
Begrijpelijk	5.44	5.32	.12

VI Mening over gebruikte agent

	Agent	Mens	Mean Difference
Goed	5.50	5.58	-.08
Modern	5.59	6.03	-.44
Realistisch	5.61	5.74	-.13
Geavanceerd	5.24	6.16	-.92 **
Bruikbaar	4.92	6.00	-1.08 *
Innovatief	4.64	5.29	-.65

VII Informatie herinnering

	Agent	Mens	Mean Difference
Wendingen	3.20	3.22	-.02
Oriëntatiepunten	3.25	3.45	-.20
Totaal	6.45	6.67	-.22

Jongeren vs ouderen

A Agent – ouderen vs jongeren

I SAM-emoties (het gevoel dat opgeroepen wordt)

	Ouderen	Jongeren	Mean Difference
Valence	5.98	4.37	1.61 **
Arousal	6.06	6.68	-.62
Dominance	4.69	4.97	-.28

II Persoonlijkheid - overtuiging

	Ouderen	Jongeren	Mean Difference
Deskundig	5.11	6.03	-.91
Geloofwaardig	5.12	6.45	-1.33 **
Realistisch	4.82	5.42	-.61
Betrouwbaar	5.90	6.05	-.15
Vriendelijk	5.01	6.71	-1.70 **
Sympathiek	4.81	6.21	-1.41 **
Dominant	5.69	5.47	.21 *

III Persoonlijkheid (Cattell)

	Ouderen	Jongeren	Mean Difference
Eigenzinnig	4.61	4.71	-.11
Emotioneel	6.04	6.89	-.85 *
Koel	3.94	4.61	-.67
Gevoelig	3.31	4.55	-1.25 **
Open	4.74	4.89	-.16
Sociaal	5.79	4.42	1.37 **
Gespannen	4.55	3.58	.97 **
Onzeker	3.58	2.97	.60
Opportunistisch	4.56	4.66	-.09
Impulsief	6.70	7.18	-.48
Enthousiast	3.67	3.39	.28
Voorzichtig	5.54	5.34	.19
Innovatief	3.68	5.55	-1.87 **
Wantrouwend	4.33	4.05	.28
Fantasierijk	3.67	3.29	.40
Abstract denkend	3.57	3.82	-.24

IV Manier waarop route wordt uitgelegd

	Ouderen	Jongeren	Mean Difference
Goed	4.09	5.21	-1.12 *
Prettig	4.38	4.92	-.54
Beleefd	6.49	3.61	2.89 **
Natuurlijk	5.04	4.53	.52
Houterig	5.60	5.82	-.22
Ontspannen	5.68	3.95	1.73 **
Energiek	5.53	4.71	.82 **
Statisch	3.86	4.47	-.61
Zorgvuldig	6.77	3.58	3.19 **
Uitbundig	3.75	5.74	-1.98 **
Kalm	6.43	3.16	3.27 **
Geïnteresseerd	5.00	4.47	.52

V Mening over routebeschrijving

	Ouderen	Jongeren	Mean Difference
Langdradig	4.66	4.05	.60
Ingewikkeld	3.68	3.82	-.14
Makkelijk	4.12	6.03	-1.90 **
Boeiend	3.47	6.05	-2.59 **
Gestructureerd	6.65	4.08	2.57 **
Bruikbaar	3.25	4.45	1.2 **
Duidelijk	5.34	4.66	.68
Begrijpelijk	5.44	5.63	-.19

VI Mening over gebruikte agent

	Ouderen	Jongeren	Mean Difference
Goed	5.50	4.00	1.50 **
Modern	5.59	5.95	-.36
Realistisch	5.61	4.08	1.53 **
Geavanceerd	5.24	4.18	1.06 *
Bruikbaar	4.92	6.24	-1.32 *
Innovatief	4.64	5.16	-.51

VII Informatie herinnering

	Ouderen	Jongeren	Mean Difference
Wendingen	3.20	3.63	-.42
Oriëntatiepunten	3.25	4.50 **	-1.25 **
Totaal	6.45	8.13 **	-1.68 **

B Mens – ouderen vs jongeren

I SAM-emoties (het gevoel dat opgeroepen wordt)

	Ouderen	Jongeren	Mean Difference
Valence	5.94	4.83	1.11 **
Arousal	4.94	6.13	-1.19 **
Dominance	4.45	5.45	-1.00 **

II Persoonlijkheid - overtuiging

	Ouderen	Jongeren	Mean Difference
Deskundig	6.19	5.05	1.14
Geloofwaardig	6.26	6.60	-.34
Realistisch	5.74	5.98	-.04
Betrouwbaar	5.10	6.05	-.95
Vriendelijk	6.90	6.43	.48
Sympathiek	6.84	6.18	.66
Dominant	5.52	5.53	-.01

III Persoonlijkheid (Cattell)

	Ouderen	Jongeren	Mean Difference
Eigenzinnig	4.39	5.05	-.66
Emotioneel	6.52	6.58	-.06
Koel	4.03	4.35	-.32
Gevoelig	4.10	4.78	-.68
Open	5.06	4.90	.16
Sociaal	5.00	4.65	.35
Gespannen	4.10	4.65	-.55
Onzeker	2.87	3.70	-.83
Opportunistisch	5.58	5.13	.46
Impulsief	7.03	7.03	.01
Enthousiast	3.77	3.18	.60
Voorzichtig	5.65	5.10	.55
Innovatief	4.03	4.15	-.12
Wantrouwend	4.26	3.78	.48
Fantasierijk	3.32	3.13	.20
Abstract denkend	3.71	4.45	-.74

IV Manier waarop route wordt uitgelegd

	Ouderen	Jongeren	Mean Difference
Goed	5.10	5.30	-.20
Prettig	5.10	4.70	.40
Beleefd	6.97	3.78	3.19 **
Natuurlijk	5.77	5.13	.65
Houterig	5.16	5.28	-.11
Ontspannen	6.13	4.65	1.48 **
Energiek	5.58	5.25	.33
Statisch	4.06	3.63	.44
Zorgvuldig	7.42	3.63	3.79 **
Uitbundig	3.84	5.98	-2.14 **
Kalm	6.94	3.13	3.81 **
Geïnteresseerd	5.29	5.18	.12

V Mening over routebeschrijving

	Ouderen	Jongeren	Mean Difference
Langdradig	4.52	3.30	1.22
Ingewikkeld	3.42	3.33	.09
Makkelijk	3.81	6.43	-2.62 **
Boeiend	3.71	7.00	-3.29 **
Gestructureerd	6.58	4.45	2.13 **
Bruikbaar	3.71	4.08	-.37
Duidelijk	6.00	5.00	1.00
Begrijpelijk	5.32	5.28	.04

VI Mening over gebruikte agent

	Ouderen	Jongeren	Mean Difference
Goed	5.58	4.13	1.46 **
Modern	6.03	6.00	.03
Realistisch	5.74	5.00	.74
Geavanceerd	6.16	5.08	1.09 **
Bruikbaar	6.00	5.53	.47
Innovatief	5.29	5.35	-.06

VII Informatie herinnering

	Ouderen	Jongeren	Mean Difference
Wendingen	3.22	3.52	-.30
Oriëntatiepunten	3.45	4.05	-.60
Totaal	6.67	7.57	-.90