Master Degree Project

UNIVERSITY OF SKÖVDE

AUDIO-NAVIGATION

What are the benefits of utilising diegetic spatial audio in audio-navigation software?

Master Degree Project in Informatics
30 ECTS
2012

Per Anders Östblad

Supervisor: Henrik Engström

Examiner: Mikael Johannesson

Abstract

The task of navigating and orienting oneself in an unknown environment might seem arbitrary to most people. For a person with a visual impairment however, it can prove to be a challenge. Much research has been conducted to provide useful software solutions to aid this problem. Nevertheless, not enough research has been put into studying the use of our everyday sounds for such software. The present thesis is aimed at evaluating the usefulness of utilising diegetic spatial audio. After a real environment was replicated in a virtual audio environment without visual input, 16 sound students from the University of Skövde were asked to perform the same set of tasks in the virtual and the real environment. The results indicated that diegetic spatial audio is an important part of navigating without vision and that the experiment would be interesting to perform on a larger scale with visually impaired participants.

Key words: diegetic audio, visual impairment, audio-navigation, spatial knowledge, spatial cognition, sound design

Acknowledgements

Thank you to Mårten Jonsson for a very interesting and mutually beneficial (and destructive) collaboration.

Thank you to Henrik Engström, Mikael Johannesson and Per Backlund for a really interesting and enlightening year and special thanks to Henrik Engström for all the help as supervisor.

Table of Content

1 Introduction

Being able to navigate our environment is something most of us take for granted. However, for a person with a visual impairment, it can be quite challenging and may require a lot of help and practise. That blind people are better at using their hearing to navigate their environment may seem obvious. However, after performing a thorough literature review into the subject of audio-navigation and spatial cognition amongst visually impaired people, it would seem that not enough focus has been put into discovering how to use this (possibly) more acute sense of hearing, in software aimed to help blind people with orientation and mobility. The research presented in the present thesis was aimed at examining the usefulness of focusing on the everyday sounds that surround us for navigation-training software for visually impaired people and to increase awareness about sound as a way of navigating.

In chapter 2, the motivations for the present study will be made clear and an overview of what research has been, and is being done on this topic will be presented. In chapter 3, the problem on which the present thesis is based, as well as the method that was used for experimentation, will be discussed. Chapter 4 outlines the experiment and the preparations leading up to it and all the collected data will be presented and analysed in chapter 5. Lastly, in chapter 6, the results of the study will be summarised and discussed from a wider perspective.

The aim of the present study was to perform a pilot study, in collaboration with Mårten Jonsson, evaluating the viability of diegetic spatial audio in audio-navigation software aimed at helping blind people. Based on two hypotheses, an experiment with 16 sound students from the University of Skövde was performed. With the use of a virtual environment called *Audiction,* modelled after a real environment, the participants were told to perform a set of tasks in *Audiction* and then to perform the same tasks in the corresponding real environment. In parallel with *Audiction*, Mårten Jonsson created a similar application but with functional audio rather than the diegetic spatial audio *Audiction* was created to evaluate.

2 Background

For a person with normal vision, orientation and navigation are fairly simple tasks. For a blind person, however, it can prove to be quite a challenge to get to know one's environment and learn to navigate it. In contrast, a blind person could be better than a sighted person at many tasks depending on tactile or auditory stimuli. Dykhoff (2002) states that the human brain has limited bandwidth available for sensory information and that the largest portion of this bandwidth is being used up by vision. If, however, there is no visual input, there is a larger amount of bandwidth available for other senses. Similarly, Sánchez et al. (2010a) report that, when asked to move along a predetermined route, blind testers activated parts of their brain normally associated with route visualisation, e.g. the visual cortex. These signs of adaptability of the brain might suggest that it is possible to train one's senses to be more acute and finely tuned. Furthermore, it may be an indication that blind people can compensate for their lack of vision with other senses, for example hearing, as suggested by Copeland (2000).

Similar indications have been observed in a study of listening practise (Östblad, 2011). Based on Chion's (1994) theories of listening modes, an experiment was conducted suggesting that people who work with sound might be better at analysing and identifying audio material than those who have little experience working with sound. It also indicated that visual stimuli may have an impact on how we interpret and identify sound material.

Sánchez et al. (2001, p.65) report: "Blind children tend to represent spatial environments with cognitive difficulty". The authors believe one solution to this issue would be to compensate for the lack of visual input by stimulating other senses. Moreover, the authors state that blind children are able to create mental maps of an environment based only on sound. Along the same lines, there is research suggesting that audio-based interfaces may work equally well to convey information about spatial structures as those that are visually-based (Oren et al., 2008). Roden and Parberry (2005) highlight the importance of the player having a clear mental image of the game world in an audio-game. Similarly, Kamel and Landay (2000) emphasise how imperative it is for a blind person's conception of the world to accurately represent graphical material in a comprehensive way.

Sanchez et al. (2010a) state that blind people often prefer to follow a wall when walking through a room. As it is not easy for a visually impaired person to determine the size of a room, it might be useful to be able to convey that information.

2.1 Concept Definitions

To be able to properly describe the problem area, there are a few basic concepts that need to be addressed and defined: virtual environment, diegetic sounds, spatial audio, spatial knowledge.

2.1.1 VE (virtual environment)

There have been quite a few serious games and VE's (virtual environments) created to research navigation issues for visually impaired people (Wallet et al., 2008; Lahav &

Mioduser, 2004). The term "game" can often spark a discussion about what qualifies as a game and the definition of games. To avoid that discussion, which is irrelevant to this project, the term VE will instead be used. VE is defined in the present study as a computer-based software-tool created to represent a fictional and/or real environment, which the user can explore using a compatible means of control, i.e. a computer keyboard. Although a VE may not be considered a game, the act of using a VE may in some cases in the present work be referred to as "playing".

2.1.2 Diegetic Sounds

To be able to properly describe the VE presented in chapter 2.2, there are a number of sound-concepts that need to be addressed. The first of these concepts is diegetic versus non-diegetic sounds; these are from computer game theory (Ekman, 2005). Ekman refers to diegetic sounds as, sounds that originate from the game world, for example the sounds of birds, footsteps or weapons. Non-diegetic sounds, on the other hand, are sounds that do not originate from within the game world, e.g. music, spoken instructions or a narrator-voice. Of course, a narrator voice or the music could be considered to be diegetic, say if the music track was being played from a radio visible in the game. If it does not have a clear origin within the game world, however, it would be considered as being non-diegetic. The VE described in chapter 2.2 makes a point of only utilising diegetic sounds.

2.1.3 Spatial Audio

Many of the VE's that have been created to test blind user navigation have utilised spatial sound as a navigation aid (Sanchez et al., 2010a). In this project, spatial audio denotes sounds that are played from a physical location within the VE and that attenuate according to the user's position and direction, as sounds would in real life. In the present work, all spatial audio will be in the form of diegetic sounds, i.e. sounds that can be found in the corresponding real environment.

2.1.4 Spatial Knowledge

For a blind person to be able to navigate an environment safely, he or she would most likely stand to benefit from having a good, spatial mental understanding of said environment. The concept of spatial mental understanding is discussed in many different terms, for instance: "mental maps", "cognitive maps", "spatial orientation", "spatial mental models", "spatial awareness", "spatial knowledge" and "spatial cognition" (Heuten et al., 2008; Quiñones et al. 2011; Tversky, 1993; Wallet et al., 2008). Although there might be distinctions, a reasonable summary of these concepts may be as follows: "Spatial cognition refers to the cognitive processes associated with the development of a comprehensive understanding of a 3D environment and the utilization of that knowledge for various purposes." (Wallet et al., 2008). The present study is focused more on the utilisation of spatial knowledge, rather than the cognitive processes of creating mental models.

2.2 Audiction

During 2011 Jonsson and Östblad have developed a software tool, called *Audiction,* to function as a means for creating audio-based VE's and games. *Audiction* was made to work with both PC and Mac. It features no graphics, but instead relies on spatial audio. It features a first-person character, controllable using a standard keyboard, with support for footstep sounds and physics-collision. It also includes a system for easy placement and attenuation of spatial audio and real-time effects. Using this engine, a VE based on a fictional office environment was created to test the functionality of the engine. The guideline for this project was to use only means of navigation that could be utilised in real life; thus, only diegetic sounds were allowed. In addition, no unrealistic aids such as fixed 90-degree turns or spoken instructions were used.

2.2.1 Features

During the development process, several tools for audio-navigation and spatial knowledge gain were implemented. These tools were:

- Reverb-zones – This was implemented to make different rooms reflect sounds differently according to their size.
- Floor materials – The material of the floor specifies how the footsteps of the character will sound.
- The virtual cane – A virtual cane was implemented, with which the user can examine his or her proximity and tap objects and walls.
- Object materials – Objects in the VE may play fitting sounds when tapped with the virtual cane.
- Wall sounds – After testing with realistically-sounding wall materials, it was decided to make all walls sound the same when tapped with the cane. This was considered to fall within the guide-line "no unrealistic aids", because it mimics the sense of touch a blind person could utilise when navigating a similar office.
- Static sounds – The initial purpose of *Audiction* was to research whether diegetic sounds could convey useful spatial information, thus increasing a player's spatial knowledge of the VE. For that reason, one of the most important tools to include was static sound sources, i.e. diegetic, spatial audio, aimed at replicating everyday objects that make noise, for example, a kitchen fan or a wall-clock.

2.2.2 Results

During the development of this VE, a pilot study was conducted. Initial test results indicated that the blind people playing the VE seemed better at using the spatial audio to navigate than the sighted players. They also suggested that by playing this VE, one could learn to find one's way around the virtual office. Furthermore, the results indicated that it might be possible to improve one's audio-navigation skill and the conclusion was that further research on this subject would be interesting. Most importantly, the tests showed that the tools implemented in the engine all appeared to be working as intended for the most part. However, a few problem areas were discovered, the most alarming being how to communicate the speed of walking and turning. There were some problems during tests, where test-subjects would tap the walk- or turn-button, instead of holding it down. This resulted in very slow movement, leading to confusion as to how far one had travelled or turned. The solution to this problem might simply be to better instruct the user on how far he or she moves with every step.

2.3 Related Research

There have been quite a few research projects aimed at identifying blind people's navigation issues and how to solve them. This research has been going on for many years, but with the emergence of GPS, mobile phones, Internet and advanced computer games and audio-engines, there are more opportunities than ever to make a difference. The creators of a narrative-based audio-game (Archambault & Olivier, 2005) proclaim that, even though it might be costly to create games for blind children, it is very important for their sense of inclusion.

In the late 80's, Passini and Proulx (1988) conducted a successful experiment on way-finding without vision. After two guided tours, the authors let 15 totally blind and 15 sighted subjects perform the same navigation tasks in a complex route, following it up with cognitive mapping tasks and interviews. The results showed that many of the blind subjects were able to navigate the route without errors and, more interestingly, demonstrated good spatial knowledge of the setting. Passini and Proulx also reported some interesting differences in the decision making of the sighted and blind groups. For instance, the blind group tended to prepare for the task more carefully. Along the same lines, Schneider and Strothotte (2000), propose that re-constructing an environment, with the help of instructions, a tactile interface and a spatial computer system, could help both sighted and blind people to get to know an area and its routes.

A great deal of the research that has been made on blind navigation has suggested that by stimulating other senses, it is possible to convey spatial information. Sánchez et al. (2001) state that blind children can create mental structures based on sound alone. They suggest the use of computer programs to help blind children explore environments and thereby increasing their spatial knowledge. Masuch and Röber (2004, p.6) claim that when designing audio-based VE's: "interaction techniques have to be as natural as possible to closely mimic the process of natural hearing. This results in a deeper immersion into the virtual environment and allows a better perception of the conveyed information".

A common practice when researching blind navigation is to create VE's with both haptic and audio feedback. One such system was created by Lahav and Mioduser (2004). They reported interesting findings with regards to the problems blind people might have with mobility due to their lack of vision. They based their work on the assumption that it is possible to increase blind people's spatial knowledge by stimulating other senses. Their test subjects were told to play in a VE modelled after a real space and then visit the real environment. The blind participants showed a greater confidence when visiting the real setting, after having explored the VE. A similar approach can be found in the work performed by Sánchez et al. (2010b). With the help of an audio-game called *MOVA3D* in combination with a haptic control-mat, the authors researched the possibility to increase blind children's spatial knowledge by first playing in a VE and then in a corresponding real setting. Encouraging results showed that the children were able to achieve successful transfer of spatial knowledge from the VE to the real environment. Sanchez et al. believe virtual training is a very good way for blind children to practise navigation tasks.

Perhaps the attempts most aimed at creating usable navigation-tools for blind people have been those that utilize mobile devices. Paterson et al. (2010), for instance, have created a location-aware mobile game. They highlight the necessity for well-designed audio to achieve immersion and engagement. The authors believe that location-aware games may be a useful way to convey spatial knowledge of an environment. Similarly, Magnusson et al. (2011) have presented a game they call *NIVINAVI*. It features GPS positioning and spatial audio-navigation. While walking about in the real world, the player receives audio cues via a mobile device about where to go and what tasks to perform. Magnusson et al. performed tests with blind children and the conclusion was that the game was fun, engaging and useful. Yang et al. (2011) address the problem of accessibility for blind travellers. With the help of mobile, location-based software providing information about the environment and how to access it, the authors suggest that an increase in spatial knowledge may be achieved, thus helping blind people with way-finding.

In conclusion, all of the attempts presented in this chapter had one thing in common, namely to increase the quality of life for blind people. Whether it is with the help of a location-aware mobile game, a haptic control mat, virtual audio reality or with GPS guidance, all researchers aim to find the solution to blind people's navigation issues. There is, however, one aspect that warrants more focus in this field of research: spatial audio. This will be discussed further in chapter 3.

3 Problem

Great deals of the research methods that have been presented on the subject of audio-navigation have shown good results. Yet, there is one area in this field that remains fairly unexplored: spatial audio, and in particular diegetic sounds. Many of the projects presented in the present study have been similar to *Audiction*, and its purpose, in many respects. However, although they may utilise spatial audio they more often than not contain additional tools for orientation and spatial representation. For instance, grid systems, speech cues, non-diegetic sound cues, tactile and haptic controls and set compass points. All of these tools are certainly viable when designing tools to increase the quality of life for the visually impaired. Nevertheless, there might be some valuable information to be gained from exploring spatial audio in this type of software in detail. Loomis et al. (1998, p.202) argue that “spatialized sound will be more effective in conveying this information than synthesized speech”, referring to information accumulation by blind travellers of the content of their surroundings. Along the same lines, Friberg and Gärdenfors (2004) suggest that designers of audio-games should pay more attention to aesthetics to increase immersion, thus increasing the entertainment value. It is arguable that increased immersion might lead to an increase in learning as well. Blind people and sighted people do not experience the same things in their everyday lives. It would be a fair assumption, that it is a more conscious and deliberate act for a blind person to learn a room’s size and layout than for a sighted person. Therefore, it would be desirable to make spatial knowledge accumulation for blind people into a more subconscious act. One way to do this might be to utilise diegetic spatial audio to a greater extent in software, than traditionally has been the case. Most importantly though, highlighted by Sánchez et al. (2010c), it is imperative to respect the needs and wishes of the target audience, in this case the visually impaired community.

Following the assumption that blind people, does in fact, have a more developed and acute sense of hearing, it seems that fact should be considered when designing software aimed at helping them. It might appear redundant when performing tests on spatial cognition and way-finding in a virtual environment or in a controlled test environment. However, if the purpose is to create useful navigation training software, to be used by blind people in their everyday lives, the question of diegetic spatial audio should be addressed. There are, after all, no artificial grid systems or speech-cues in the real world. Furthermore, it might be dangerous for a blind person to have headphones on, listening to instructions from a GPS or mobile game, when navigating in traffic. When walking around at work or in school, a blind person needs to rely on what he or she can hear, smell and feel. If it is possible to train one's hearing to be more acute, or at least learn what to listen for, then maybe it would be a good idea to practise using the sounds that surround us in our daily lives to help us navigate. Particularly for children or adults who lose their sight late in life, who are not used to navigating their environment without visual input, practising audio-navigation in a VE would seem like a good option.

3.1 Aims

The aim of the present research project was to evaluate the viability of diegetic spatial audio, in navigation aid software for blind people, and how it affects spatial knowledge transfer from a VE to a real corresponding environment.

To evaluate this problem, the following hypotheses were tested:

Hypothesis 1: *The group playing Audiction, modelled after a real environment and utilising diegetic spatial audio, will have better spatial knowledge of the environment, and will therefore be better prepared for the real world test than the other groups.*

Hypothesis 2: *Participants that have played Audiction, modelled after a real environment and utilising diegetic spatial audio, will be able to achieve spatial knowledge transfer from the VE to the real environment.*

3.2 Method

In short terms, the experiment contained one session of navigation tasks in a VE, and one session of the same navigation tasks in a corresponding real environment. Inspired by Passini and Proulx's (1988) successful research on audio-navigation, these navigational tasks were followed up by map drawing tasks to help assess the spatial knowledge of the participants. These sessions will be described in greater detail further on. To achieve the aim of the present project there were several methodological aspects that needed to be considered.

Before testing could begin there were a number of objectives that needed to be fulfilled. The first objective in designing the experiment was to define how to measure spatial knowledge, as well as specify the criteria for when spatial knowledge transfer has been achieved. After doing that, the second objective was to perform a pre-test of the experiment with a small number of participants to identify flaws in the test design.

In parallel with the VE presented in the present thesis, Mårten Jonsson designed a contrasting VE, one that was aimed at evaluating functional, non-diegetic sounds, rather than diegetic spatial sounds. In order to collect useful data, these two VE's were tested in the same environment and on the same premises. Modelling these VE's after the same setting allowed for a comparison of data, which would hopefully show on positives and negatives with both applications. Moreover, to get even more reliable results, a control group that did not play any of the VE's was also evaluated.

To get as much useful information as possible out of the experiment, a combination of qualitative and quantitative methods were utilised (Bryman, 2002). The quantitative side of testing consisted of collecting data during test sessions about right and wrong navigation. Specifying criteria for what was considered right and wrong navigation was therefore a required prerequisite. The qualitative side consisted of questionnaires and subjective evaluations of the performances of the participants (Bryman, 2002). To gather additional qualitative information during the tests the concept of "undisguised participant observation" (Shaughnessy et al, 2005) was used.

3.2.1 Ethical considerations

Since the present experiment was performed with human participants there were some ethical issues that had to be considered. Therefore, the APA Ethical Standards (Shaughnessy et al., 2005) for psychological research was consulted on matters of participant safety, privacy and informed consent.

3.2.2 Experiment

An important step in designing the experiment was to make alterations to *Audiction,* in order for it to suit the experiment. By recording sounds in the chosen environment and implementing them in *Audiction,* the aim was to create as realistic a representation of that setting as possible. It had to be an environment which contained some static sound sources or there would be no sounds to translate into diegetic spatial audio in the VE.

The second step was to design the actual test. As test subjects a number of sound students from the *University of Skövde* were used. The plan was to let a number of participants play the VE (a second group played Jonsson's application) and complete a set of missions, for instance collect three items, and then let them complete the same set of missions in the real environment, all the time without visual input. The aim was then to perform the same, real-world test, with a control group that did not play any of the VE's, but received alternative instructions about the environment, e.g. text-based. For the data collected from the different groups to be comparable, it was important that the instructions they received was analogous. If the control group were allowed to study the text-based instructions for the same amount of time as the other group were allowed to play the VE, it seemed plausible that their results would be comparable. A text-based form of instruction also seemed to be the best alternative due to its translatability into braille, in case the experiment will also be performed with blind testers.

The present experiment was designed to investigate whether the VE based on diegetic spatial audio was better at conveying spatial information than the other tested means of spatial instruction.

The last step in this experiment was to analyse all collected data, and compare them to Jonsson's collected data. As mentioned earlier, Jonsson conducted similar research in parallel, but with a navigation tool aimed at testing functionality rather than realistic spatial audio. All tasks and locations in the test sessions were identical for the two research projects, thus the data would be comparable and hopefully better conclusions could be drawn. In co-operating and comparing results with Jonsson, the hope was to identify key aspects of navigation amongst blind people as well as how they orient themselves in their environment. Identifying these key aspects would serve as a good basis, on which to create a framework for the development of useful and creative navigation aids for the visually impaired community. This could possibly lead to opportunities for both continued research and commercial game development.

3.2.3 Delimitations

For practical reasons the tests were not performed with blind testers. Instead, sound students from the *University of Skövde* were used. Blind participants would also be invited to join the test session if time permitted. This was not considered to be a major problem however, as it is fair to assume that if a sighted person with a blind fold can perform audio-navigation tasks, so can a blind person.

To avoid being overwhelmed, the present research was treated as a pilot study aimed at creating a good base for further research. Moreover, to be able to perform the tests with as many participants as possible, the experiment and instructions were kept fairly simple and to the point.

3.2.4 Expected results

The expected results from the tests were somewhat dependant on the success of mimicking a real environment in *Audiction*. If that task was successful, it was considered plausible that the participants playing *Audiction* would perform better in the real-world test than the control group. Moreover, it was highly likely that *Audiction* and the VE created by Jonsson would bot show some good and some bad features respectively. Analysis of the data and the interviews, would hopefully highlight which features were useful and not for future co-operative research and game development.

4 Experiment

The following chapter contains a description of the preparations and finalisation of the pilot study performed by Jonsson and Östblad.

4.1 Preparation

The first step of the preparation for the pilot study was to model the software application *Audiction* after a real environment. To do this, a suitable setting had to be selected. With the help of a teacher, an apartment that met the set requirements was found. There were a couple of criteria that this apartment fulfilled: It was fairly well isolated from outside noise, it was of manageable size, it had a good layout with several rooms all connected to the entrance hallway and it contained a number of sound sources that could be used as diegetic sounds in the VE.

After all the measurements of the apartment had been taken and a sketch of the layout had been made, all rooms and sound sources were recorded. A four-channel surround recording was made in each room to document the ambiance and the reverb in the different areas of the apartment.

After collecting all sound samples, the next step was to modify *Audiction* to replicate the apartment (Figure 1). After the layout had been created and all relevant sound sources had been placed, reverb zones were created for each room to match the ambiance of the real environment. All objects, such as tables or chairs, were given realistic sound effects for when being tapped with the virtual cane. The floors were given specific foot step sounds to match the material of the corresponding real floors. Lastly, three positions were decided upon and implemented, with voice feedback for when the player of the VE found them: one in the bedroom, one in the living room and one in the kitchen. The diegetic sounds in the apartment, replicated in *Audiction,* were:

- Kitchen fan/timer over the stove.
- TV in the living room.
- Small stereo playing music on the desk in the bedroom.
- Bathroom fan.

In addition to these diegetic sound sources, each room contained its own ambiance track. There was also an ambiance track for the whole apartment, played on very low volume. With both *Audiction* and Jonsson's application finished, the next step was to design the experiment itself.

Figure 1 Overview of the apartment in *Audiction*. 1. The player's starting position. 2. The desk in the bedroom. 3. The bookshelf in the living room. 4. The backpack on the table in the kitchen.

4.2 Pre-test

In order to test the quality of the experiment, a pre-test with three teachers at the University of Skövde, was performed. Chapter 4.2.1 will outline the pre-test in chronological order.

4.2.1 Execution

When the three participants arrived at the location, they were asked to close their eyes before entering the apartment. They were then led into a room which was not included in the VE and therefore not in the real-life task either. Hence, they would not have any preconceptions about the apartment when participating in the test. They were divided into three groups: one group for *Audiction*, one group for *Audionome* (Jonsson's VE) and one control group. The control group was told to read detailed instructions about the layout of the apartment and the positions they were going to locate.

After receiving some basic instructions about how the test was going to work, they each got five minutes of instructions about the group-specific task they were going to perform. For *Audiction* and *Audionome,* the participants were told to listen to pre-recorded instructions about how to play the VE's and which three positions they were supposed to try to find. The control group was given a written instruction that outlined what tasks to perform. After receiving these instructions, ten minutes of play-time followed.

The groups playing the VE's used headphones and did not receive any visual input. *Audiction* was played on a standard PC keyboard and *Audionome* on a smartphone with a touchscreen. The control group studied the written instructions. After the ten minutes of reading or exploration of the VE's, the participants were asked to fill out questionnaire 1 (Figure 2).

1. How many of the 3 positions did you find?
2. How many rooms did you find?
3. Draw the rooms in relation to each other (don't worry about scale).
4. Grade the rooms in your drawing from biggest to smallest, 1 is the biggest.
5. Mark the positions (with an X) of the locations you were asked to find, in your drawing.
6. How well would you say you know the layout of the apartment? (Circle you answer)
   Not at all 1 2 3 4 Very well

Figure 2 Questionnaire

This questionnaire was aimed at assessing the test subjects' spatial knowledge of the apartment after performing their first tasks. Throughout the whole experiment, where participants were asked to grade questions, a 1 – 4 scale was used. The reason for this was that, with the few participants available for testing, it would hopefully show slightly more tendencies towards one direction or another than, for instance, a 1 – 5 scale.

The next step was to perform the real-life test. While two groups listened to music with headphones on, the third was lead to the entrance hallway and placed right at the front door, as was the case in the respective VE's. Wearing a blindfold, the participant was asked to explore the apartment and find the three positions (the desk in the bedroom, the bookshelf in the living room and the stove in the kitchen) as fast as possible, while being careful. When a position was believed to have been located, the subjects were asked to state so verbally. When one person had found all three positions or a maximum of ten minutes had passed, that person was led back into the spare room and the next group were told to perform the real-life test. During this part of the test, Jonsson and Östblad independently assessed the participants' performances and their confidence in navigating the environment. Since these assessments were going to be subjective, it was decided that notes would be taken independently and then compared in order to be as reliable as possible.

When all groups were done, they were asked to fill out the same questionnaire one more time. The same questions were used in order to see if there had been an increase in spatial knowledge before and after the real-life test. If so, which group had the best spatial knowledge after playing a VE/studying the written instructions?

After filling out the second questionnaire, the participants were allowed to see the apartment for the first time. They were shown the different rooms and where the three positions were located. They were then asked to answer a follow-up interview (Figure 3) with control questions.

1. Did the layout of the apartment match your expectations? (Circle your answers)
2. What tools did you use to perform the real-life tasks? (Circle the letters, Multiple answers allowed)
   A. I listened to sounds in the apartment
   B. I remembered the layout from the VE
   C. I counted footsteps
   D. I felt my way around
   E. I had no idea where I was and just walked around
   F. Other

Figure 3 Follow-up interview for the pre-test

After the pre-test, our teachers were asked for their input on how the test had been.

4.2.2 Results

A great deal was learned about the execution of the experiment from the pre-test and from suggestions from the three participants. The most important change to make for the upcoming study was to remove the written instructions for the control group. It was decided that it would be a more useful method of control to just let the control group explore the apartment without any training/instructions at all. That way, it would hopefully show if any spatial knowledge transfer from the VE's to the real-life test occurred.

Another change that was made for the pilot study was to translate everything into Swedish, including verbal instructions and all questionnaires and instructions. Since all participants were going to be Swedish, this was decided upon to avoid misunderstandings due to language. Following input from one of the teachers, a ranking aspect to question 2 in the follow-up interview was added. For the pilot study, it was also clear that three test subjects at once were too hard to manage. Therefore, only two participants at a time were invited to take the test.

4.3 Pilot study

The main task of the present thesis was to perform a pilot study to evaluate the method of testing and the usefulness of the application *Audiction*. Chapter 4.3 describes the execution and outcome of this pilot study.

4.3.1 Selection

Participants for the experiment were chosen selectively. Ideally, the test would have been performed with visually impaired participants. However, due to time constraints this was not an option. Therefore, sound students from the computer game development program at the University of Skövde were used. This was considered a viable option, following a hypothesis that if sound students are able to complete the test wearing a blindfold, so can visually impaired people. If anything, they would probably perform better than sighted people as they are more accustomed to navigating without visual input. Out of the 24 invited students, 16 performed the test. Out of the 16 participants, six played *Audiction*, six played *Audionome* and the remaining four constituted the control group. The reason for not having a randomly selected test group was to get a homogenous population i.e. to minimise the effect of random outliers. All participants reported having normal hearing and were all considered to have adequate computer and game experience. Only one woman participated and no effort was made to balance genders in the selection since it was not considered to be of importance for the present study. All students in the test group were between the ages of 20 and 30.

4.3.2 Execution

The pilot study was performed in roughly the same manner as the pre-test. The experiment was executed on two occasions with eight students each time, with two participants simultaneously. The participants were either in the VE group, i.e. the two students played *Audionome* and *Audiction* in parallel, or they were assigned to the control group and performed only the real-life test.

For the VE group, the experiment started off with the participants being greeted and brought in to the spare room with their eyes closed. There they received a brief introduction to the experiment and were assigned one of the VE's. Wearing headphones they then received spoken instructions about how to play the VE and that they were supposed to locate three positions in the VE. The three positions were: the desk in the bedroom, the bookshelf in the living room and the backpack on the table in the kitchen. In addition to the spoken instructions, they were given the same text on paper in order to be able to better grasp the instructions in the five minutes they had at their disposal.

After they had received their instructions, the participants were told to play on their own for ten minutes in their respective VE. If they got stuck or lost their way, they could easily reset their positions at the starting point. Even if they found all three positions, they still had to play the full ten minutes and explore and try to memorise the environment. When the ten minutes of play time were up, they answered questionnaire 1 (figure 2).

The next step was the real-life test. One participant was asked to listen to music wearing headphones while the other performed the test in the parts of the apartment represented in the VE. In order to avoid the results being affected by in which order they performed the real-life test, the starting order of *Audionome* and *Audiction* players were alternated throughout the experiment. Wearing a blindfold the person performing the real-life tasks was placed at the front door facing the living room, i.e. exactly the same starting position as in the VE's. Instructions were then given that the tester was supposed to find the three positions in the apartment within three minutes. Participants were encouraged to verbalise their thoughts as they explored the apartment. When they thought they had found a position, they stated so and continued without comment from the experiment leaders (Jonsson and Östblad). During the ten minutes of exploration, notes were taken independently by Jonsson and Östblad, and times and the number of errors was recorded. The number of errors was evaluated according to a list of set criteria, namely if participants:

- Entered a room that did not contain a position.
- Verbalised a thought that demonstrated they were lost.
- Left a room containing a position without finding it.
- Entered a room containing a position again after already having found it.

When both VE players had completed the real-life test, they were asked to fill out questionnaire 2, which was an exact replica of questionnaire 1. When they had done so, they were allowed to view the apartment for the first time and the test ended with the students filling out the follow-up interview (Figure 4).

1. Did the layout of the apartment match your expectations? (Circle your answer)
   Not at all 1 2 3 4 Very well
2. How important were the following aspects for your navigation? (Circle your answer)
   A. I listened to sounds in the apartment (1-4)
   B. I remembered the layout from the VE (1-4)
   C. I counted footsteps (1-4)
   D. I felt my way around (1-4)
   E. I had no idea where I was, I just walked around at random (1-4)
   F. Other (1-4)

Figure 4 Follow-up interview for the pilot study

The control group performed exactly the same tasks as the VE group but without playing a VE and did therefore not fill out questionnaire 1. Furthermore, alternative 2.B (I remembered the layout from the VE) was removed from the follow-up interview. Before the real-life test, the control group was given a brief introduction and told about the three positions and what to do. They did not receive any instructions on how to find the three positions or about the layout of the apartment.

In conclusion, the pilot study went very well and seemed to be easily understood. All the participants found the experiment to be interesting and enjoyable. The experience of navigating an unknown environment without sight seemed to be both scary and exciting.

5 Results

In order to evaluate all collected data, there were a few aspects of the study that had to be addressed. The subjective assessments of the participants' navigation, made by Jonsson and Östblad, were one such issue. To neutralise the subjective differences between scoring, the mean value for each evaluation, made by Jonsson and Östblad, was calculated. Since one of the main tasks of the present study was to compare the groups, all the results for each group were added and then divided by the number of participants, to get mean values. Furthermore, to get useful data out of the participants' map drawings, Jonsson and Östblad worked together with a 1 – 4 grading scale to score them.

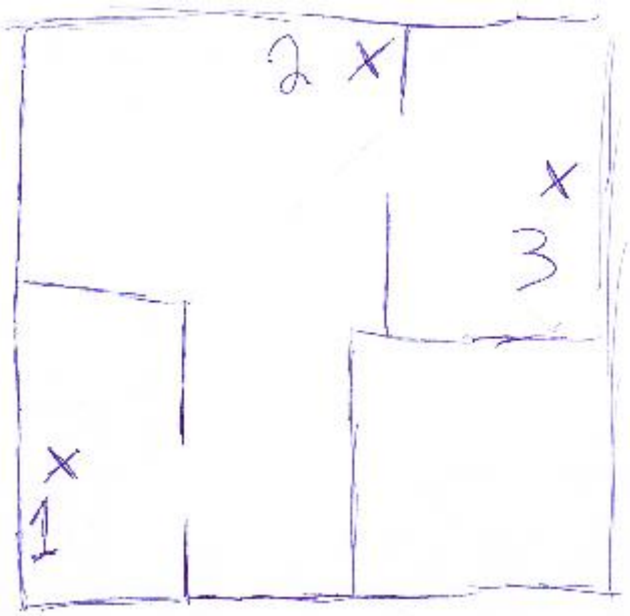

Figure 5 Example of map drawn in questionnaire 1, after playing *Audiction*

Although the drawing shown in Figure 5 does not have any clear indications about which room is which, it is quite clear that the participant has not received a very good overview of the rooms in relation to each other. The representation of room sizes is also quite poor; however, the positions are quite accurate in relation to door positions and in relation to each other. The score for this drawing was (on a scale from 1 – 4:

- Map score - 2
- Room size score – 1,5
- Position score - 3

Unfortunately the instructions for drawing maps were too vague, therefore a number of maps were not clear enough to score and had to be discarded. These maps could not just get a low score, because it was not clear enough if they were good or bad or if they had understood the instructions.

After all subjective measurements had been translated into scores, all data was then inserted into tables (Appendix A) and made into diagrams. Total time was converted to minutes in decimal form and rounded to the nearest quarter minute, i.e. four minutes and seventeen seconds would be expressed as 4,25.

There was one participant in group 1 who might be considered to be an outlier, and should perhaps therefore be excluded from the results. This person seemed to have a harder time navigating the environment than the other students. The participant made eight errors and only found two out of three positions within the ten minutes. All other participants found the three positions within the time limit and no one else made more than five errors. However, the outlier will not be excluded in the diagrams in the following chapters since most of the results were not as deviant. Where the deviations might have a big impact on the results, this will be discussed.

5.1 Comparison between Groups 1 and 2

First to be examined are the results of groups 1 (*Audionome*) and 2 (*Audiction*). *Audionome* was play on a smartphone with touch screen, from a top down perspective with only functional audio rather than diegetic audio. *Audiction* on the other hand was played from a first-person perspective and only contained diegetic audio. Both VE's were played completely without graphics.

5.1.1 Real-Life Test

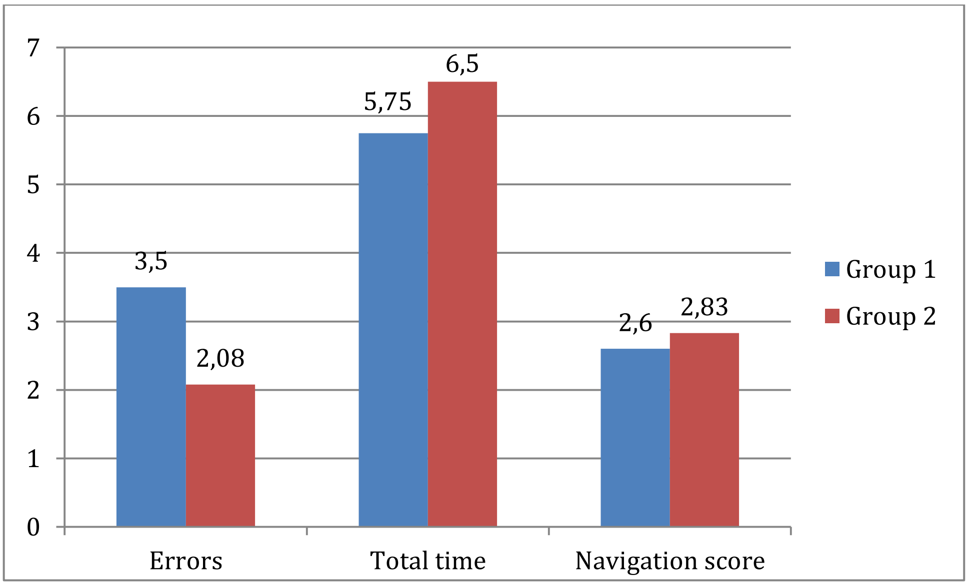

Diagram 1 Scores recorded by Jonsson and Östblad during the real-life test

As seen in Diagram 1, group 2 made fewer errors on average than group 1 during the real-life test. Group 2 needed more time to complete the objectives but they also received higher navigation scores from the experiment leaders. With the outlier excluded, the results for group 1 are: 2,7 errors, 5,0 total time and 2,87 navigation score. Only the navigation score changes slightly in strength ratio when excluding the outlier.

5.1.2 Questionnaire 1

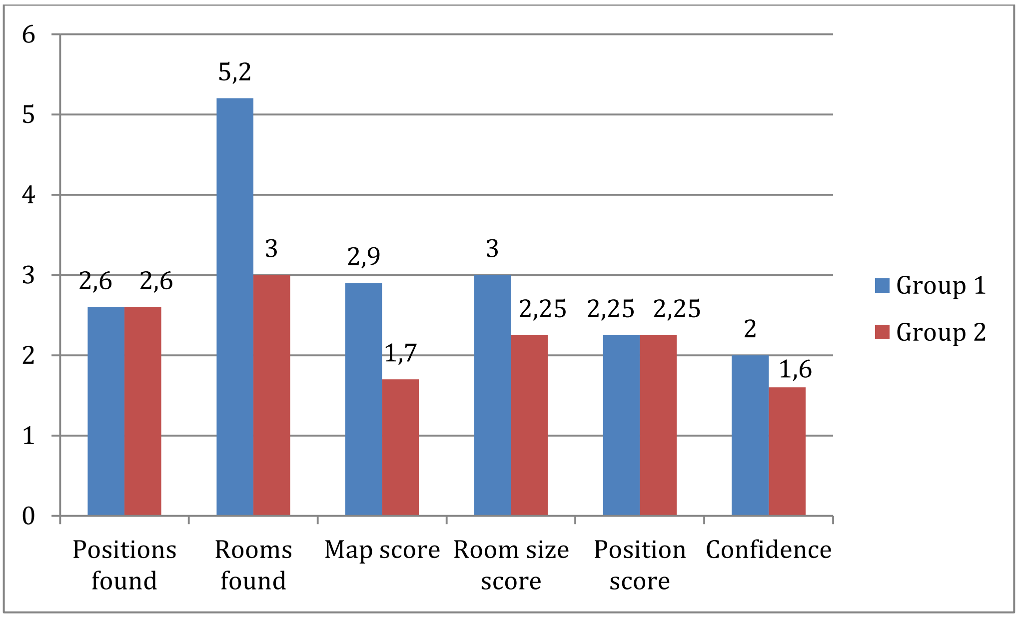

Diagram 2 Scores recorded in questionnaire 1

In questionnaire 1 (Diagram 2), it is clear that the groups found the same amount of positions when they played the VE's. Group 2 found significantly fewer rooms however. Group 2 scored a lot lower at drawing the general layout of the apartment and they also scored lower at determining the sizes of the rooms. When asked to mark the positions in their drawings, the groups performed equally. Group 2 showed a little less confidence in their spatial knowledge of the apartment after only playing the VE.

5.1.3 Questionnaire 2

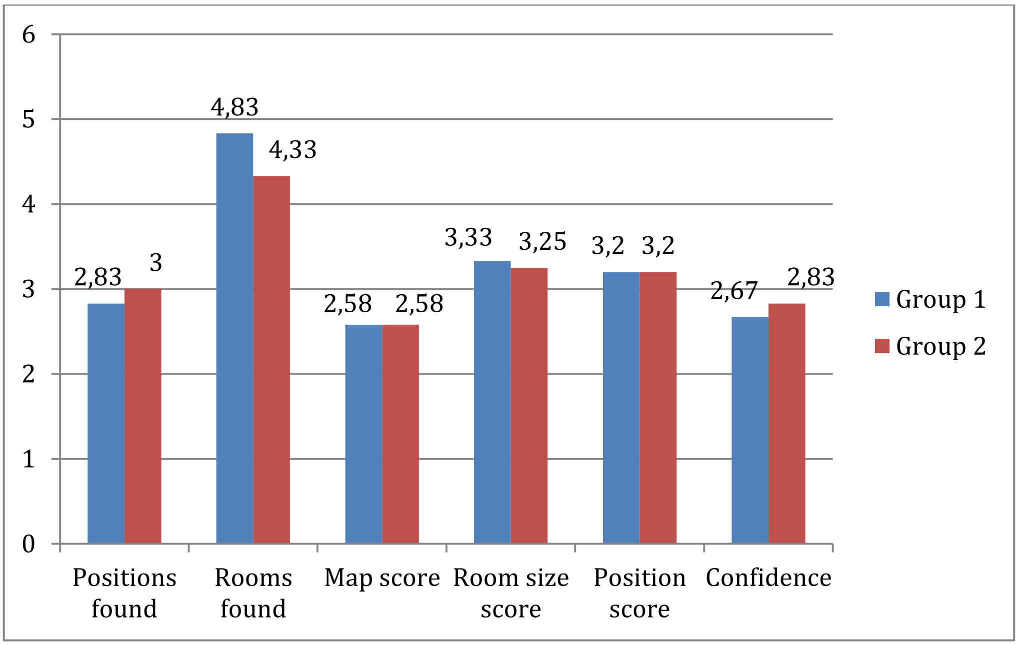

Diagram 3 Scores recorded in questionnaire 2

When filling out questionnaire 2 after performing the real-life test, the results were more equal (Diagram 3) between the two groups. Group 2 were a little better at locating the positions, but without the outlier included in the results, the groups performed equally. Group 2 found slightly fewer rooms on average than group 1. The scores for the three map drawing tasks were equalised after the real-life test, and group 2 showed a greater increase in confidence than group 1 compared to questionnaire 1.

5.1.4 Follow-Up Interview

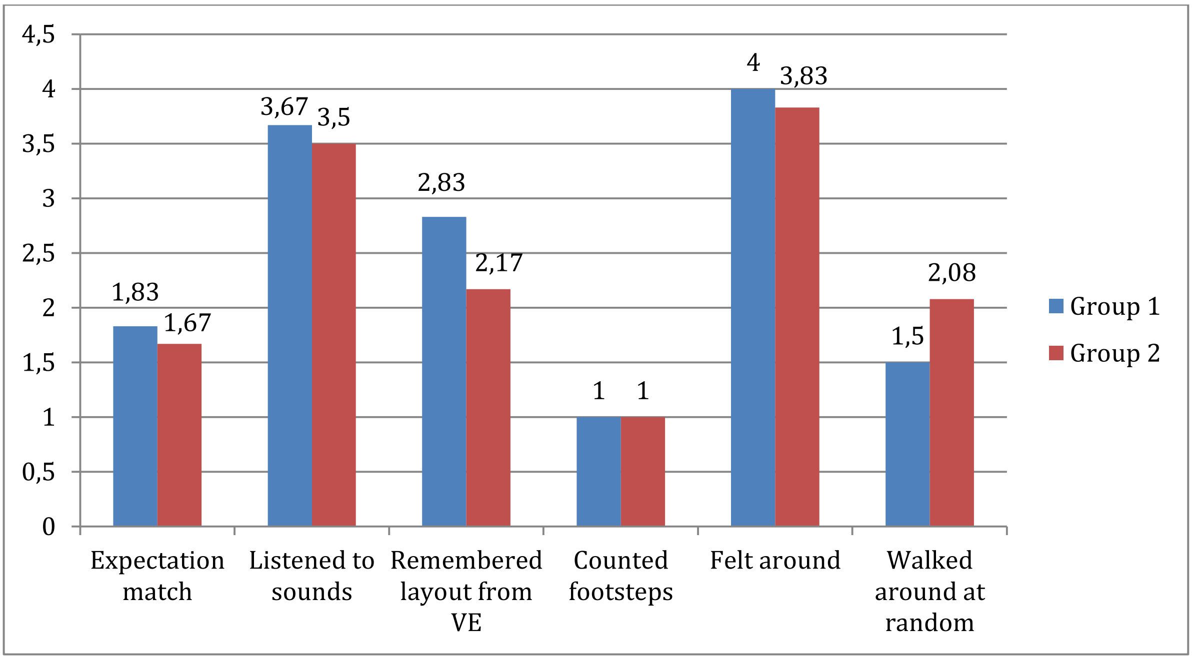

Diagram 4 Scores recorded in the follow-up interview

From the follow-up interview (Diagram 4), a rather low match of the participants' expectations about the layout of the apartment can be seen, slightly lower for group 2 than for group 1. When the participants were asked to grade what aspects of navigation they utilised and how important each aspect was, very clear trends could be seen in the results. All groups graded "listened to sounds" and "felt around" as being very important. No participant in the whole test answered that they counted footsteps. Group 2 had a harder time remembering the layout from the VE than group 1. Group 2 also answered that they walked around somewhat more at random.

5.2 Comparison between Groups 2 and 3

Next to be presented is a comparison between groups 2 (*Audiction*) and 3 (control group). The control group only performed the real-life test and filled out questionnaire 2 and the follow-up interview.

5.2.1 Real-Life Test

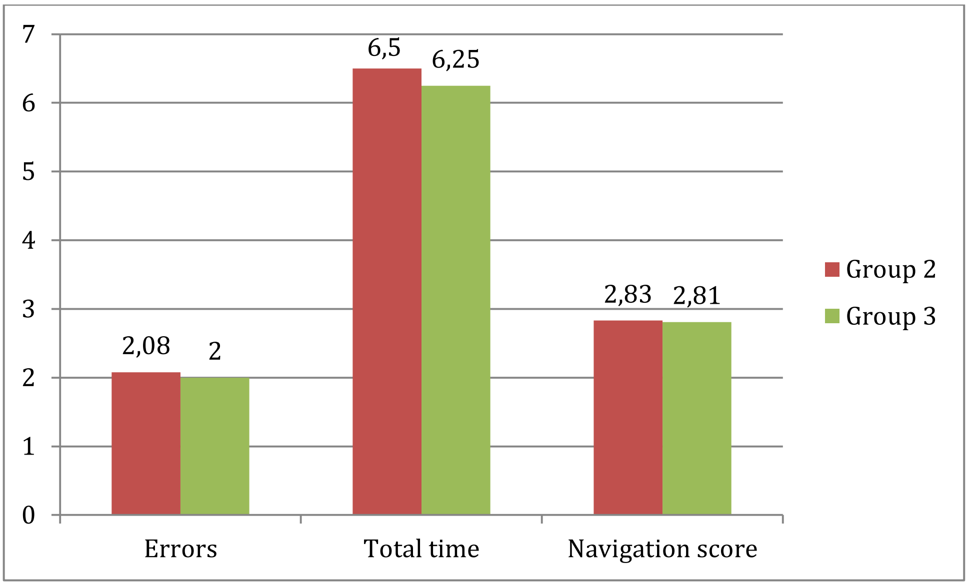

Diagram 5 Scores recorded by Jonsson and Östblad during the real-life test

As seen in Diagram 5, the number of errors between the groups is very similar, as is the total time and navigation score. It is notable that group 3 has the least errors and lowest time, although the differences are not significant.

5.2.2 Questionnaire 2

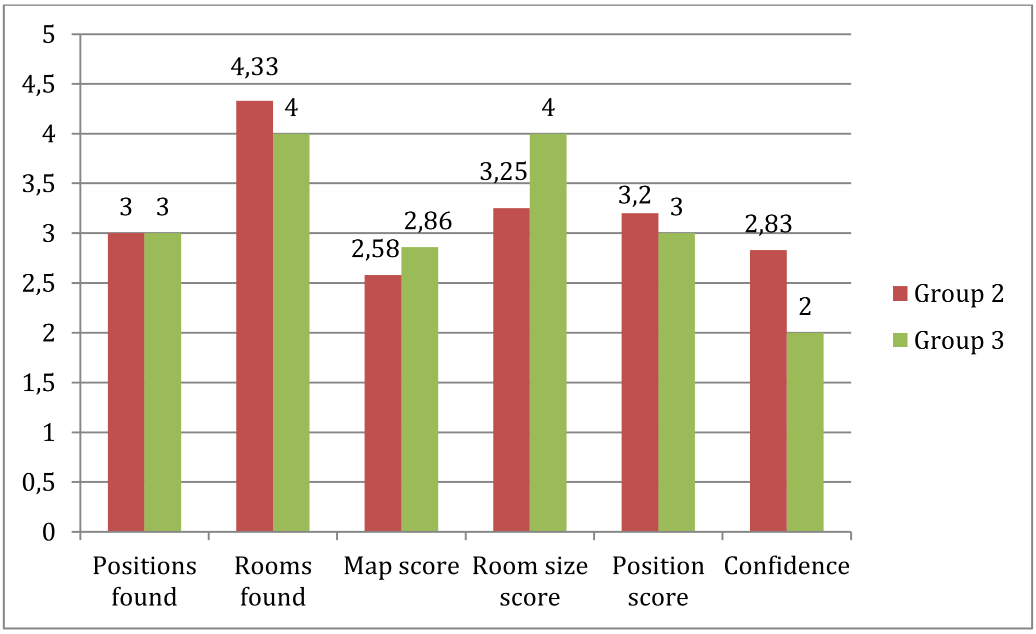

Diagram 6 Scores recorded in questionnaire 2

All the participants of groups 2 and 3 found the three positions within the time limit during the real-life test (Diagram 6). Group 2 found slightly more rooms than the control group. Group 2 got a lower map score than group 3; notable is that all participants of group 3 scored a maximum of 4 at drawing room sizes. However, group 2 were better at marking the positions. Not correlating with their map scores, group 3 showed lower confidence in their spatial knowledge of the apartment than group 2.

5.2.3 Follow-Up Interview

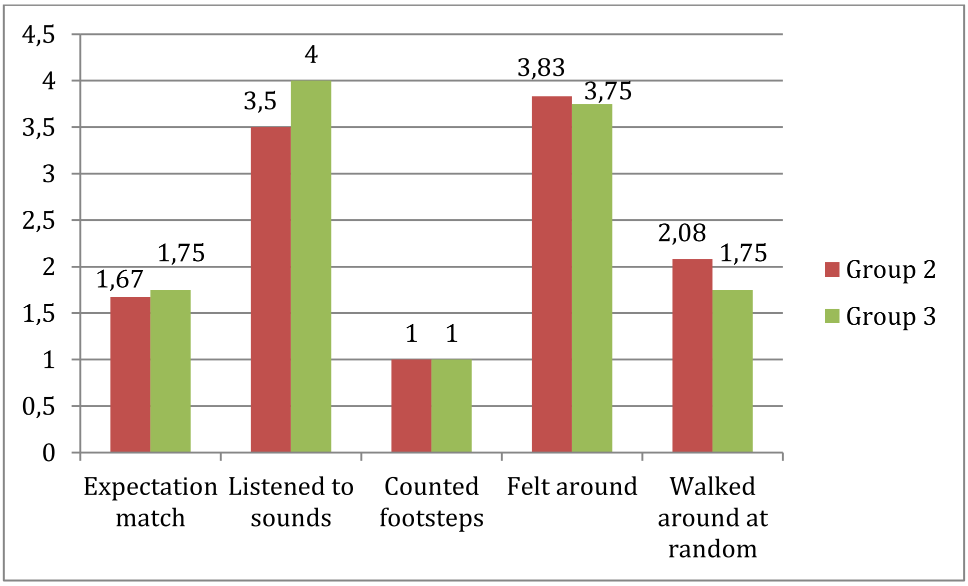

Diagram 7 Scores recorded in the follow-up interview

Lastly, a look at the follow-up interview (Diagram 7) showed that the expectations of the apartment did not match quite as well for group 2 as for group 3. As seen in the comparison between groups 1 and 2, the answers from group 3 follow roughly the same trends about what aspects of navigation were more or less important.

5.3 Analysis

In the following chapter, the results presented above will be analysed and discussed in relation to the two hypotheses on which the present study was based. Since the present work has been a pilot study with rather few participants, the results presented in this chapter are not statistically significant. There are, however, some interesting findings worth discussing.

When group 2 was compared to the scores of all the groups combined it showed that group 2 had fewer errors and a higher navigation score than average. This would suggest that the participants that played *Audiction* were better prepared for the real-life test than the other groups. In contrast, group 2 scored significantly lower when asked to draw a map of the layout and they displayed less confidence in their spatial knowledge of the apartment after having played *Audiction*. These results may seem conflicting. One explanation for this contrasting outcome might be found with the diegetic sounds. Although group 2 displayed very poor spatial knowledge of the apartment after playing the VE compared to group 1, they still performed better in finding the positions in the real-life test. The fact that *Audiction* contained the same diegetic sounds as the real environment might have helped the participants to recognise the positions, without having good overall spatial knowledge of the layout. Group 1, on the other hand, did not have any diegetic sounds related to the positions in *Audionome*, which might explain the higher error rate in the real-life test. In other words, the diegetic sounds may have provided group 2 with some recognition from the VE, thus leading the participants towards the positions. This notion is strengthened by the fact that group 2 found fewer rooms but the same or more positions. Since the participants were instructed in which rooms the positions were located, a low score on “rooms found” in correlation with a high score on “positions found” in the real-life test should be considered a positive. Furthermore, the fact that group 2 marked the positions in their maps as well as group 1, suggests that they knew where the positions were but did not have a very good sense of the layout in general. The big increase in map scores and confidence for group 2 might, again, be explained by a sense of recognition awarded by the diegetic sounds.

Group 1 performed the test faster than average but made more errors. They drew more accurate maps of the layout and showed more confidence in their spatial knowledge. These results indicate that they got a good overview of the apartment from playing *Audionome*, but received poor knowledge about where to find the positions. Moreover, these results might suggest that group 1 had, somewhat, false confidence in their spatial knowledge after playing the VE. This is also indicated by the much larger improvement in map scores, displayed by group 2 between questionnaire 1 and 2, as well as by the decrease in confidence for group 1 after the real-life test.

The match of expectations about the layout, when the groups were allowed to look at the apartment, showed no substantial differences. All the participants answered 2 or lower, except one person from group 1 and one from group 2, who both answered 3 out of a maximum 4. This goes to show how hard it can be to create a mental map of an unknown area, when not allowed to look at it. All the groups answered roughly the same when asked which aspects of navigation they had utilised and how important they were.

5.3.1 Observations

Perhaps the most notable and interesting result, was the fact that the control group performed as well and utilised sound as much as they did. Especially one participant in the control group showed remarkable skill at using acoustics to identify the environment. For instance, this participant identified the bedroom, without touching anything, just by hearing the dampened ambiance and isolation of the walls. Several other participants, from all groups, said they used the acoustics to try to identify their surroundings. These traits, shown by several students, might not be representative for people in general as they were all sound students and may be more used to listening actively in this manner (Östblad, 2011).

The insignificant differences between the performances of groups 2 and 3 might, at first glance, suggest that *Audiction* is not a viable solution to the problem of conveying spatial information without visual stimuli. However, that may not be the case. It was concluded when evaluating the results that, ten minutes of playtime with the VE was not a sufficient amount of time to properly benefit from it. Furthermore, it was concluded that the apartment used for the experiment, was not complex enough and the positions were too easy to find to fully evaluate the usefulness of the VE’s.

Before the experiment, the participants were asked if they believed themselves to have a good sense of orientation in general. No indications that their answer to that question had any impact on their performances were found.

On the subject of diegetic audio, some surprising aspects were observed. Even the groups that did not receive any diegetic sound information, before the real-life test, displayed a high degree of recognition from the sounds in the apartment. For instance, the sound of the TV prompted participants to walk in that direction, due to the preconception that the TV was in the living room. The same went for the kitchen fan, the radio in the bedroom and the acoustics of the bathroom. There were some exceptions however; one student stated that he clearly heard the TV in absolutely the wrong direction, although he did identify it as a TV. These seemingly preconceived notions about sounds in our environment might be useful to exploit when trying to convey spatial knowledge via audio.

One clear flaw with the experiment was the fact that some questions had to be subjectively evaluated, e.g. the map scores and the navigation scores. If this experiment was to be performed on a larger scale, it would be preferable to have better specified criteria for all scoring.

5.3.2 Hypotheses

Hypothesis 1: *The group playing Audiction, modelled after a real environment and utilising diegetic spatial audio, will have better spatial knowledge of the environment, and will therefore be better prepared for the real world test than the other groups.*

Group 2 were not better at drawing maps and did not have as much confidence in their spatial knowledge after playing *Audiction*. Nevertheless, they did not make the most errors and they received the highest navigation score. This hypothesis can neither be confirmed nor rejected without further study. For Hypothesis 1 to be confirmed or rejected, there would have to be significant differences between the groups in number of errors, positions found and map scores.

Hypothesis 2: *Participants that have played Audiction, modelled after a real environment and utilising diegetic spatial audio, will be able to achieve spatial knowledge transfer from the VE to the real environment.*

Group 2 displayed a fair bit of recognition in the real-life test and they connected positions to sounds. All the participants found all three positions within the allotted time but proved to have very bad spatial knowledge overall. This hypothesis can, therefore, neither be confirmed nor rejected without further research. To consider this hypothesis confirmed, the participants in group 2 would have had to display better understanding about the general layout of the apartment. Even though the map drawings were subjectively evaluated it was quite clear that they had not obtained very accurate spatial knowledge after playing the VE.

As stated in chapter 3.1 “The aim of the present research project was to evaluate the viability of diegetic spatial audio, in navigation aid software for blind people, and how it affects spatial knowledge transfer from a VE to a real corresponding environment.” The fact that all groups answered that sounds in the apartment were an important part of their navigation, would suggest that diegetic spatial audio is indeed viable when creating audio-navigation tools. Although not conveying a good overview of the layout, *Audiction* did manage to provide recognition and positional information to the players.

6 Conclusion

In the following chapter the results of the present experiment are summarised and discussed.

6.1 Summary of results

The present pilot study by Jonsson and Östblad was performed with 16 sound students from the University of Skövde. After a suitable environment was selected and replicated in the VE's, *Audiction* and *Audionome*, the students were invited to perform the experiment and were divided into three groups. Group 1 performed tasks in *Audionome,* group 2 performed the same tasks in *Audiction* and all three groups then performed corresponding tasks in the real environment in a real-life test. The experiment went well and all participants found the experience exciting and engaging.

With the relatively few participants available for this pilot study, no statistically significant results could be found. Nevertheless, there were some interesting findings on the topic of diegetic spatial audio. *Audiction* contained the same diegetic audio sources as the real environment and the group that played Audiction were able to recognise the sounds in the real-life test and relate them to the tasks they were asked to perform. The diegetic spatial audio seemed good at conveying specific positional information, but Audiction as a tool was not successful in giving the participants a good overview and spatial knowledge of the environment.

The hypotheses on which the experiment was based could neither be confirmed nor rejected. However, the aim of the research was to evaluate the viability of diegetic audio in audio-navigation software for visually impaired people. On that point, the results showed promising indications that spatial audio plays a big part when navigating without vision. It was concluded that the concept of *Audionome* with the diegetic spatial audio of *Audiction*, would most likely be a good way to convey spatial information about an environment, without the use of visual stimuli. The method utilised in the present pilot study showed promise and would be interesting to perform on a larger scale and preferably with blind participants.

6.2 Discussion

When performing the present experiment, a number of the sound students that participated showed surprising skill at using acoustics to navigate. This might be attributed to the fact that they are used to working with sound. This theory coincides with previous indications that it might be possible to get better at active and analytical listening through practise (Östblad, 2011). Sánchez et al. (2010a) showed that blind people activate parts of their brain normally associated with sight when asked to perform spatial tasks. When discussing *Audiction*, a blind test player said that to what degree blind people use sound to navigate, generally varies between different people and even between different countries. If the theory of listening practise is accurate, it is fair to assume that blind people could stand to benefit from practising their hearing in order to acquire spatial information to a greater extent.

Perhaps the most surprising result was that there seemed to be a lot of preconceptions about the sounds in the apartment. Several of the participants, from all three groups, stated that they heard the TV during the real-life test and walked towards it because they believed they would find the living room. The same occurred with the sound of the fridge and fan in the kitchen and with the radio in the bedroom. Although these particular sounds may have been too common to provide any incredible discoveries, it is interesting to ponder what could be achieved by exploiting these types of sounds deliberately to provide spatial information. One example where spatial audio is used in this manner is traffic lights. In many countries traffic lights give off sounds to provide pedestrians with information. These sounds work both as a beacon to locate the crossing as well as to give information about when to cross the street. Most people would probably recognise the sounds these traffic lights give off immediately; perhaps it would be a good idea to provide similar sound cues in other parts of society as well.

A problem for many people with disabilities is accessibility. This is an issue that has to be considered when designing public environments. For an open society to be open for everyone, it is imperative that all public places are as accessible as possible. To feel included in a society or a community, it is important to be able to take part in the same activities and visit the same places as everyone else. Train stations, for instance, often have signs with tactile maps of the station placed at strategic locations. However, in order to find this map, a person with a visual disability might have to acquire the aid of a sighted person or be familiar with the layout beforehand. This would seem as good example for when spatial audio might be utilised, in the same way as with traffic lights. If the layout of that train station could be modelled in a VE, with corresponding diegetic sounds and obstacles, it might be possible for a visually impaired person to learn the layout of the train station in a safe environment before visiting it. At the very least, this type of application might provide some recognition and thus, increase the confidence of the visitor. Taking it one step further, it might even be possible to create an application for a mobile device, with which a user can get an overview and other useful information about an environment on site. In the future, it may even be possible to use GPS in a mobile phone to quickly access such a navigation-aid for the place you are visiting, in real time online. A big problem with these types of public accessibility tools is that there is not much money to be made. Not many companies can see large financial potential in this sector and it is therefore up to the public domain to provide the funds.

The theories about diegetic audio and listening practise are not only applicable for navigation-aids and serious games research. These topics should be addressed in commercial game development as well. Whether to create better audio games or to make regular games more accessible, sound designers might benefit from creating games that utilise the way sound is used in our everyday lives. As mentioned earlier, a sense of inclusion is very important to be an integral part of society, and for children growing up today, computer games are a big part of their lives. If computer games can be created in way to include a wider range of the population, a greater sense of inclusion for people with disabilities might be achieved.

The present experiment was well received by the participants. All the students that took part found it to be an interesting and enlightening experience to navigate an unknown environment without use of their vision. Many of the participants even found it scary and felt a lot of uncertainty. Taking part in an experiment like this could benefit anyone, especially those who work with visually impaired people and especially blind children. If nothing else, to experience some of the challenges a blind child might face, first hand. Many seeing people are so used to have their vision that it might be hard to communicate with someone who has never experienced sight. It would seem that being able to understand what the person you are trying to help experiences on a daily bases, would be very beneficial for both parties.

6.3 Future work

Following the conclusions from the present experiment, it is clear that more research on the topic of diegetic spatial audio would be interesting. The pilot study performed by Jonsson and Östblad showed that VE's such as *Audiction* and *Audionome* might be viable, although they would have to be altered to some extent. A proposed continuation of the present study would be to combine the two VE concepts. Meaning, a VE played on a mobile touch screen device, from a top down view that provides a good spatial overview, as is the case with *Audionome*, but with the diegetic spatial audio found in *Audiction* to provide recognition and specific positional information. To properly evaluate such a tool there are a few alterations that would have to be made to the experiment design. More play time would be a prerequisite, as it is a complex concept that has a certain learning curve. The VE would also have to be modelled after a more complex environment to fully evaluate the differences between the group playing the VE and the control group. Another prerequisite would be to have a larger population to neutralise the impact of outliers and to get statistically significant results. Since the idea of the VE is to help blind people with navigation and orientation, any such study would preferably be executed with blind participants. It would be very interesting to have an equal amount of blind and seeing participants in each group to also study the differences in audio-navigation skill between them. Lastly, all criteria for evaluation and scoring in such an experiment would have to be well designed, to avoid having to grade performances subjectively.

References

Archambault, D. & Olivier, D. (2005) *How to make games for Visually Impaired Children*. ACE 2005, Valencia, Spain.

Bryman, A. (2002) *Samhällsvetenskapliga metoder*. Liber.

Chion, M. (1994) *Audio-vision: sound on screen*. Columbia University Press.

Copeland, D. (2000) *Associative Listening*. Soundscape: The Journal of Acoustic Ecology 1:1.

Dykhoff, K. (2002) *Ljudbild eller synvilla?: en bok om filmljud och ljuddesign*. Liber ekonomi.

Ekman, I. (2005) *Meaningful Noise: Understanding Sound Effects in Computer Games*. In Proc. Digital Arts and Cultures, 2005, Kopenhagen, Denmark.

Friberg, J. & Gärdenfors, D. (2004) *Audio Games: New perspectives on game audio*. Ace conference 2004, Singapore.

Heuten, W., Henze, N., Boll, S. & Pielot, M. (2008) *Tactile Wayfinder: A Non-Visual Support System for Wayfinding*. NordiCHI 2008: Using Bridges, 18-22 October, Lund, Sweden.

Kamel, H. M. & Landay, J. A. (2000) *A Study of Blind Drawing Practice: Creating Graphical Information Without the Visual Channel*. ASSETS'00, November 13-15, 2000, Arlington, Virginia.

Lahav, O. & Mioduser, D. (2004) *Blind Persons' Acquisition of Spatial Cognitive Mapping and Orientation Skills Supported by Virtual Environment*. Proc. 5th Intl Conf. Disability, Virtual Reality & Assoc. Tech., Oxford, UK, 2004.

Loomis, J. M., Golledge, R. G. & Klatzky, R. L. (1998) *Navigation System for the Blind: Auditory Display Modes and Guidance*. Presence, April 1998, Vol. 7, No. 2, 193–203.

Magnusson, C., Waern, A., Rassmus-Gröhn, K., Bjernryd, Å., Bernhardsson, H., Jakobsson, A., Salo, J., Wallon, M. & Hedvall, P-O. (2011) *Navigating the world and learning to like it - mobility training through a pervasive game*. MobileHCI 2011, Aug 30–Sept 2, 2011, Stockholm, Sweden.

Masuch, M. & Röber, N. (2004) *Interacting with sound: An interaction paradigm for virtual auditory worlds*. Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia, July 6-9, 2004.

Oren, M. A., Harding, C. & Bonebright, T. (2008) *Evaluation of Spatial Abilities within a 2D Auditory Platform Game*. ASSETS'08, October 13–15, 2008, Halifax, Nova Scotia, Canada.

Östblad, P. A. (2011) *Analytisk lyssning: Hur påverkar bilder vad vi hör och kan vi träna vår hörsel*? Bachelor thesis, University of Skövde, Sweden.

Passini, R. & Proulx, G. (1988) *Wayfinding without Vision: An Experiment with Congenitally Totally Blind People*. Environment and Behavior, March 1988, vol. 20 no. 2, 227-252.

Paterson, N., Naliuka, K., Jensen, S. K., Carrigy, T., Haahr, M. & Conway, F. (2010) *Design, Implementation and Evaluation of Audio for a Location Aware Augmented Reality Game*. Fun and Games 2010, Leuven, Belgium.

Quiñones, P-A., Greene, T., Yang, R. & Newman, M. (2011) *Supporting visually impaired navigation: a needs-finding study*. In Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems (CHI EA '11). ACM, New York, NY, USA.

Roden, T. & Parberry, I. (2005) *Designing a Narrative-Based Audio Only 3D Game Engine*. ACE 2005, Valencia, Spain.

Sánchez, J., Lumbreras, M. & Cernuzzi, L. (2001) *Interactive Virtual Acoustic Environments for Blind Children: Computing, Usability, and Cognition*. In CHI '01 extended abstracts on Human factors in computing systems (CHI EA '01). ACM, New York, NY, USA, 65-66.

Sánchez, J., Sáenz, M., Pascual-Leone, A. & Merabet, L. (2010a) *Navigation for the Blind through Audio-Based Virtual Environments*. CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.

Sánchez, J., Sáenz, M. & Garrido, J. M. (2010b) *Usability of a Multimodal Video Game to Improve Navigation Skills for Blind Children*. ACM Transactions on Accessible Computing, Vol. 3, No. 2, Article 7, Pub. date: November 2010.

Sánchez, J. & de la Torre, N. (2010c) *Autonomous Navigation through the City for the Blind*. ASSETS'10, October 25 - 27, 2010, Orlando, Florida, USA.

Schneider, J. & Strothotte, T. (2000) *Constructive Exploration of Spatial Information by Blind Users*. ASSETS '00, November 13-15, 2000, Arlington, Virginia.

Shaughnessy, J. J., Zechmeister, E. B. & Zechmeister, J. S. (2005) *Research methods in psychology*. 7. ed. Boston: McGraw-Hill Education

Tversky, B. (1993) *Cognitive Maps, Cognitive Collages, and Spatial Mental Models*. A.U. Frank, I. Campari (Eds.), Spatial Information Theory. A Theoretical Basis for GIS, European Conference, COSIT'93, Lecture Notes in Computer Science, 716, Springer, Berlin (1993), pp. 14–24.

Wallet, G., Sauzéon, H., Rodrigues, J. & N'Kaoua. (2008) *Use of virtual reality for spatial knowledge transfer: Effects of passive/active exploration mode in simple and complex routes for three different recall tasks*. VRST'08, October 27-29, 2008, Bordeaux, France.

Yang, R. et al. (2011) *Supporting Spatial Awareness and Independent Wayfinding for Pedestrians with Visual Impairments*. ASSETS'11, October 24-26, 2011, Dundee, Scotland, UK.

Appendix A - Tables

Table 1 Results for group 1 - *Audionome*

Audionome													
	Notes					**Questionnaire 1**							
Participant	Errors	Total time	Navigation score	Good sense of orientation	Our notes	Positions found	Rooms found	Map score	Room size score	Position score	Confidence		
P1	3	5,25	2,83	YES	Mixed u	3	6	3,5	3,5	2	2		
P2	8	10	1,25	NO	Didn't re	3	6	3	3,5	2	1		
P3	3	6	2,5	YES	Heard TV	-	-	-	-	-	-		
P4	3	4	3,25	NO	Wrong tu	2	5	3,5	3	-	2		
P5	3	5,75	3	NO	Seemed	3	4	1,5	3	1,5	2		
P6	1,5	3,5	2,75	NO	Aggresiv	2	5	3	2	3,5	3		
TOTAL	**21,5**	**34,5**	**15,58**	**-**	**-**	**13**	**26**	**14,5**	**15**	**9**	**10**		
Mean	**3,5**	**5,75**	**2,6**	**-**	**-**	**2,6**	**5,2**	**2,9**	**3**	**2,25**	**2**		
	Questionnaire 2						**Follow-up interview**						
Participant	Positions found	Rooms found	Map score	Room size score	Position score	Confidence	Expectation match	Listened to sound	Remembered	Counted footste	Felt around	Walked aroun	Other
P1	3	6	3,5	4	3	3	1	4	2	1	4	1	3
P2	2	5	1,5	3	2	1	1	4	3	1	4	1	1
P3	3	6	2	4	3,5	3	2	3	2	1	4	2	1
P4	3	5	3	3	-	3	2	4	3	1	4	2	3
P5	3	3	2	3,5	3,5	3	2	4	3	1	4	2	1
P6	3	4	3,5	2,5	4	3	3	3	4	1	4	1	1
TOTAL	**17**	**29**	**15,5**	**20**	**16**	**16**	**11**	**22**	**17**	**6**	**24**	**9**	**10**
Mean	**2,83**	**4,83**	**2,58**	**3,33**	**3,2**	**2,67**	**1,83**	**3,67**	**2,83**	**1**	**4**	**1,5**	**1,67**

Table 2 Results for group 2 - *Audiction*

Audiction

	Notes					Questionnaire 1					
Participant	Errors	Total time	Navigation score	Good sense of orientation	Our notes	Positions found	Rooms found	Map score	Room size score	Position score	Confidence
P7	2,5	5,75	3,1	NO	Listens a	3	4	2,5	3	2,5	1
P8	3	9,25	2,75	NO	Listened	3	3	2	1,5	3	2
P9	1	8,5	2,38	NO	Rememb	-	-	-	-	-	-
P10	3	3,5	2,9	YES	Sounded	2	2	1	1	1	1
P11	1	6,75	3,1	NO	Sweeps	2	3	1	-	-	2
P12	2	4,5	2,75	YES	Wants to	3	3	2	3,5	2,5	2
TOTAL	**12,5**	**38,25**	**16,98**	**-**	**-**	**13**	**15**	**8,5**	**9**	**9**	**8**
Mean	**2,08**	**6,5**	**2,83**	**-**	**-**	**2,6**	**3**	**1,7**	**2,25**	**2,25**	**1,6**

	Questionnaire 2						Follow-up interview						
Participant	Positions found	Rooms found	Map score	Room size score	Position score	Confidence	Expectation match	Listened to sound	Remembered	Counted footste	Felt around	Walked around	Other
P7	3	5	3,5	3,5	-	2	3	3	3	1	4	2	3
P8	3	5	1,5	1,5	2,5	3	2	4	3	1	3	2	1
P9	3	4	1	3,5	4	2	1	3	1	1	4	3	1
P10	3	4	3,5	3,5	3,5	4	1	4	2	1	4	1	1
P11	3	4	3	4	3	3	1	3	2	1	4	2	1
P12	3	4	3	3,5	3	3	2	4	2	1	4	2,5	1
TOTAL	**18**	**26**	**15,5**	**19,5**	**16**	**17**	**10**	**21**	**13**	**6**	**23**	**12,5**	**8**
Mean	**3**	**4,33**	**2,58**	**3,25**	**3,2**	**2,83**	**1,67**	**3,5**	**2,17**	**1**	**3,83**	**2,08**	**1,33**

Table 3 Results for group 3 – control group

Control group											
	Notes					Questionnaire 1					
Participant	Errors	Total time	Navigation score	Good sense of orientation	Our notes	Positions found	Rooms found	Map score	Room size score	Position score	Confidence
P13	1	7	3,5	NO	Listened	-	-	-	-	-	-
P14	1,5	4,75	2,7	NO	Aggresiv	-	-	-	-	-	-
P15	4,5	7,75	2,42	YES	Heard TV	-	-	-	-	-	-
P16	1	5,75	2,6	YES	Heard TV	-	-	-	-	-	-
TOTAL	**8**	**25,25**	**11,22**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
Mean	**2**	**6,25**	**2,81**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**

	Questionnaire 2						Follow-up interview						
Participant	Positions found	Rooms found	Map score	Room size score	Position score	Confidence	Expectation match	Listened to sounds	Remembered layout	Counted footsteps	Felt around	Walked around at	Other
P13	3	5	3,5	4	3	2	1	4	-	1	4	1	4
P14	3	4	2,5	4	-	2	2	4	-	1	4	2	1
P15	3	4	2,5	4	3	2	2	4	-	1	3	2	-
P16	3	3	3	4	3	2	2	4	-	1	4	2	1
TOTAL	**12**	**16**	**11,5**	**16**	**9**	**8**	**7**	**16**	**-**	**4**	**15**	**7**	**6**
Mean	**3**	**4**	**2,86**	**4**	**3**	**2**	**1,75**	**4**	**-**	**1**	**3,75**	**1,75**	**2**