D3.1 User expectations and cross-modal interaction

Alce, Günter; Liodenot, David; Hermodsson, Klas; Lemordant, Jacques; Lasorsa, Yohan; Chippendale, Paul

2013

SEVENTH FRAMEWORK PROGRAMME

FP7-ICT-2011-1.5 Networked Media and Search Systems
b) End-to-end Immersive and Interactive Media Technologies

Specific Targeted Research Project

VENTURI

(FP7-288238)

immersiVe ENhancemenT of User-woRld Interactions

D3.1 User expectations and cross-modal interaction

Due date of deliverable: 31-01-2013
Actual submission date: [31-01-2013]

Start date of project: 01-10-2011
Duration: 36 months

Summary of the document

Document Code:	D3.1 User expectations and cross-modal interaction - v0.4
Last modification: State:	31/01/2013 Add audio content creation
Participant Partner(s): Editor & Authors (alphabetically): Fragment:	INRIA , SONY Editor: Jacques Lemordant Authors: Alce, Günter (SONY), Hermodsson, Klas (SONY), Lasorsa, Yohan (INRIA), Liodenot, David (INRIA), Paul Chippendale (FBK) No
Audience:	☒ public ☐ restricted ☐ internal
Abstract:	This document is deliverable D3.1 "User expectations and cross-modal interaction" and presents user studies to understand expectations and reactions to content presentation methods for mobile AR applications and recommendations to realize an interface and interaction design in accordance with user needs or disabilities.
Keywords:	Interaction, cross-modal, Augmented Reality, navigation, visually impaired people, 3D audio
References:	Refer to the corresponding section at the end of the deliverable

Document Control Page

Version number	V0.4
Date	31/01/2013
Modified by	Paul Chippendale
Comments	Final Quality check
Status	☐ draft ☒ WP leader accepted ☒ Technical coordinator accepted ☒ Project coordinator accepted
Action requested	☐ to be revised by partners involved in the preparation of the deliverable ☐ for approval of the WP leader ☐ for approval of the technical coordinator ☐ for approval of the project coordinator Deadline for action: 31/01/2013

Change history

Version number	***Date***	***Changed by***	***Changes made***
0.1	11/10/2012	G. Alce	Preliminary version
0.2	21/12/2012	D. Liodenot	Integrated user study for audio and experiments with visually impaired people
0.3	02/01/2013	D. Liodenot	Add audio content creation
0.4	31/01/2013	P. Chippendale	Final Quality check

Table of Contents

Executive Summary

Scope

This document provides the deliverable contents related to the T3.1 User expectations from Mixed Reality and cross modal interaction.

Audience

This deliverable is public.

Summary

In this report, the objective is to investigate user's expectations and reactions towards content presentation in a mixed reality fashion. AR gaming, interactive audio scene and navigation with visually impaired people are considered.

Structure

This deliverable is structured as follows: Section 1 is an introduction explaining the objective of the deliverable. Section 2 describes the methodology, the outcome and design recommendations of the AR gaming user study. In section 3, the audio integration based on the MAUDL format and user study are described. Finally, section 4 considers audio for indoor and outdoor navigation and experiments with visually impaired people.

1 Introduction

User studies have been undertaken to understand the expectations and reactions of users to content presentation methods for mobile AR applications, taking into account usability aspects. The use-cases defined in WP2 provide the background for this study. In a user-centred design approach, current and future audio and visual technologies are explored to learn how to improve the efficiency and quality of AR applications and assistive AR technologies. In this report, AR gaming, interactive audio scene and navigation are considered. The expectations from visually impaired people with regards to mixed reality applications will be investigated, especially with pedestrian navigation applications.

2 AR gaming

This section describes the methodology, the outcome and design recommendations of the AR gaming user study.

2.1 Introduction

The user studies are part of an iterative design process (Figure 2.1). User centred design means keeping the user at the centre of focus when developing and designing. The users should be directly involved from the beginning of the design process, where they can actually influence the design and not only be asked to help in validation. The outcome of the user studies should then be reused in the next version of the product.

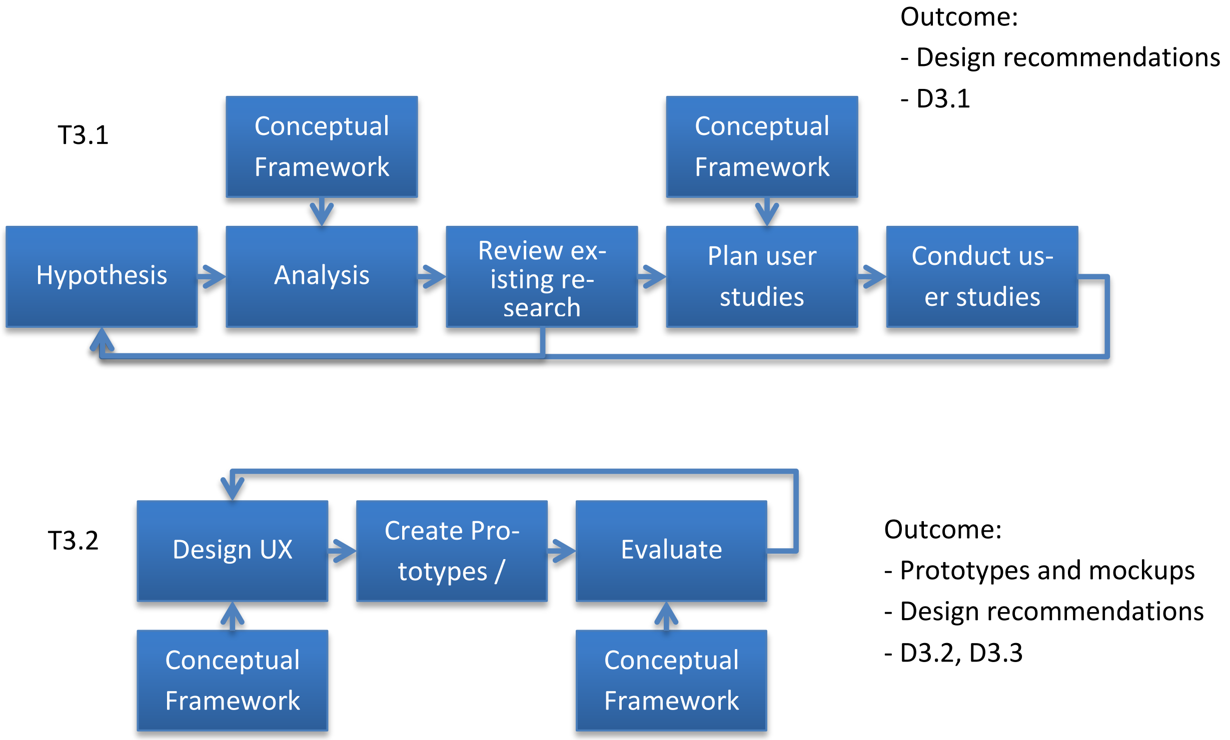


FIGURE 2.1 THE ITERATIVE DESIGN PROCESS

The use-cases defined in D2.1.1 were considered and since the first year use-case is based on gaming, a decision was made to investigate how people react to current AR marker based games and compare this with future marker-less games such as those illustrated by VeDi 0.9.

Since we are still in an early stage of the project, we did not focus on quantitative results. Instead our focus was on participants' opinions of the experience of AR gaming, and, to see if participants find the technology flaws irritating. Subjective data was gathered through observations, semi-structured interviews, questionnaires and NASA-TLX. The NASA-TLX workload questionnaire enabled us to look at how difficult participants found the playability of each AR game [2].

2.2 AR gaming user study

User studies were conducted on existing marker based AR games, since one of the objectives was to investigate the pros and cons of current AR interfaces as the game developed by VENTURI was still not available. The idea was to understand how the user experiences are influenced by technical instabilities and to understand what the users think about the new concept of AR gaming in terms of presentation and interaction.

When preparing user studies, a conceptual framework is useful both for the design phase and for the evaluation phases; in order to keep the user and context in mind. A concept framework assigns a priority for issues during the designing phase, and provides clues about what to include in the evaluation. It also gives one an understanding for real usage scenarios, especially when the usage scenario is not very clear. According to Wikipedia, 2012-01-05:

Conceptual frameworks (theoretical frameworks) are a type of intermediate theory that attempts to connect to all aspects of inquiry (e.g., problem definition, purpose, literature review, methodology, data collection and analysis). Conceptual frameworks can act like maps that give coherence to empirical inquiry. Because conceptual frameworks are potentially so close to empirical inquiry, they take different forms depending upon the research question or problem.

Examples of conceptual frameworks are: Grounded theory, Activity Theory (AT) and distributed cognition. For the AR gaming user study, AT was used both for preparing the user studies and for analysing the collected interview data.

2.3 Method

Our approach was to design the user studies similar to a case study, where the main objective is rather small and we have a small number of participants. The focus was on in-depth investigations, multiple data sources and an emphasis on qualitative data [4]. Furthermore, to analyse the qualitative data an AT checklist [3] was used. The activity checklist is intended to elucidate the most important contextual factors of human–computer interaction. It is a guide to the specific areas in which a researcher or practitioner should pay attention when trying to understand the context in which a tool will be, or is, used. An AT checklist lays out a kind of “contextual design space” by representing the key areas of context specified by AT, the areas are:

- Means/Ends
- Environment
- Learning/cognition and articulation
- Developments

2.3.1 Participants

We conducted interviews with six people to start with. In the eventuality answers diverged too much plans were made to interview more users. However similar answers and views of AR marker based gaming were identified and a level of saturation was achieved.

Interesting population for the user studies is non-engineering but with interest of new technology. The assumption was that people from marketing and administration have little or no technical background. Further it was assumed that those interested in participating in the user study are interested in new technology. Invitation mail was sent out to all administrative staff within Sony Mobile Communications. Pilots where used to ensure the relevance of the questions and to get an indication of how long the user study will last. Four of the participants were females and two were males with ages varied between 30 and 47. Additional six participants tried out the AR games and answered the questionnaire and NSA-TLX. This was done to enlarge the number of participants filling in the questionnaires to be able to illustrate trends with graphs.

2.3.2 AR Games

The objective was to let participants try both unstable and stable AR games. AR Blitz and AR Defender are two games that easily lose tracking while Danger Copter and NerdHerder are more stable. It was decided to let participants test the games both on phone form factor and tablet form factor. However Danger Copter worked only on Nexus phone thus it could not be tested on tablet. Halfway of the studies AR Blitz stopped working with the tablet and NerdHerder was released at the same point so in order to have at least two games which are played with tablet it was decided to let participants play NerdHerder with tablet. The complete list of the games and the hardware form factor that were used are listed below:

1. AR Blitz with phone (Sony Xperia P) (Figure 2.2)
2. AR Blitz with tablet (Sony Tablet S)
3. AR Defender with phone (Sony Xperia P) (Figure 2.3)
4. AR Defender with tablet (Sony Tablet S)
5. Danger Copter with phone (Nexus) (Figure 2.4)
6. NerdHerder (Sony Tablet S) (Figure 2.5)

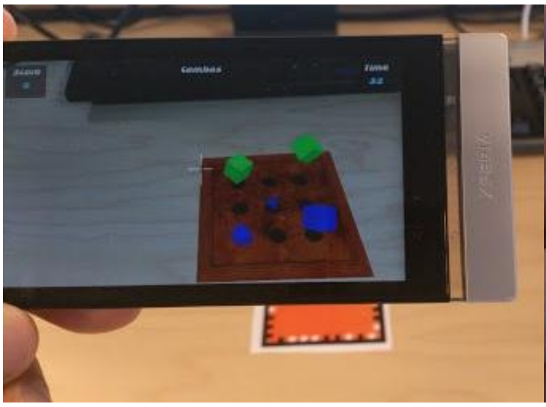


FIGURE 2.2 AR BLITZ

FIGURE 2.3 AR DEFENDER

FIGURE 2.4 DANGER COPTER

FIGURE 2.5 NERDHERDER

The order of the games and form factor were mixed. For example, if the first participant tried out with a phone form factor the next one would start with tablet form factor and so on. Also the order of the games was mixed.

2.3.2.1 AR Blitz

AR Blitz is a game where user needs to hit shapes popping out of a hole (Figure 2.2). It is a very simple game where users must touch the screen to hit the different shapes. Sometimes it has some difficulties finding the marker. Users can move the phone but only with small movements otherwise it loses tracking. Video link: http://youtu.be/bSFo_U30lWw .

2.3.2.2 AR Defender

AR Defender is a game involving a tower, which users need to defend by using different weapons to target the enemies (Figure 2.3). Users need to move the phone to target the enemies and need to press a button to shoot. Also sensitive and loses tracking easily. Video link: http://youtu.be/rB5xUStsUs4 .

2.3.2.3 Danger Copter

Danger Copter is a game where you are a pilot of a fire-fighting helicopter (Figure 2.4). Most of the interaction is done by moving around the phone in all directions (sideways, up and down etc.). The game is very stable and very intuitive. Video link: http://youtu.be/LlFryaZwD6Y .

2.3.2.4 NerdHerder

In NerdHerder users are IT-managers and need to tempt the workers back to their office with a donut (Figure 2.5). The interaction with NerdHerder is similar as Danger Copter but the metaphor is not as easy to understand as the Danger Copter. NerdHerder is also a very stable game and warns users when it is about to lose tracking which is a great advantage. Video link: http://youtu.be/RSxImyFXSXw .

2.3.3 Procedure

The in-depth interviews were performed at the Usability Lab of Sony Mobile Communications. It started with some refreshments, in order to get the participant relaxed and letting her/him getting used to both the environment and the moderator (one of the authors). After the refreshments the session started with a short introduction of Azuma's definition of AR [1] followed by explaining the objective of the study.

The user study session was one hour long and it was recorded with two video cameras and a table microphone (Figure 2.6).

FIGURE 2.6 THE SETUP OF THE USER STUDY

Two cameras were used in order to cover both the participants' view of the phone and the participants' face to observe the participants' reactions. All participants signed an informed consent form (Appendix 7.4).

The session continued as follows, participant played the game using a phone or tablet device, followed by a semi-structured interview that was followed by a questionnaire and NASA TLX. These steps were repeated for all tested AR games and with both phone and tablet form factor.

The video material was transcribed and then colour coding was used to identify patterns about participant's thoughts about AR presentation and interaction. After that the Activity Theory checklist [3] was used to go through the transcribed material again.

2.4 Results and Discussions

This section presents and discusses results of questionnaires and interviews.

2.4.1 Questionnaire

In appendix 7.5 – 7.12 all results from the questionnaires of the phone form factor and tablet form factor are presented. The results from the questionnaires shows that the participants found that it was easy to understand how to play the AR games (figure 7.1) this is also evident in the NASA –TLX (figure 7.15). Also most participants found the game sufficiently stable. Technical instabilities indicate some diversion. However the majority thinks that the games are stable enough and that both the responsiveness and how the camera picture followed the movements are stable enough.

It should be noted graphs presented in figure 7.1 – figure 7.7 is the results from phone form factor. Similar trend can be seen also for tablet form factor shown in figure 7.8 – figure 7.14 in appendix.

From the NASA-TLX questionnaires, it can be found that the majority of participants think that mental demand and the physical demand are low similar to the answers from figure 7.15 and figure 7.16, but note that the temporal, effort and frustration diverse figure 7.17, figure 7.19 and figure 7.20.

2.4.2 Interviews

The following topics emerged in the analysis of the interviews: interaction, engagement and environment. These topics are discussed below.

2.4.2.1 Interaction

All participants commented about interaction techniques. In the beginning participants found it strange to move the phone instead of as usual touching the display when playing. However after a couple of minutes they got into it and started to discover new "features" such as zoom in and out by simply moving the phone away and towards the marker. Some of the participants pointed out that especially Danger Copter is very intuitive, probably since the participants could immediately relate the movements with real life scenario.

2.4.2.2 Engagement

The majority of participants showed signs of engagement in the game. One such signal was the use of spontaneous quotes during the game: "*come I will save you*", "*don't run away*", *"like the idea of play in "real surroundings""*, "*wow fun, my kids should see this*"(author's translation). Also, participants found the AR games fun. Despite this, however, none of the participants would consider playing them again. One main reason is that the games are too simple, and you need markers to be able to play. The latter view is illustrated in the quote: "*I would have to plan when to play, it would not be something I do spontaneously*". It was fairly noticeable that the participants found the Danger Copter game much more engaging than the others probably since the metaphor of being a helicopter is much easier to relate with real life scenario and participants liked the idea of rescuing people.

2.4.2.3 Environment

The participants addressed issues when it comes to the environment, in terms of the location in which participants play preferable. This was related to the fact that most participants generally play (mobile games) before going to bed. This means that in order to play the game with markers the room needs to be bright with a flat surface e.g. a table in order to put the marker on. One participant suggested that she would like to play while waiting for the buss if the markers where e.g. on the ground or placed on the environment.

2.5 Conclusions

The objective of the questionnaires and interviews was to find out how technical instabilities influence the user experience and try to see how people react with the new way of interaction.

Both from the questionnaires and interviews it is certain that participants were irritated but still found the technical issues sufficiently stable such as detection of the game board, location of the virtual objects and responsiveness. Perhaps, however, since it is "just" a game it is not that important but if it was something important such as buy a ticket or be guided to a place while in a hurry it is essential that it is stable.

Moving the phone instead of as usual just touching the display was strange interaction in the beginning. However participants got used to it fairly quickly. Utilizing movements with good scenarios that people can relate to, is emphasized.

2.6 Design recommendations

This sections summaries recommendations emerged both from AR gaming user studies and expert evaluation of VEDI 0.9. Expert evaluation is described in detail in delivery report D6.4.

Recommendations from AR gaming:

- Avoid markers or at least "hide" markers in the environment
- Incorporate real objects in the game
- Create game which relate to real life scenario

Recommendations from expert evaluations:

- Give feedback when tracking is lost
- Give feedback when tracking is about to be lost
- Have on screen prompts. For the VEDI game relevant prompt could be; "Pickup burger", "Deliver burger"
- Have feedback sounds/vibration to make user aware of new goal
- Have floating burger on top of car to show state of the game (and remove it when dropped off)
- Increase time to drop off hamburger
- Make pickup and delivery easier
- Pause the countdown timer while tracking is lost
- Improve location and visibility of drop off location
- Steer the car with a joystick instead of current UI

3 Audio integration & User study

This section describes the audio integration based on the MAUDL format and user studies. We will describe how to create audio content specifically adapted to the mobile context. Then, We will show user tests and results on an interactive audio scene based on the VeDi 1.0 game scenario (burger delivery). Finally, user tests are performed using pre-rendering HRTF samples.

3.1 Audio content creation

Creating audio content for mobile application must take in account the specifics of the target platform and audio API limitations. Special care must also be taken when preparing the sound files to optimize the quality and clarity of their reproduction on any conditions: headphones, portable speakers or larger sound systems. This includes some prerequisites on audio file formats and sound processing.

3.1.1 Choosing an audio format

The choice of the output audio formats depends on many aspects:

- The target audio API, which must be able to handle the audio format.
- The hardware limitations of the target platform, as the available quantity of RAM may limit the choice of usable audio formats.
- The length of the sound, and its trigger frequency: small and frequently triggered sounds should be provided in uncompressed or linearly compressed audio formats (like PCM or ADPCM) when very long audio content (such as background music) is better to handle using good compression formats (like MP3, OGG or M4A).

- The playback type of the sound: for sounds that will need to be looped, the use of compressed formats is discouraged since the compression and the overhead due to the decompression stage may alter the looping capability of the sound, resulting in pops and cracks.

When possible, the use of the ADPCM format for short or looping sounds is encourage since it reduces the size of output files with no noticeable penalty on audio quality of playback. For sounds with longer duration, MP3 is a good choice when available as it often can be decoded in hardware, and have low decoding overhead when it is not the case.

The number of channels to use (generally 1 or 2) also depends on the audio content and limitations of the platforms:

- When space limitation for sounds is tight (in RAM or for application size restrictions), you better use mono sounds (1 channel).
- For music or sound ambiances that are not aimed to be spatialized using 3D rendering, using stereo sounds (2 channels) can enhance the sound quality and user experience. UI sounds may also benefit from using stereo format in some cases.
- For voice, sound effect or spatialized sound it is better to use mono (1 channel).

Finally, for the sample rate, most of the time you should stick to using 44100 Hz sample rate frequency to ensure best sound quality. If sound size is really a concern, then for voice sample or simple sound effects you can reduce their sample rate frequency to 22050 Hz without losing too much on audio quality.

3.1.2 Preparing sounds for mobile usage

As you may use audio content coming from various sources (recordings, internet, sound banks…) the volume and spectral balance of the sounds may differ drastically. Thus it is necessary to harmonize these sounds using audio processing. To perform this processing, audio editor software is required such as Adobe Audition [5], Steinberg Wavelab [6] or Audacity [7], a good open-source and cross-platform solution for audio editing (http://audacity.sourceforge.net.

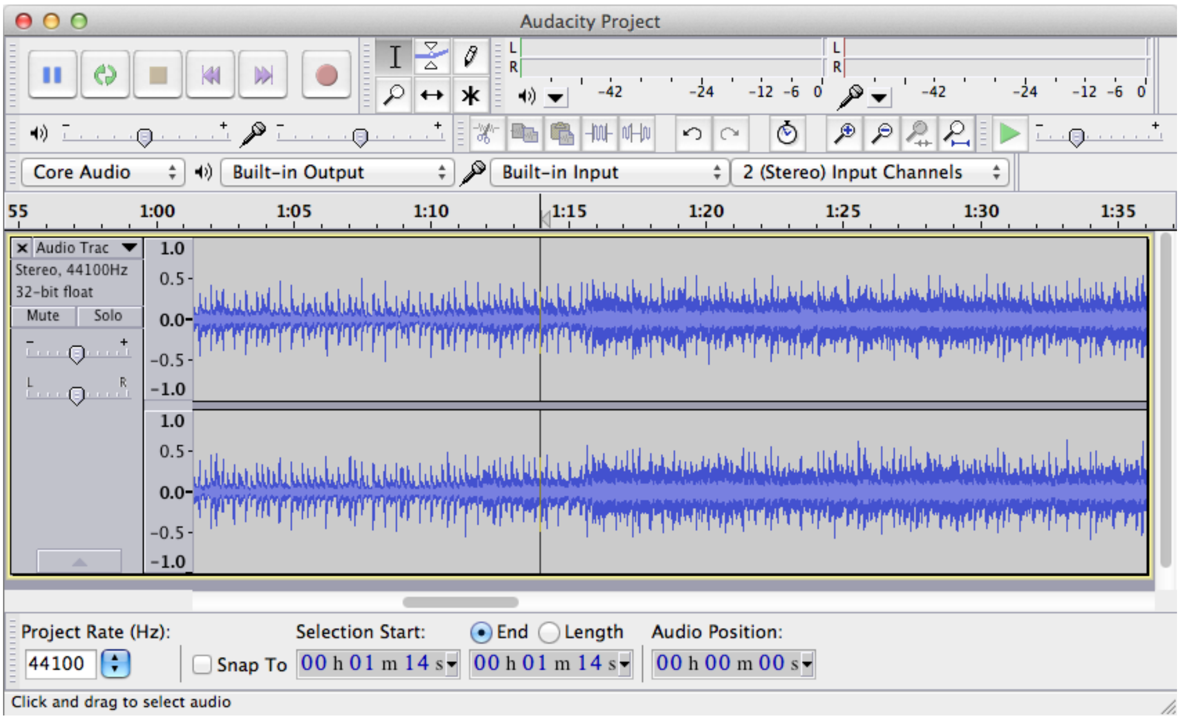


FIGURE 3.1 AUDACITY, AN OPEN-SOURCE AUDIO EDITOR

Denoising

When using audio content coming from recordings made with low-quality microphones, there may be a lot of noise in the sound. Then we need to use the denoiser plugin or processor from the audio editor to reduce this unwanted noise that occurs in the higher frequencies.

The denoising process is essentially done in 2 stages:

- First, you need to determine the noise profile. To do this, select a blank part of your recording that contains no significant content other than noise, then train the denoiser tool with this.

- Then select the noise reduction to apply, in dB: 6-12 dB allow reducing the noise with no or little side artefacts, while greater values further reduce the noise but may alter the quality of the original recording. Once you have determined the optimal noise reduction value, apply the process to your entire sound.

Equalization

Audio contents from various sources may have very different spectral balance. It is particularly important to adjust the frequency spectrum of sounds targeted at mobile applications, since most of these will be played through the integrated speaker or low-quality ear buds.

Special care must be taken especially for low frequencies, which eats most of the energy power of sounds, but will be inaudible on such listening systems. Too much high frequencies may become unpleasant to the ear, while too much mediums frequencies may cause the sound to be aggressive. A lack of these frequencies may make the sound dull of empty though. Everything is a matter of balance and harmonization. But since equalization is a very subjective process and depends heavily on the initial audio content, it is hard to provide general advices or recipes, everything is best done using testing and A/B comparisons.

The equalization process can be done using equalizer plugins, and may come in various forms. For these kinds of process, it is best to use parametric equalizer and limit the correction to removing or attenuating unwanted frequencies rather than boosting some frequency ranges.

There are 3 typical parameters to understand for parametric equalizers:

- The base frequency of the filter is the centre frequency at which the equalizing process will occur.
- The Q factor or quality factor corresponds to how steep will be the correction curve, in other words how large will be the correction range around the base frequency. Higher values means that the correction will be very localized and precise, while lower values affects greater frequency ranges.
- The correction factor, expressed in dB, determined how much the selected frequencies would be cut or boosted. Remember as a general rule that is always better to cut frequencies off than to boost some.

Volume normalization

This is one of the most critical and tricky processes. Sound files from various sources can have drastic differences in perceived audio volumes. There is notions to separate here: sound intensity corresponding to the actual peak volume of audio content (expressed in dB, from 0 to -96 dB for 16 bits files) and perceived volume (expressed in RMS [Round Mean Square] dB) corresponding to an average of the perceived sound energy. Peak volume normalization (as you may found in most audio editors) does not adjust the volume differences in audio files, it only maximize the volume based on the most intense audio peak in the files. On the other hand, perceived volume normalization may need special care depending of the audio editor and the audio content: in order to adjust the mean perceived volume, a compression of the sound dynamics must be performed, with the use of an audio limiter to prevent digital saturation of the sound.

In order to perform this process, you can use either a compressor/limiter plugin or two separate plugins, or special RMS normalization plugin as some audio editors provides.

The target value for perceived volume normalization heavily depends on the content, but values around -10 or -12 RMS dB are good candidates for normalization of all kind of audio content. You may even push it to -6 or -8 RMS dB for special effects or UI sounds, to add more impact.

Trimming and fading ends

The final preparation stage is to select in an appropriate manner the specifically desired sound part from the current audio file. Silences or noises at the beginning or the end of a sound may disturb the user or the timing of interactions, so you must make sure your sounds start and ends at the right time. To do this, just delete the unwanted parts of the sounds at the beginning and at the end of the audio file using the audio editor.

After your sound is correctly trimmed, the last stage is to ensure your sound has a smooth start and end at 0-value point, using fades. Using the built-in functions of your audio editor, perform a fade-in on the beginning and a fade-out at the end or your sound. Good all around values is around 5-10 ms for fades. If you have sounds with ending audio tails like reverbed sound, you may want to perform a longer fade-out (1 or more seconds) in order to make the sound ends smoothly. Applying fades is essential to avoid unpleasing clicks and cracks at the beginning and end of sounds when they are played.

3.1.3 Creating seamless loops

The creation of perfectly looping audio files is not an easy task: you must have appropriate audio content to create seamless loops, with not too much variations in the sound and perform manual mixing operation to make the looping perfect and pleasing at the ear.

The basic theory of this process can be decomposed like this:

- Find audio content that is suitable for looping (an ambiance, a repetitive sound, etc.).
- Locate a part in the audio content that would be appropriate for creating a loop: is must have not too much variation in volume and general tone. You must also make sure you have additional appropriate audio content before the place the loop will be actually created, as it will used to create the seamless transition.
- Create the loop markers in your audio editor, to serve as a reference.
- Now comes the tricky part: crossfade the end part of the loop with the audio content preceding the beginning of the loop (see fig. 10). The fade curves must be of constant energy (logarithmic). The best duration is to determine by ear, but 10% of the loop duration is a good starting point. It is also better to keep the crossfade duration between 500 milliseconds and 5 seconds to not alter the loop too much while keeping a pleasant transition.
- Make sure your loop begins and ends on zero-crossing points, to avoid clicks.
- Trim the audio content before and after your loop. Do not perform additional audio fades after this or it will alter the seamlessness of you audio loop!

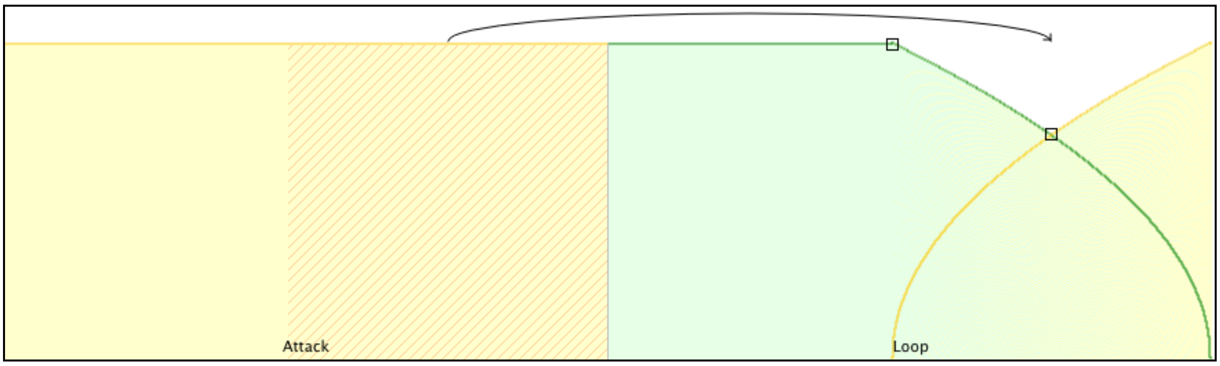


FIGURE 3.2 CROSS-ADE FOR SEAMLESS LOOPING

The most difficult stage is the crossfading part. Some audio editors can automate this process for you, on others you having to performs the different steps manually:

- Perform a fade-out of a given duration (for example 1 second) on the end of the loop, with a logarithmic curve.
- Select and copy the same duration of audio content just before the actual loop, in a new audio file.
- Perform a fade-in of the same duration on this new file, using a logarithmic curve.

- Mix this new file at the end of the loop, like on fig. 10.

Sometimes this process must be tweaked a little to find the optimum results.

3.1.4 Pre-rendering HRTF samples

If you want to pre-render various HRTF directions, for example to make a 3D audio beacon, you still have to prepare your sound with the processing stages explained in section 4.2. After you have a final “simple” version of your sound, you may proceed to the HRTF pre-rendering. In order to do so, you have to get an audio plugin able to perform the HRTF processing, like IRCAM Spat [8] or WaveArts Panorama [9]. Note that mono sounds are better suited for HRTF rendering.

FIGURE 3.3 WAVEARTS PANORAMA, A HRTF SPATIALIZATION PLUGIN

Then you have to configure the plugin for your needs: choose the right HRTF to use (most plugins include a generalized HRTF profile name Human or so), set up the reverb (a small reverb helps improving the 3D perception) and save this as a new pre-set.

Finally, you have to apply the result to your sound, changing each time the location of the source for each direction you want to pre-render and saving the result to a new audio file. You should end up with as many audio files as directions you chose to use.

3.2 Audio integration & User tests using IXE demonstrator

3.2.1 Scenario

We have designed an interactive audio scene based on the VeDi 1.0 game scenario (burger delivery) to demonstrate the features of our content authoring system and sound manager based on XML and its future integration in the VeDi demonstrator. In order to do this, we have recreated the OpenStreetMap navigation network of the Venturi city model (figure 3.4).

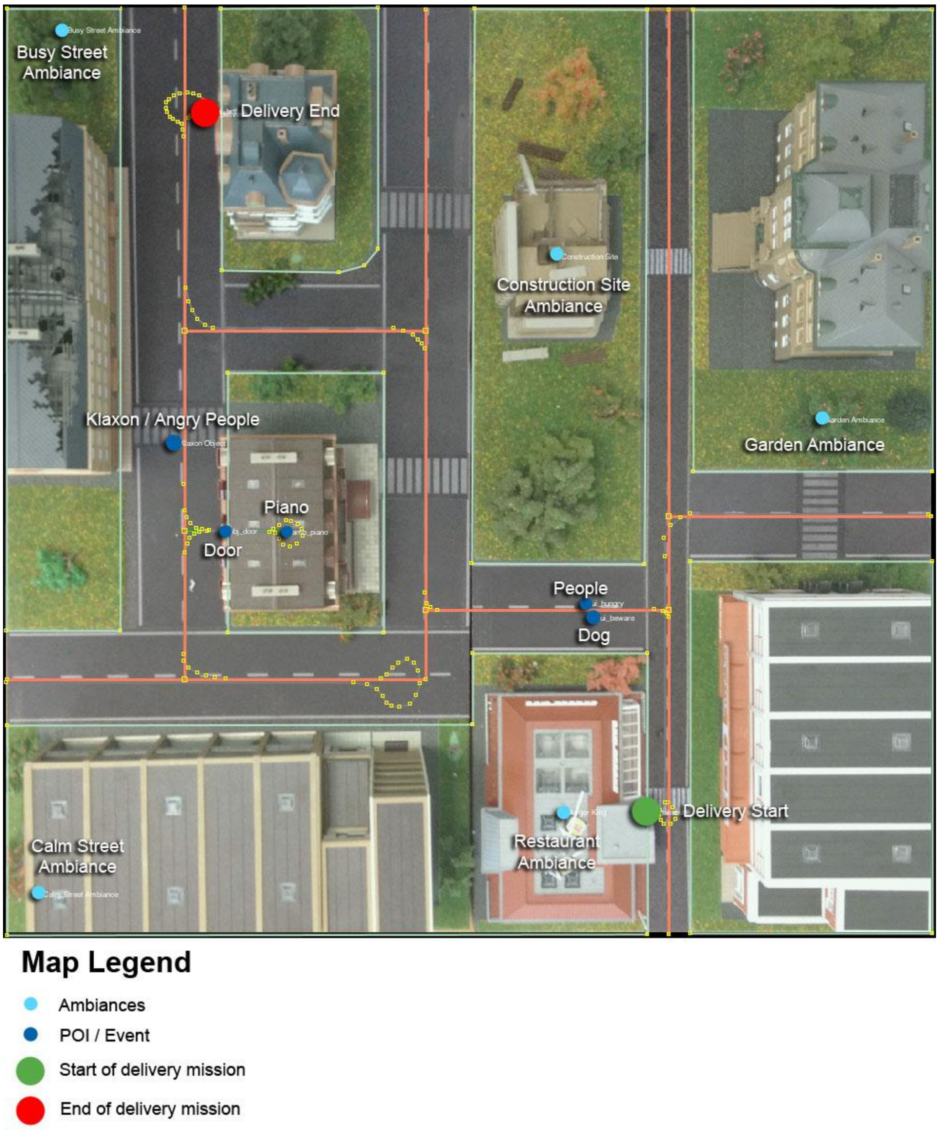


FIGURE 3.4 OSM NETWORK OF THE CITY

Based on this OSM network, we will trigger audio events using our IXE navigation demonstrator. The audio events will by interpreted by the sound manager, based on the XML audio document in MAUDL format we provide (see next section).

This scenario is designed to illustrate the following features of audio language:

- Event synchronization and triggering
- Simple stereo ambiances with distance attenuation (garden, busy street, calm street)
- 3D spatialized ambiances (restaurant, construction site)
- Sound object 3D spatialization with rear attenuation (klaxon / angry people, piano) or not (people, dog)
- Sound randomization (klaxon / angry people)
- Internal synchronization (piano with its reverb, door with ambiances)
- Sound queues with priority classes and validity discrimination (Dog, People)
- Pre-rendered HRTF beacon to indicate delivery target using multiple sound sources
- Mix groups (ambiance, objects, UI)

To demonstrate these features, we have built a scenario based on the VeDi 1.0 game, for which the goal is to deliver a burger from one location to another using a car. Using our IXE demonstrator, the car (represented here by the black head) will move on the OSM network, following this predefined route:

FIGURE 3.5 SIMULATION ROUTE OF THE SCENARIO

The simulation starts on the right at the green pin and ends on the left at the red pin. The car will move following the path specified in blue. First it will go to pick the burger to deliver at the restaurant. Once the burger is taken, the HRTF beacon starts to indicate the delivery location. On its way to the delivery target, the car will go past 2 events: the barking dog and a person talking. Depending of its speed, only the dog or both of these may be heard, as these sounds are put into a sound queue (see 3.2.2.3). The car will pass near the construction site and then arrive at the delivery point, which ends the beacon. The car will then move down the street, meet a random event (klaxon / angry guy) and then make a stop. The driver will go out of the car and enter the building. As it passes the door, the car engine and exterior ambiances stop. The driver will move forward to hear a piano player in a concert room with a lot of reverbed sound, walk around him and finally return outside. As he passes the door again, the exterior ambiances will be heard again and the car engine will restart. The car will finally move into the neighbourhood, make a U-turn as it was going the wrong way and then go out of the city.

3.2.2 Audio scene description

The audio scene is defined in the XML audio format, using object names and events mapped on the OSM document. This is the audio document we use in this scenario:

```
<?xml version="1.0" encoding="UTF-8"?>
<maudl xmlns="http://gforge.inria.fr/projects/iaudio/maudl/1.0" id="audio_stylesheet">

  <!-- use standard 3D rendering settings, with a little rear attenuation to improve focus effect -->
  <sounds rolloff="linear" listenerRearAttenuation="0.2" listenerPos="0 0 0" listenerLookAt="0 0 1" scale="1.0"
```

```
classOrder="objective danger info">

    <!-- local spatialized ambiances -->
    <sound id="amb.restaurant" play="amb.restaurant.trigger obj.door.started" replay="stop" render3D="advanced" min="3" max="30"
    src="amb_restaurant.wav" loopCount="-1"/>
    <sound id="amb.construction" play="amb.construction.trigger obj.door.started" replay="stop" render3D="advanced" min="10"
    max="42" src="amb_construction.wav" loopCount="-1"/>

    <!-- piano ambiance with reverb -->
    <sound id="amb.piano" play="amb.piano.trigger" stop="obj.door.trigger" render3D="advanced" min="1" max="4"
    src="amb_piano.wav" loopCount="-1"/>
    <sound id="amb.piano.reverb" play="amb.piano.started" stop="amb.piano.ended amb.piano.stopped" render3D="no" min="1"
    max="4" src="amb_piano_reverb.wav" loopCount="-1" volume="0.1"/>

    <!-- global ambiances, with attenuation only -->
    <sound id="amb.busystreet" play="amb.busystreet.trigger obj.door.started" replay="stop" render3D="simple" pan3DFactor="0"
    min="5" max="75" src="amb_busystreet.wav" loopCount="-1"/>
    <sound id="amb.calmstreet" play="amb.calmstreet.trigger obj.door.started" replay="stop" render3D="simple" pan3DFactor="0"
    min="5" max="80" src="amb_calmstreet.wav" loopCount="-1" volume="0.5"/>
    <sound id="amb.garden" play="amb.garden.trigger obj.door.started" replay="stop" render3D="simple" pan3DFactor="0.5" min="5"
    max="50" src="amb_garden.wav" loopCount="-1"/>

    <!-- spatialized ponctual random object -->
    <sound id="obj.klaxon" pick="random" play="obj.klaxon.trigger" render3D="advanced" pos="0 0 0" min="5" max="20">
      <soundsource src="obj_klaxon1.wav"/>
      <soundsource src="obj_klaxon2.wav"/>
    </sound>

    <!-- car engine (not spatialized since the scene listener is on the car) -->
    <sound id="obj.car" play="navigation.start resume_navigation.play obj.door.trigger" stop="pause_navigation.play" replay="stop"
    render3D="none" loopCount="-1" volume="0.1" src="obj_car.wav"/>

    <!-- pre-rendered HRTF spatialized beacon for the target -->
    <sound id="obj.hrtf" loopCount="0" pick="fixed" render3D="simple" pan3DFactor="0.0" play="ui.takeburger.ended obj.hrtf.ended"
    stop="ui.deliverburger.trigger" min="1" max="150">
      <soundsource src="hrtf_0.wav"/>
      <soundsource src="hrtf_1.wav"/>
      <soundsource src="hrtf_2.wav"/>
      <soundsource src="hrtf_3.wav"/>
      <soundsource src="hrtf_4.wav"/>
      <soundsource src="hrtf_5.wav"/>
      <soundsource src="hrtf_6.wav"/>
      <soundsource src="hrtf_7.wav"/>
      <soundsource src="hrtf_8.wav"/>
      <soundsource src="hrtf_9.wav"/>
      <soundsource src="hrtf_10.wav"/>
      <soundsource src="hrtf_11.wav"/>
      <soundsource src="hrtf_12.wav"/>
      <soundsource src="hrtf_13.wav"/>
      <soundsource src="hrtf_14.wav"/>
      <soundsource src="hrtf_15.wav"/>
    </sound>

    <!-- entering inside building -->
    <sound id="obj.door" play="obj.door.trigger" render3D="none" src="obj_door.wav"/>

    <!-- user interaction sounds -->
    <sound id="ui.takeburger" enqueue="ui.takeburger.trigger:ui.queue" render3D="none" src="ui_takeburger.wav" class="objective"/>
    <sound id="ui.deliverburger" enqueue="ui.deliverburger.trigger:ui.queue" render3D="none" src="ui_deliverburger.wav"
    class="objective"/>
    <sound id="ui.dog" enqueue="uidog.trigger:ui.queue" render3D="simple" src="ui_dog.wav" class="danger"/>
    <sound id="ui.people" enqueue="ui.people.trigger:ui.queue:-1:3" render3D="simple" src="ui_people.wav" class="info"/>
    <sound id="ui.pause" play="ui.pause.trigger pause_navigation.play resume_navigation.play" render3D="none"
    src="ui_pause.wav"/>

  </sounds>
  <queues>

    <!-- priority queue for playing UI sounds -->
    <queue id="ui.queue" autoPlay="true" sort="class" timeBase="realtime"/>

  </queues>
  <mixers>
```

```
    <!-- mixer for ambiance sounds -->
    <mixer id="mix.ambiance" volume="0.9">
      amb.burger, amb.construction, amb.busystreet, amb.calmstreet, amb.garden, amb.piano, amb.piano.reverb
    </mixer>

    <!-- mixer for objects sounds -->
    <mixer id="mix.objects" volume="0.5">
      obj.klaxon, obj.car, obj.door
    </mixer>

    <!-- mixer for UI sounds -->
    <mixer id="mix.ui" volume="1.0">
      ui.queue, ui.paused
    </mixer>

  </mixers>
</maudl>
```

Most of the sounds specified here correspond to the elements declared in the OSM document, with few exceptions (that will be explained later). We will now see in details the main features illustrated in this simulation.

3.2.2.1 3D Spatialization

There are various spatialization techniques used in this demonstration, with different goals:

- Global ambiances, with large listening radius, only use distance attenuation along their initial stereo rendering (garden, busy street, calm street). The goal is that these global ambiances are heard when the listener is within their range, and attenuate when he goes farther than the source of the ambiances.

- Local ambiances and some sound objects are spatialized using 3D positioning, distance attenuation with rear attenuation enabled (restaurant, construction site, piano, klaxon). Since these sounds only occur locally, they are spatialized using a natural positioning and attenuation model. The rear attenuation focus the listener attention on what is in front of him.

- Some user interactions are spatialized using simple 3D positioning with distance attenuation (dog, people). These represent local events, but since they are quite long sounds (people talk) they should be heard while the car is moving further from them from behind, hence why no read attenuation is used here.

- The audio beacon indicating the delivery target is composed by many sound sources representing 16 pre-rendered directions using HRTF processing. The selection of the right sound source to play depending of the listener orientation is done in the IXE application. Then distance attenuation is used to indicate whether the listener is close or far to its destination.

These different techniques are used to illustrate the possibilities offered by the MAUDL format in various contexts, and to compare the drawbacks and benefits of each method.

3.2.2.2 Synchronization and interactions

The MAUDL format is entirely based on an event synchronization system derived from SMIL. External events (sent by the application) are used to trigger the audio objects during the car navigation. In addition, the audio objects themselves generate internal events when they start, stop, etc. All these events are used to build the dynamic soundscape of this demonstration.

Internal events are here used to create interactions between sounds:

- When the car arrive at the restaurant to take the burger for delivery (*ui.takeburger.trigger* event), the sound ui.takeburger will be added to the sound queue ui.queue to be played. When it has finished playing (*ui.takeburger.ended* event), the HRTF beacon obj.hrtf will start playing in loop. It will end as soon as the burger is delivered (*ui.deliverburger.trigger* event).

- When the listener goes past the door (*obj.door.trigger* event), the sound obj.door is played. When this sound is started (*obj.door.started* event) the first time, the ambiances already playing and the car engine sound obj.car will stop, and the piano sound amb.piano will start. The reverb of the concert room amb.piano.reverb is also synchronized on the start and stop of the piano sound. When the listener goes out of the building, the sound obj.door is started again, caused all the previously stopped ambiances and the car engine to start again.

Using internal makes it easy to create complex interactions between sounds in a simple way.

3.2.2.3 Sound queue

In order to demonstrate the priority and filtering system of sound queues, we have created one named ui.queue to play the user interaction sounds. We have defined 3 priority classes, ordered by decreasing priority: objective, danger and info.

- The sounds ui.takeburger and ui.deliverburger are set in the *objective* class, as they must always be played with the highest priority. However, this is just for safety and these sounds should not be part of the test cases explained later.

- The sound ui.dog is set in the *danger* class, since dangerous cues should be notified with high priority.

- The sound ui.people is set in the *info* class, as it is only informative and may be skipped for higher priority sounds. This sound has a validity distance of 3 meters. If the listener has moved more than this distance before this sound can be played, it will be automatically skipped.

This setup allows performing 3 test cases to assert that the queue behaviour is the one expected:

- If the car moves very fast, the sound ui.dog and ui.people will be put in the sound queue at the same time. But due to the higher priority class of the ui.dog sound, it will always be played first.

- If the car moves fast, the sound ui.dog will obviously be played and the sound ui.people may be skipped completely if it goes farther than the 3 meters radius before the sound ui.dog has finished playing.

- If the car moves slowly and is still within the 3 meters radius after the sound ui.dog has finished playing, the sound ui.people will be played.

Finally, simply changing the movement speed of the car modify the playback behaviour of the sound queue, allowing to put in evidence the benefits of its usage in a navigation context.

3.2.3 User tests

In order to test the perception of audio of this demonstration scenario by various users, we have put in place a series of tests based on a questionnaire. The goal of these tests to better understand how the various audio elements are perceived, how effective are the different spatialization methods and what can be achieved to improve the user immersion in such scenarios.

3.2.3.1 Testing Methodology

We will present the demonstration scenarios to a group of testing users. The general soundscape context (areas of the city) and simulation objective (driving inside a city, delivering burger) is explained and the unmarked map of the city will be showed, so the users know have an idea of what they should expect to hear. They will hear the simulation a first time with the car moving at an average speed, then a questionnaire (see next section) will be given to them for answer. They will hear the simulation a second time before answering the questions.

3.2.3.2 Questionnaire

After they have heard the simulation a second time, they will be asked to answer each question using a scale ranging from one to ten, one corresponding to “I strongly disagree” and ten corresponding to “I strongly agree”.

> S1: “I have a good spatial conception of the sound locations.”

Before the users will answer this question, we will ask them to concentrate on the spatialized ambiances (restaurant, construction site) and punctual sound objects like the piano.

> S2: “I can determine if I am going closer or farther of a sound source easily.”

To answer this question, we will ask the users to focus on the ambiances sounds and punctual sound objects, for example the construction site.

> S3: “I can determine approximately in which area of the city the car is currently present.”

During their second hearing of the simulation, we will make 3 pauses at different places of the map: before arriving at the restaurant, near the construction site, and after delivering the burger. We will each time ask the users if they can approximately tell in which area of the city they think they currently are, and after the end of the hearing we will show them the simulation route and the correct answers, so they can see if they were right or not.

> S4: “I can determine easily if I’m inside or outside.”

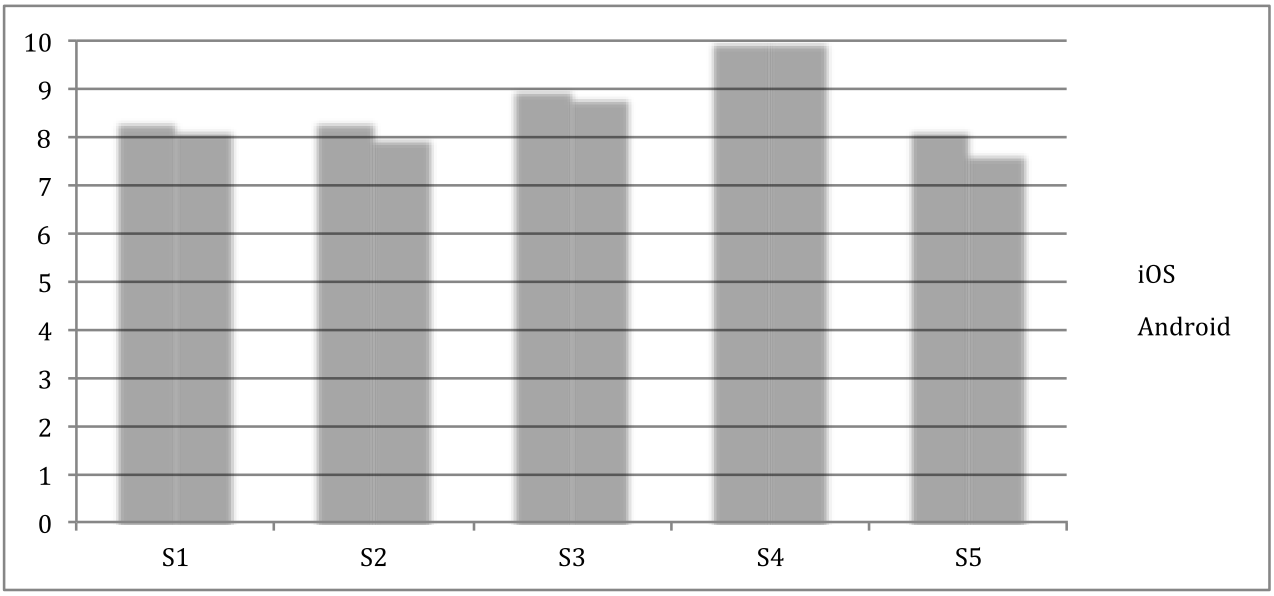
During the briefing before the tests, the users will be told that at some point the user will enter a building, do something inside and return outside. We will ask them after the second hearing at which point they think this occurred, and tell them the correct answer (between the door sounds, when they hear the piano), so they can make their opinion.

> S5: “When I am hearing multiple sounds concurrently, I can distinguish and understand them without effort.”

After the second hearing, we will ask them to tell with their own words what they think occurred during the whole scenario, and after that we will remind them what was the original scenario so they can compare and make their opinion.

3.2.4 Results and conclusion

We have performed the tests with a group of 12 users, 6 for iOS demonstrator and 6 participants for Android version. The mean score for each question can be found in the following graphic:



From these results, it can be found that iOS and Android results are similar. The majority of participants can localize easily sound objects with the different spatialization methods. They understand that the volume attenuation allows determining if they go closer or farther of a sound source. The general soundscape with specific ambiances helps users to know quickly in which area the car is present. Due to the sounds chosen, all participants can determine if they are inside or outside easily. Finally, some users think that it is sometimes difficult to distinguish and understand multiple sounds that are playing concurrently.

The objective of the questionnaire was to find out how the various audio elements are perceived and how effective are the different spatialization methods. All participants think that adding audio helps to make the user experience in a game more immersive. The MAUDL XML format and its sound manager implementation are very useful to describe a rich soundscape. This demonstration scenario shows the various possibilities offered by the audio language and its usage in a concrete example. It is also a first step before the audio integration into the VeDi game demonstrator.

3.3 3D HRTF Audio integration & User tests

3.3.1 Scenario

A second interactive audio scene based on the VeDi 1.0 game scenario (burger delivery) was designed to test sound objects using pre-rendering HRTF samples for different directions. Figure 3.6 below shows four sound objects using HRTF samples (bird, cat, dog, church). The OpenStreetMap document describing the navigation network of the city contains these four sound objects as POIs. Circles show the area where the objects can be heard.

FIGURE 3.6 MAP OF THE CITY WITH HRTF SOUND OBJECTS

Based on the OSM document, a node element localizes a sound object and defines a distance in meter to trigger audio events:

```
<node id='-362' action='modify' visible='true' lat='0.0042435083850540006' lon='0.0101376635687379665'>
   <tag k='cat' v='hrtf' />
   <tag k='triggering' v='11' />
 </node>
```

We will trigger audio events using tag elements in our IXE navigation demonstrator. In this example, tags are parsed and the audio events “cat.hrtf.trigger” will be created using <tag k='cat' v='hrtf' />. When the listener enters in the triggering circle of this node (<tag k='triggering' v='11' />), the event is sent to the sound manager and will be interpreted based on the XML audio document in MAUDL format we provide.

To demonstrate the pre-rendered HRTF feature, we have built a scenario based on the VeDi 1.0 game, for which the goal is to move around the city from one location to another using a car. Using our IXE demonstrator, the car (represented here by the black head) will move on the OSM network, following a predefined route or a route computed by the OSM router embedded in the application.

Using our simulator, the car will move following the path specified in blue. When the simulator computes a new location and orientation, this information is sent to the Sound Manager to set the listener location and orientation. Then, our application looks at POIs information to trigger events with the new location simulated. Figure 3.7 is a screenshot of the IXE demonstrator where the simulator is following a predefined route.

FIGURE 3.7 SIMULATION ON A PRE-DEFINED ROUTE

3.3.2 Audio scene description

The audio scene is defined in the XML audio format, using object names and events mapped on the OSM document. This is the audio document we use in this scenario:

```
<?xml version="1.0" encoding="UTF-8"?>
<maudl xmlns="http://gforge.inria.fr/projects/iaudio/maudl/1.0" id="audio_stylesheet">

   <!-- use standard 3D rendering settings, with a little rear attenuation to improve focus effect -->
   <sounds rolloff="linear" listenerRearAttenuation="0.2" listenerPos="0 0 0" listenerLookAt="0 0 1" scale="1.0" classOrder="objective
danger info">


      <!-- car engine (not spatialized since the scene listener is on the car) -->
      <sound id="obj.car" play="navigation.start resume_navigation.play" stop="pause_navigation.play" replay="stop" render3D="none"
loopCount="-1" volume="0.1" src="obj_car.wav"/>

         <!-- pre-rendered HRTF spatialized bird -->
      <sound id="bird.hrtf" loopCount="0" pick="fixed" render3D="simple" pan3DFactor="0.0" play="bird.hrtf.trigger bird.hrtf.ended"
stop="pause_navigation.play" min="5" max="20">
                <soundsource src="hrtf_bird_0.wav"/>
                <soundsource src="hrtf_bird_1.wav"/>
                <soundsource src="hrtf_bird_2.wav"/>
                <soundsource src="hrtf_bird_3.wav"/>
                <soundsource src="hrtf_bird_4.wav"/>
                <soundsource src="hrtf_bird_5.wav"/>
                <soundsource src="hrtf_bird_6.wav"/>
                <soundsource src="hrtf_bird_7.wav"/>
```

```
                <soundsource src="hrtf_bird_8.wav"/>
                <soundsource src="hrtf_bird_9.wav"/>
                <soundsource src="hrtf_bird_10.wav"/>
                <soundsource src="hrtf_bird_11.wav"/>
                <soundsource src="hrtf_bird_12.wav"/>
                <soundsource src="hrtf_bird_13.wav"/>
                <soundsource src="hrtf_bird_14.wav"/>
                <soundsource src="hrtf_bird_15.wav"/>
        </sound>

        <!-- pre-rendered HRTF spatialized dog -->
    <sound id="dog.hrtf" loopCount="0" pick="fixed" render3D="simple" pan3DFactor="0.0" play="dog.hrtf.trigger dog.hrtf.ended"
stop="pause_navigation.play" min="5" max="20">
                <soundsource src="hrtf_dog_0.wav"/>
                <soundsource src="hrtf_dog_1.wav"/>
                <soundsource src="hrtf_dog_2.wav"/>
                <soundsource src="hrtf_dog_3.wav"/>
                <soundsource src="hrtf_dog_4.wav"/>
                <soundsource src="hrtf_dog_5.wav"/>
                <soundsource src="hrtf_dog_6.wav"/>
                <soundsource src="hrtf_dog_7.wav"/>
                <soundsource src="hrtf_dog_8.wav"/>
                <soundsource src="hrtf_dog_9.wav"/>
                <soundsource src="hrtf_dog_10.wav"/>
                <soundsource src="hrtf_dog_11.wav"/>
                <soundsource src="hrtf_dog_12.wav"/>
                <soundsource src="hrtf_dog_13.wav"/>
                <soundsource src="hrtf_dog_14.wav"/>
                <soundsource src="hrtf_dog_15.wav"/>
        </sound>

        <!-- pre-rendered HRTF spatialized cat -->
    <sound id="cat.hrtf" loopCount="0" pick="fixed" render3D="simple" pan3DFactor="0.0" play="cat.hrtf.trigger cat.hrtf.ended"
stop="pause_navigation.play" min="5" max="15">
                <soundsource src="hrtf_cat_0.wav"/>
                <soundsource src="hrtf_cat_1.wav"/>
                <soundsource src="hrtf_cat_2.wav"/>
                <soundsource src="hrtf_cat_3.wav"/>
                <soundsource src="hrtf_cat_4.wav"/>
                <soundsource src="hrtf_cat_5.wav"/>
                <soundsource src="hrtf_cat_6.wav"/>
                <soundsource src="hrtf_cat_7.wav"/>
                <soundsource src="hrtf_cat_8.wav"/>
                <soundsource src="hrtf_cat_9.wav"/>
                <soundsource src="hrtf_cat_10.wav"/>
                <soundsource src="hrtf_cat_11.wav"/>
                <soundsource src="hrtf_cat_12.wav"/>
                <soundsource src="hrtf_cat_13.wav"/>
                <soundsource src="hrtf_cat_14.wav"/>
                <soundsource src="hrtf_cat_15.wav"/>
        </sound>

                        <!-- pre-rendered HRTF spatialized church -->
    <sound id="church.hrtf" loopCount="0" pick="fixed" render3D="simple" pan3DFactor="0.0" play="church.hrtf.trigger
church.hrtf.ended" stop="pause_navigation.play" min="5" max="50">
                <soundsource src="hrtf_church_0.wav"/>
                <soundsource src="hrtf_church_1.wav"/>
                <soundsource src="hrtf_church_2.wav"/>
                <soundsource src="hrtf_church_3.wav"/>
                <soundsource src="hrtf_church_4.wav"/>
                <soundsource src="hrtf_church_5.wav"/>
                <soundsource src="hrtf_church_6.wav"/>
                <soundsource src="hrtf_church_7.wav"/>
                <soundsource src="hrtf_church_8.wav"/>
                <soundsource src="hrtf_church_9.wav"/>
                <soundsource src="hrtf_church_10.wav"/>
                <soundsource src="hrtf_church_11.wav"/>
                <soundsource src="hrtf_church_12.wav"/>
                <soundsource src="hrtf_church_13.wav"/>
                <soundsource src="hrtf_church_14.wav"/>
                <soundsource src="hrtf_church_15.wav"/>
        </sound>
```

```
    </sounds>


    <mixers>

        <!-- mixer for ambiance sounds -->
        <mixer id="mix.ambiance" volume="0.9">
          bird.hrtf, church.hrtf, dog,hrtf
        </mixer>

        <!-- mixer for objects sounds -->
        <mixer id="mix.objects" volume="0.5">
          obj.car, cat.hrtf
        </mixer>

    </mixers>

</maudl>
```

Most of the sounds specified here are using pre-rendering HRTF sound sources for 16 different directions according to listener position and orientation. Samples are created using section *4.4 Pre-rendering HRTF samples of T.5.1.2 3D Audio Content Creation.* Sound sources are ordered in clockwise to cover all the directions from 0 to 360 degrees. For example, the first source `<soundsource src="hrtf_church_0.wav"/>` is the pre-rendering HRTF sample for a sound at 0 degree (in front of the listener), `<soundsource src="hrtf_church_12.wav"/>` is the HRTF sample for a sound at 270 degrees (at the left of the listener).

3.3.2.1 Synchronization and interactions

The MAUDL format is entirely based on an event synchronization system derived from SMIL. External events (sent by the application) are used to trigger the audio objects during the car navigation. In addition, the audio objects themselves generate internal events when they start, stop, etc. All these events are used to build the dynamic soundscape of this demonstration.

- External events trigger a sound to start playing (`play="church.hrtf.trigger"`).
- Internal events are here used to loop a sound when it finishes playing (`play="church.hrtf.ended"`).

Also, an algorithm is implemented in IXE demonstrator to select the sound source to play according to the angle between the source and the listener positions for each HRTF sound objects:

```
    // update all HRTF objects
    for (ADSound *hrtfSound in hrtfSounds) {
        float hrtfAngle = 0.0;

        // set the listener as the center of our world
        point2f_t newPos = point2f_init(hrtfSound.position.x - soundManager.listenerPosition.x, hrtfSound.position.z - soundManag-
er.listenerPosition.z);

        if (!(fequalzero(newPos.x)|| fequalzero(newPos.y))) {
            // get the angle between the source and the listener positions
            hrtfAngle = atan2f(newPos.y, newPos.x);
            if (!(fequalzero(soundManager.listenerOrientation.x) && fequalzero(soundManager.listenerOrientation.z))) {
                // get the angle of the direction the listener is looking at
                float orientationAngle = atan2f(soundManager.listenerOrientation.x, soundManager.listenerOrientation.z);
                hrtfAngle = hrtfAngle - orientationAngle;
            }
        }

        hrtfAngle = deg_angle_normalize(RAD_TO_DEG(hrtfAngle) - 270);
        int dir = (int)roundf(hrtfAngle * 16.0 / 360.0) % 16;
                // set the next soundsource to play
        [hrtfSound setNextSoundsource:dir];
    }
```

3.3.3 User tests

The goal of these tests is to evaluate how effective is the 3D spatialization using pre-rendered HRTF 3D sounds. We will compare a scenario using HRTF 3D sounds with a scenario using simple stereo sounds and distance attenuation to determinate if pre-rendered HRTF sounds improve the user experience.

3.3.3.1 Testing Methodology

We will present the demonstration scenarios to a group of testing users. The soundscape (four sound objects) and the simulated route are explained and the map of the city will be showed, so the users have an idea of what they should expect to hear. Then a questionnaire (see next section) is given to them in order to help them to concentrate on what we expect. They will hear the simulation a first time (using simple stereo sounds) with the car moving at an average speed, then they will answer the questionnaire. They will hear the simulation a second time (using HRTF 3D sounds) and they will answer the same questionnaire.

3.3.3.2 Questionnaire

After they have heard each scenario, they will be asked to answer each questions using a scale ranging from one to ten. One corresponding to “I strongly disagree” and ten corresponding to “I strongly agree”.

S6: “I have a good spatial conception of the sound locations.”

Before the users will answer this question, we will ask them to show where the four sound objects are on the map.

S7: “I have the feeling that the sources are outside of my head.”

During their hearing, we will make a pause near the dog and the church to help users to focus on these two sound objects.

S8: “I can determine easily if a source is in front of or behind the car.”

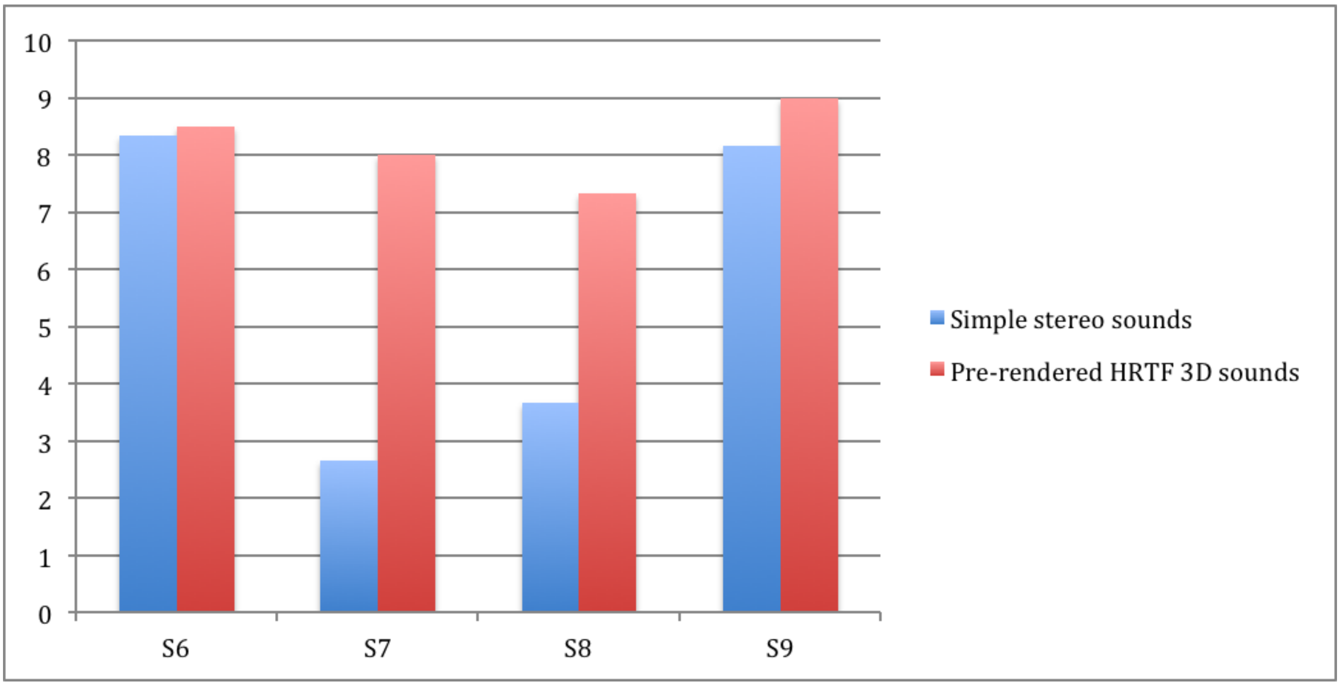
During their hearing, we will make a pause near the bird and we use gyroscope to change the listener orientation. Then, we will ask the participant to tell if the sound object is in front of or behind him.

S9: “The listening experience gives me the feeling of moving.”

After the route simulation, we will ask them to tell if the sound objects give them the feeling of moving by coming closer of sound objects or not.

3.3.4 Results and conclusion

We have performed the tests with a group of 6 users. The mean score for each question can be found in the following graphic:



From these results, we can compare the scenario using simple stereo sounds with the one using pre-rendered HRTF 3D sounds. In both cases, users can localize easily sound objects and have the feeling of moving. HRTF samples make the user experience realer because participants have the feeling that sound sources does not come from their headphones, but rather from a speaker or an object close to them. HRTF samples also help users to determine if a source is in front or behind them.

The MAUDL XML format and its sound manager implementation are very useful to describe a rich soundscape. This demonstration scenario shows the possibility of using pre-rendered HRTF 3D sounds in the audio language and its usage. From the questionnaire, we find out that the audio perception is realer if we use HRTF 3D sounds because users have less the feeling of using headphones and they can determine easily where a sound is, especially if its location is in front of or behind them.

4 Experiments with visually impaired people (June and July 2012 Grenoble)

This section considers audio for indoor and outdoor navigation. The objective of these tests was to understand how visually impaired people will use a mobile phone audio navigation system on an indoor-outdoor route within an unknown environment.

4.1 Methodology

For the tests conducted with professional ergonomists, we use a prototype of the application that will be developed in task 4.1 of WP4.

The objective was threefold:

- Removing critical ergonomics errors of the voice application audio-guide, for example:
  a. Can the user navigate in an unknown environment with the audio guide?
  b. Does the application speak the same language as the user?
  c. Can the users explicitly control the various functions of the application?

- Evaluate how the users felt about the usefulness of the application
  a. Effectiveness of the implementation of the task felt
  b. Perceived ease of use

- Test the prototype with the goal to integrate it into a continuous improvement cycle

The tests were conducted with five visually impaired people, three of them using a white stick to sense obstacle after validation of the route with a visually impaired pre-tester (the one shows on the figure 4.1).

FIGURE 4.1 VISUALLY IMPAIRED PRE-TESTER

4.1.1 Plan of a typical day

10 AM-11 AM	Initiation stage Presentation of the application and of the experimentation's context Training on the user interface and audio navigation: Calibration and walk in a corridor
11 AM-12 PM	Preparation of the course Route A Route B
12 PM-1.30 PM	Lunch
1.30 PM – 3 PM	Testing on both routes
3 PM – 4.30 PM	Interview after testing

4.1.2 Routes

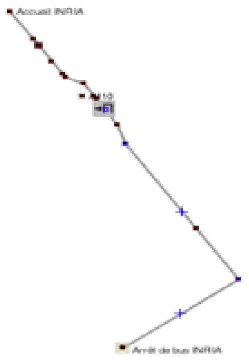
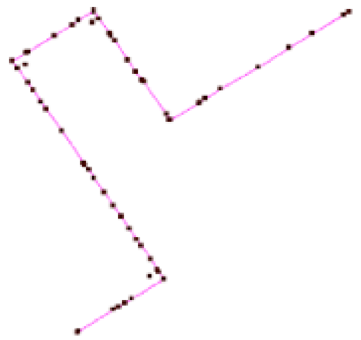
Each route is performed 2 times with a different 3D audio guide each time: continuous beacon enabled versus beacon activated only when the user stops walking. The beacon is a sound that indicates the position of the body relative to the direction of travel. The sound is more or less strong, with a change in frequency according to the difference between the orientation of the body and the route.

This use of 3D audio can be better understood by hearing this video:
http://www.youtube.com/watch?v=h2b8yfCauZ8

A discovery phase was introduced at the start of testing. It consists of two phases:

- A walking model calibration phase requiring walking along a straight line of 30 meters
- A straight-line route of 50 meters to better understand the use of the beacon and vocal instructions were presented. For the test in real conditions, two routes were available. The testers were able to use the route's simulator incorporated in the prototype application to learn it before trying it in the real world.

The table below presents the two routes used for testing:

Route A : From the bus stop to the INRIA reception	Route B : Inside the INRIA building
	
This route is 60 meters long with two stairs. The first is outside with an unusual step length and no ramp. It will be made more accessible very soon.	This route is 153 meters long in 2 meters wide corridors and open spaces.
Segment 1 = 12m Segment 2 = 38m Stairs 1: 10 steps 70cm wide Stairs 2: 9 steps, one landing of 1m then 9 steps	Segment 1 = 20m Segment 2 = 50m Segment 3 = 18m Segment 4 = 25m Segment 5 = 40m

4.1.3 Interview post-tests

The following questions were used to guide the interview and explore the utility of the application perceived by the users after the test session:

- Does the application was easy to use?
- What are the features that surprised you?
- What do you not like about the application?
- Does the use of the application allowed you to guide you effectively?
- Was its use comfortable?
- Did you understand quickly how the application works?
- Were the information and texts accessible and understandable?
- Do you like texture of sounds? Would you change? Do you have any preferences?
- Did you enjoy using the application?
- What would you like to improve? For the wording of instructions, texture of sounds, guide mode, calibration.

4.2 Conclusion and recommendations

4.2.1 Key points for user testing

Usability testing with IXE INRIA navigation application led to 5 elements of conclusion:

- IXE should be improved at the calibration and localization level. Indeed, the system calculating the positioning of users is taken in default by two identified factors: problems with walking speed and those linked to irregular orientations received by the Smartphone hung on the torso of the user. This generates critical errors in the announcement of the instructions, the use of 3D audio and renders inoperative the system. System errors are also too heterogeneous for users to adapt the system.

- The wording of the instructions is efficient. Apart from few problems easily correctable, the announced instructions are recognized and interpreted correctly by users. Therefore, the tracking system must require more work.

- Route A is the best route as there was usually 3 meters maximum error. Route B has been a source of error for the system that was successful three times out of 10 to guide users into room G220.

- The users found the system useful. All participants found relevant the use of such a system in an indoor environment. Moreover, they expect that IXE can describe more precisely the environment in order to benefit from the exploration of a place or a building.

- IXE's UI available on the Smartphone and headphones can be improved and could be made more accessible to a visually impaired people. It is impossible for a disabled person to calibrate the device or access management options for the route.

4.2.2 Vigilance points for the design

These remarks apply to technical defects that hinder the efficient use of the system and are not dependent on the design of audio and/or visual interfaces.

- The calibration process does not take into account the conditions of real steps to end-users. Indeed, it appears that the blind tend to walk unevenly, especially in unfamiliar places. Therefore, the system must adapt to this irregularity to allow an effective presentation of audio instructions. Otherwise the system leads to critical errors, making the system unusable. It is therefore necessary to allow the system to anticipate this irregularity.

- The gyro system is a sensitive and smartphones may malfunction if the phone is shaken by something other than the walking pace. Therefore, it is necessary to reduce the risk of misdirection through real-time filtering of gyroscope data

- The cycle of mobility is not fully integrated. Indeed, the model "walk-stop and take stock of the situation-go back or forward" is not fully taken into account and this makes the walk not enough secure. This is the phase "stop and take stock of the situation" that is missing, because it is currently impossible to stop during the walk to review past and future sections of the route.

- The calibration phase must be integrated into the design of interfaces. Indeed, the current system does not allow a blind person to calibrate the system himself. It lacks the notions of distance and voice guidance to define the beginning and the end of the calibration phase.

4.2.3 Recommendations

4.2.3.1 Calibration and respect of the course by the system

The calibration system is not operating due to the walking pace of the users. Therefore it must overcome this irregularity by using a repositioning system (automatic or manual) on the map.

- The system must allow users to reposition the localisation system themselves in the environment when they stopped walking. To do this, they must explore their surroundings (touch, asks third party, etc.) to validate their position. This command must be run through the headphones or the screen of the smartphone. For example, users can scroll through the instructions or POI via the buttons on the headset then validate their choice by double clicking.

- Propose the user to manually reposition themselves in key areas of the route through the use of buttons. We need to ask the user to validate sections of the course. For example, if the system tells the user to go down using a stairway consisting of 10 steps and located at a distance of 5 meters, we must ask the user to confirm the meeting between the stairs and him. This implies that the user interrupts its progress to validate the obstacle. It is also necessary that route's instructions of the audio-guide incorporate elements sufficiently identifiable to allow the user to validate the stages of his route.

- Ideally, the system can automatically capture elements of the environment to reposition the user as it travels through waypoints. For example, the project explores the Venturi system possibilities to put visual helping tags at key stages of the course.

4.2.3.2 Interpretation of Instructions

- Users must understand sounds and vocalizations used in the system. The visually impaired have different interpretations because of the absence of vision. For example, it is better able to understand "a quarter turn to the right" than "turn right".
- We must focus on a regular announcement of distances. For example: 5 meters, 10 meters and 15 meters. Instructions using announcement of heterogeneous distances such as "in 8 meters turn right then, in 1 meter turn left " should be avoided.
- Avoid announcing predictive instructions at less than 10 meters.
- You have to give distances before actions. For example "in 10 meters turn left".
- Instructions for a change of direction must be given in real time with a slightly ahead of time to allow people to anticipate the turn.
- To change direction, you must provide instructions including the representation of the user's body. For example: "upstairs, turn half to your left into the corridor B".

4.2.3.3 Soundscape

The navigation system uses four types of audio: vocal instructions, announcement of POI, sonification of the steps and positioning beacon. These four kinds of audio information can conflict when they overlap in information-rich situations or when the system bugs. This causes confusion for the user.

- The sonification of steps reassures users when they start walking or when they restart after a shutdown. But it is less useful when the user walks because he wants to focus eon other sounds. Therefore, we propose to stop this sonification after the second vocal instruction.
- A beacon indicating the true heading is useful to help people to go straight or to turn right. However, it depends on the quality of the direction computed from the gyroscopes. It is therefore necessary to know precisely when the algorithm is susceptible of not producing the good heading and then disable the beacon.
- The announcement of POIs is made in a timely manner by the system. It is preferable to announce the label of a POI (Office A) or group of POIs (multiple offices) rather than giving the information that there is POI. For example, it is better to say "4 offices" rather than "4 POI". Then the user can stop and explore his environment following the first information.
- The exploration phase must be based on three distances: less than a meter, between 1-5 meters and more than 5 meters. It must adapt quickly to the rotation of the user and vocally indicate the spatial distribution of POI. To do this, the application must allow the user to choose one among the three distances and then listen to the short list of POIs available. Then, during exploration, the POI is announced when the user is in front of him, but not before.

4.2.3.4 Recommendations for the accessibility of IXE under iOS

Voice Over [10] on iOS is a very effective system, however, it requires a significant exploration time with the user's finger. Items must be findable quickly.

- Increase the size of buttons on the interface to facilitate the exploration and selection of components with Voice Over. A good size for buttons is 2 * 2cm.
- Create buttons to scroll lists. This allows faster access in mode "Voice-over". For example: buttons "up" and "down".
- For more flexibility, it is necessary to integrate two ways to activate one action: one on the headphones and one on the screen of the smartphones.
- Provide a headphones with buttons easily distinguishable by touching them for a better understanding of commands

5 Results and Conclusions

User studies and expert evaluation were undertaken to understand users' expectations and reactions in order to improve the efficiency and quality of AR applications. This report is an input to T3.2 and WP2. It provides results, recommendations and requirements for AR gaming (see 2.5 and 2.6), interactive audio scene (see 3.2.4 and 3.3.4) and navigation (see 4.2), taking into account current and future visual and audio technologies in order to fulfil user needs or disabilities to a maximum.

6 References

[1] Azuma, R. A Survey of Augmented Reality, August 1997.
[2] Cairns, P., Cox, A. L., Research Methods for Human-Computer Interaction, Cambridge University Press, 2008, 12.
[3] Kaptelinin, V, Nardi, B., Macaulay, C., The Activity Checklist: A Tool for Representing the "Space" of context, interactions, July, 1999
[4] Lazar, J., Feng, H. J., Hochheiser, H., Research Methods – in Human-computer interaction, WILEY (2010), 144-150
[5] Adobe Audition, http://www.adobe.com/fr/products/audition.html
[6] Steinberg Wavelab, http://www.steinberg.net/en/products/wavelab.html
[7] Audacity, http://audacity.sourceforge.net
[8] IRCAM Spat, http://www.fluxhome.com/products/plug_ins/ircam_spat
[9] WaveArts Panorama, http://wavearts.com/products/plugins/panorama/
[10] Voice Over, http://www.apple.com/accessibility/voiceover/
[11] Corona: Audio Augmented Reality in Historic Sites, Florian Heller and Jan Borchers, MobileHCI Workshop on Mobile Augmented Reality: Design Issues and Opportunities, Stockholm, Sweden, August 2011.

7 Appendix for AR gaming study

7.1 Interview questions

- How was your gameplay experience?
  - Fun?
  - Difficult?
  - Something was annoying?
  - Enjoyable?
- Do you play mobile phone games?
  - What kind of mobile games do you enjoy playing?
  - When?
  - While waiting / in the transport?
  - How often?
- Do you think you would play these games?
  - When?
  - Where?
  - How often?
  - If the marker is not needed?
- Are you usually buying new inventions or do you wait until the technology is mature?
- Can you comment on how easy or hard it was to aim before shooting / smashing?
- Did you notice any technical malfunctions?
  - Which ones?
  - The recognition of the marker was lost?
  - Instability of spatial placement of graphic overlays?
  - camera picture lag?
  - Correct placement of graphic overlays?
- Did some events cause disturbance for you?
  - Why?
  - Why not?

- Is this a game that you would recommend to your girlfriend or boyfriend?
- Describe the advantages/disadvantages using phone/tablet?

7.2 Questionnaires

Questionnaire, AR game: ____________________ with ______________

Please mark how these statements fit your experience.

Q1: I easily understood how to play the game. Agree ______________________Disagree

Q2: I became physically tired of playing the game. Agree ______________________Disagree

Q3: Technical instabilities irritated me during the game. Agree ______________________Disagree

Q4: The detection of the game board was stable enough for me. Agree ______________________Disagree

Q5: The location of the virtual objects was stable enough for me. Agree ______________________Disagree

Q6: The game responded fast enough on my input. Agree ______________________Disagree

Q7: The camera picture followed my movements. Agree ______________________Disagree

Please write your positive experiences of playing the game.

____
____
____

Please write your negative experiences of playing the game.

____
____
____

Other comments

____
____
____

7.3 NASA-TLX

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

7.4 Informed consent

Informed consent

Informed consent for the study of ”Augmented Reality gaming”

I hereby confirm that I have received information about the study mentioned above. I am aware that:

- the information about me that is collected as part of the study will only be studied by the research team at Lund University (Design sciences) and the research team of Sony Mobile Communications involved in the study
- the recorded video from the study will only be viewed by those who are involved in the study
- participation is voluntary and that I may at any time end my participation in the study

Date:_______________

Name:____

Signature:____

7.5 Graphs from questionnaires phone form factor

Below graphs are presenting the results for each question that participants answered for the phone form factor.

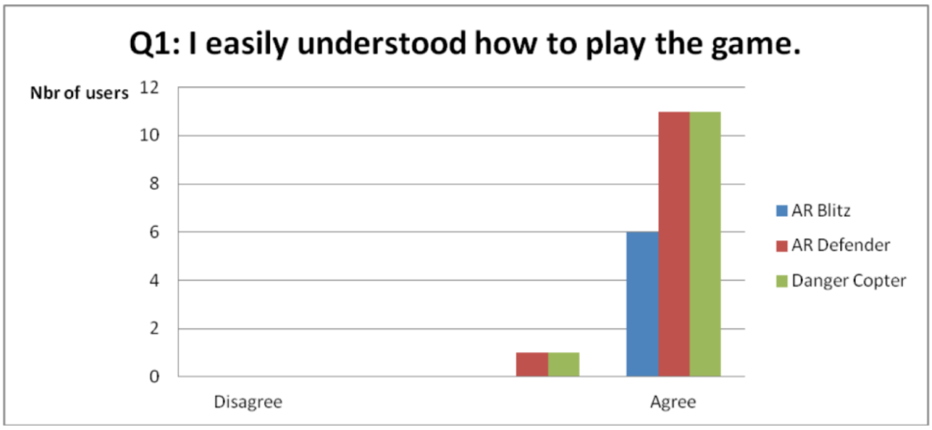


FIGURE 7.1 RESULTS FOR Q1: I EASILY UNDERSTOOD HOW TO PLAY THE GAME

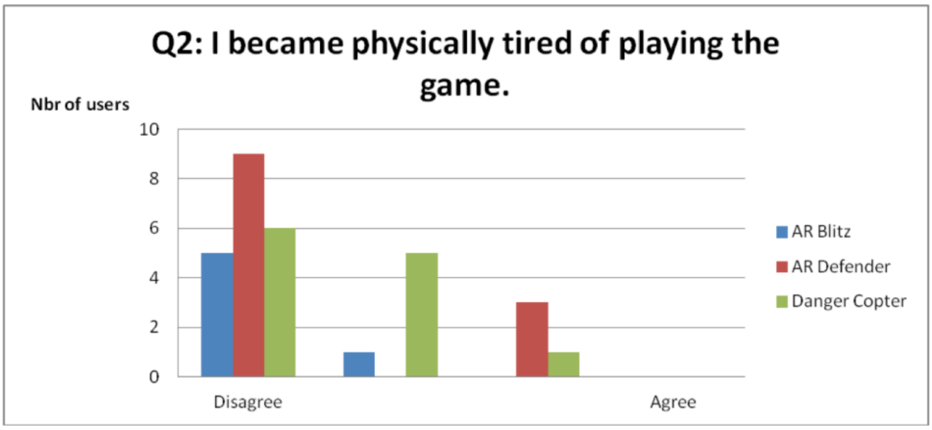


FIGURE 7.2 RESULTS FOR Q2: I BECAME PHYSICALLY TIRED OF PLAYING THE GAME

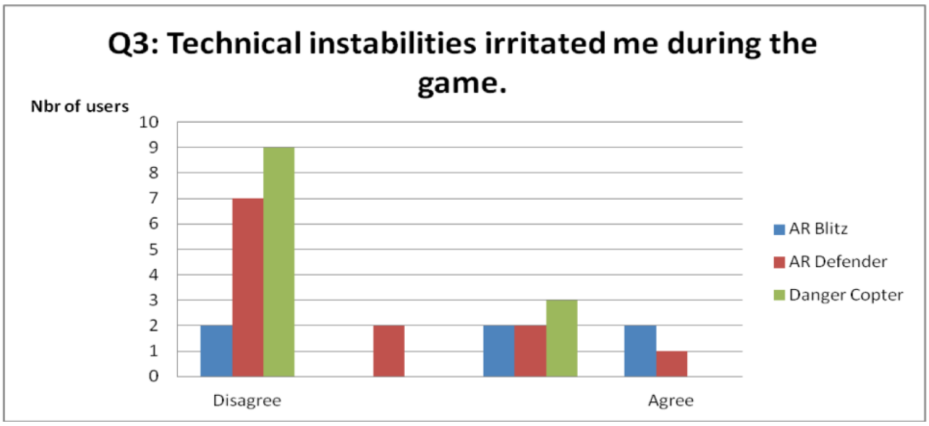


FIGURE 7.3 RESULTS FOR Q3: TECHNICAL INSTABILITIES IRRITATED ME DURING THE GAME

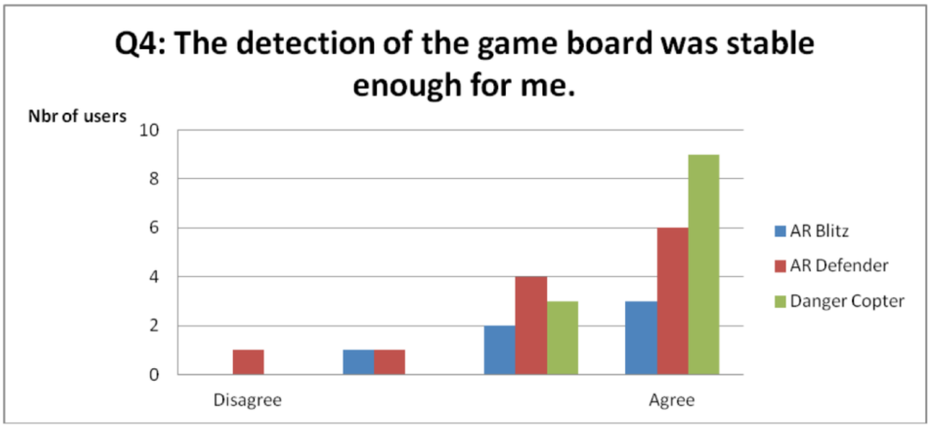


FIGURE 7.4 RESULTS FOR Q4: THE DETECTION OF THE GAME BOARD WAS STABLE ENOUGH FOR ME

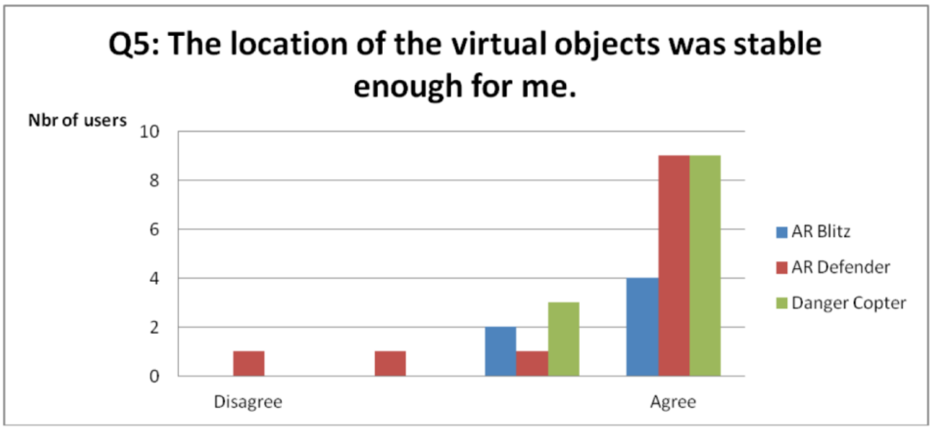


FIGURE 7.5 RESULTS FOR Q5: THE LOCATION OF THE VIRTUAL OBJECTS WAS STABLE ENOUGH FOR ME

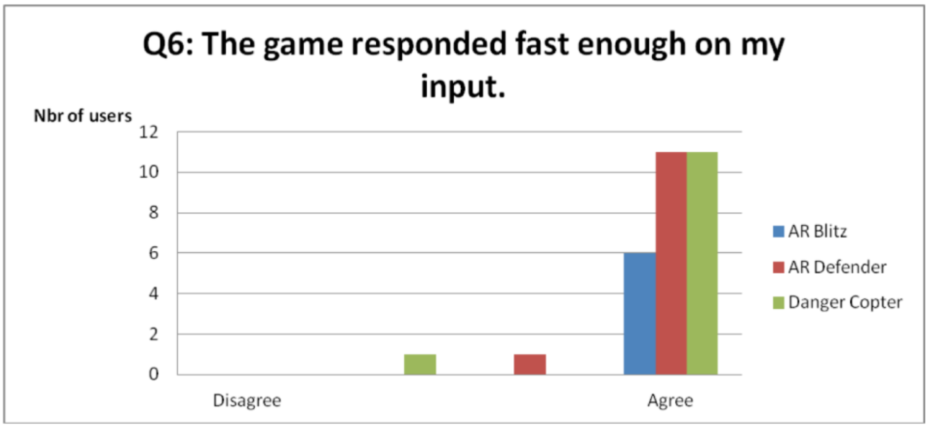


FIGURE 7.6 RESULTS FOR Q6: THE GAME RESPONDED FAST ENOUGH ON MY INPUT

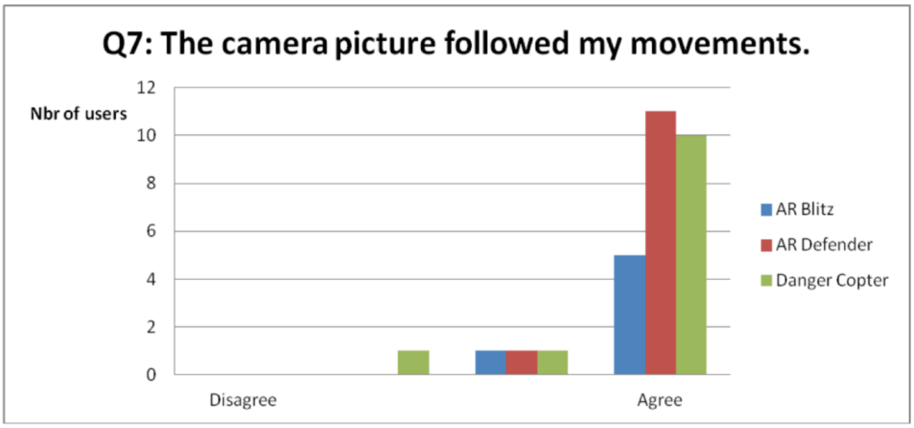


FIGURE 7.7 RESULTS FOR Q7: THE CAMERA PICTURE FOLLOWED MY MOVEMENTS

7.6 Answers from questionnaires phone form factor

Below tables are presenting the results for each game that participants answered for the phone form factor.

TABLE 7.1 SHOWS PARTICIPANTS ANSWERS FOR AR BLITZ

AR Blitz Phone				
	Disagree			Agree
I easily understood how to play the game.	0	0	0	6
I became physically tired of playing the game.	5	1	0	0
Technical instabilities irritated me during the game.	2	0	2	2
The detection of the game board was stable enough for me.	0	1	2	3
The location of the virtual objects was stable enough for me.	0	0	2	4
The game responded fast enough on my input.	0	0	0	6
The camera picture followed my movements.	0	0	1	5

TABLE 7.2 SHOWS PARTICIPANTS ANSWERS FOR AR DEFENDER

AR Defender Phone				
	Disagree			Agree
I easily understood how to play the game.	0	0	1	11
I became physically tired of playing the game.	9	0	3	0
Technical instabilities irritated me during the game.	7	2	2	1
The detection of the game board was stable enough for me.	1	1	4	6
The location of the virtual objects was stable enough for me.	1	1	1	9
The game responded fast enough on my input.	0	0	1	11
The camera picture followed my movements.	0	0	1	11

TABLE 7.3 SHOWS PARTICIPANTS ANSWERS FOR DANGER COPTER

Danger copter Phone				
	Disagree			Agree
I easily understood how to play the game.	0	0	1	11
I became physically tired of playing the game.	6	5	1	0
Technical instabilities irritated me during the game.	9	0	3	0
The detection of the game board was stable enough for me.	0	0	3	9
The location of the virtual objects was stable enough for me.	0	0	3	9
The game responded fast enough on my input.	0	1	0	11
The camera picture followed my movements.	0	1	1	10

7.7 Graphs from questionnaires tablet form factor

Below graphs are presenting the results for each question that participants answered for the tablet form factor.

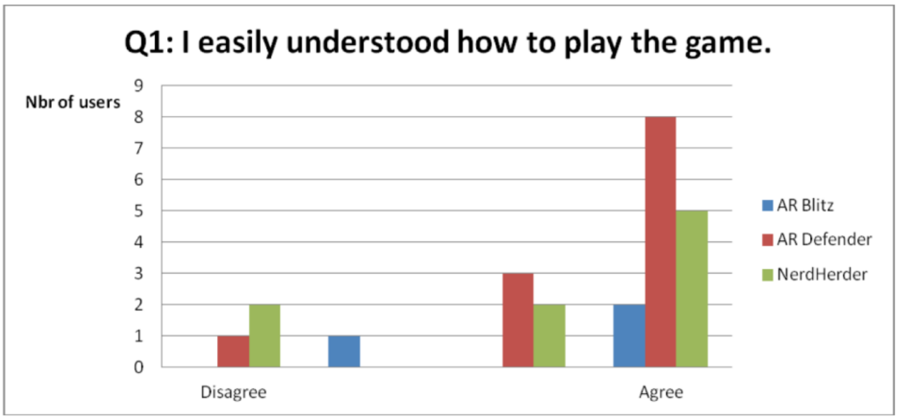


FIGURE 7.8 RESULTS FOR Q1: I EASILY UNDERSTOOD HOW TO PLAY THE GAME

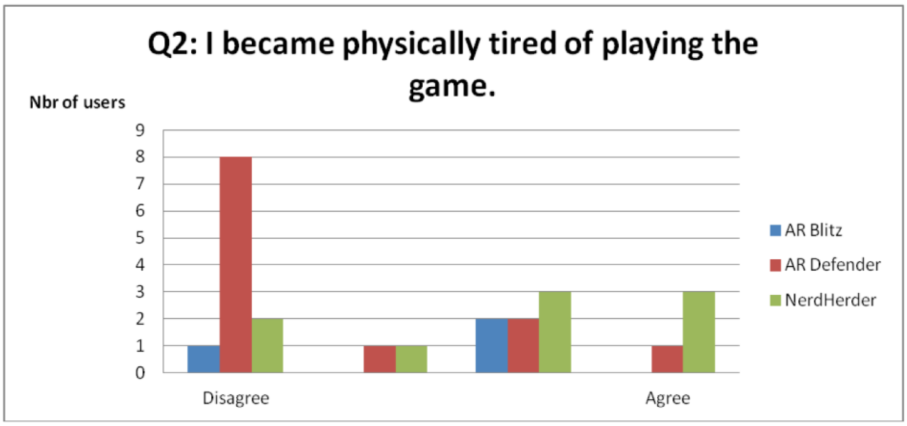


FIGURE 7.9 RESULTS FOR Q2: I BECAME PHYSICALLY TIRED OF PLAYING THE GAME

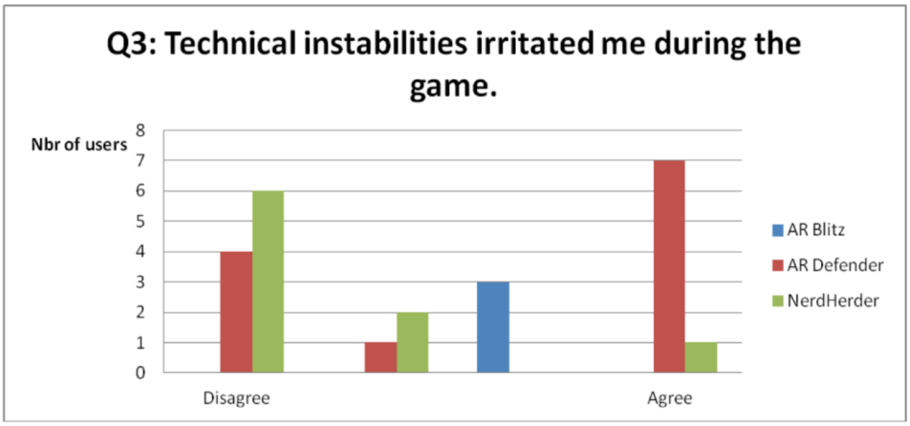


FIGURE 7.10 RESULTS FOR Q3: TECHNICAL INSTABILITIES IRRITATED ME DURING THE GAME

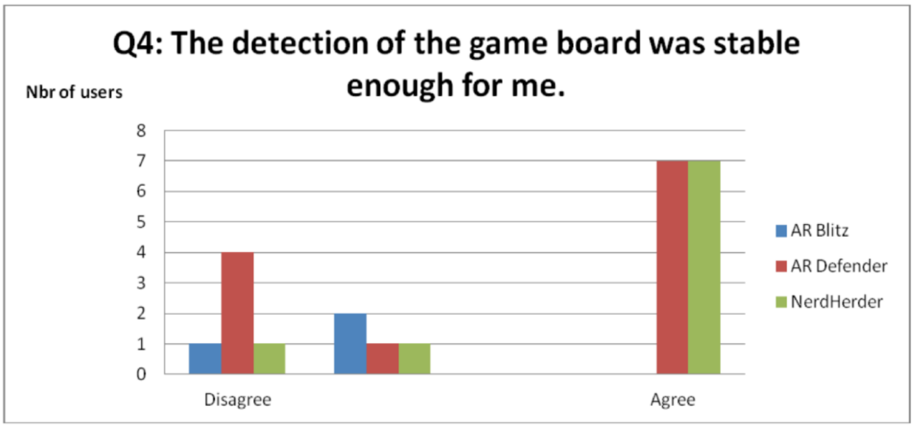


FIGURE 7.11 RESULTS FOR Q4: THE DETECTION OF THE GAME BOARD WAS STABLE ENOUGH FOR ME

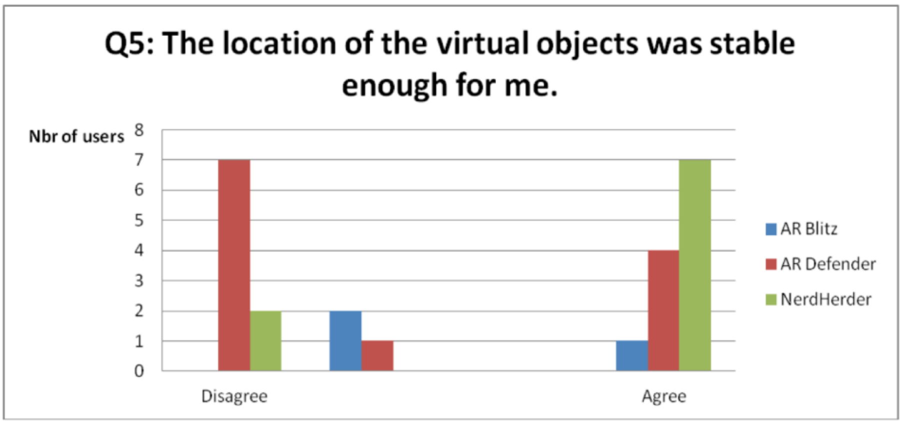


FIGURE 7.12 RESULTS FOR Q5: THE LOCATION OF THE VIRTUAL OBJECTS WAS STABLE ENOUGH FOR ME

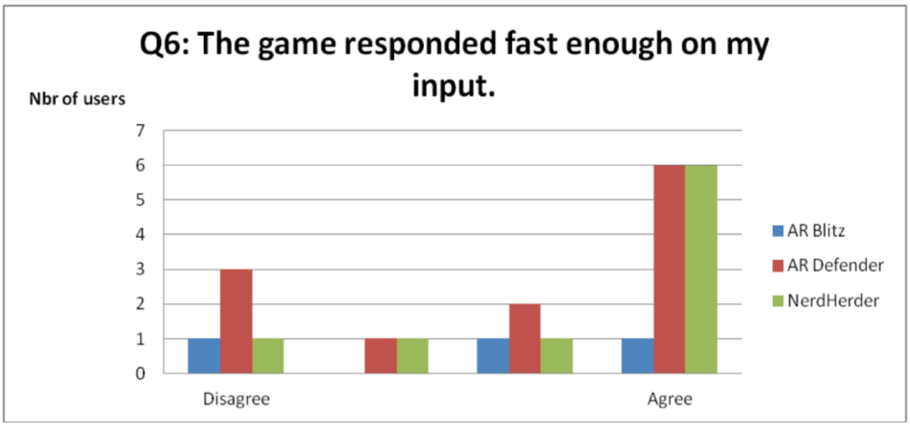


FIGURE 7.13 RESULTS FOR Q6: THE GAME RESPONDED FAST ENOUGH ON MY INPUT

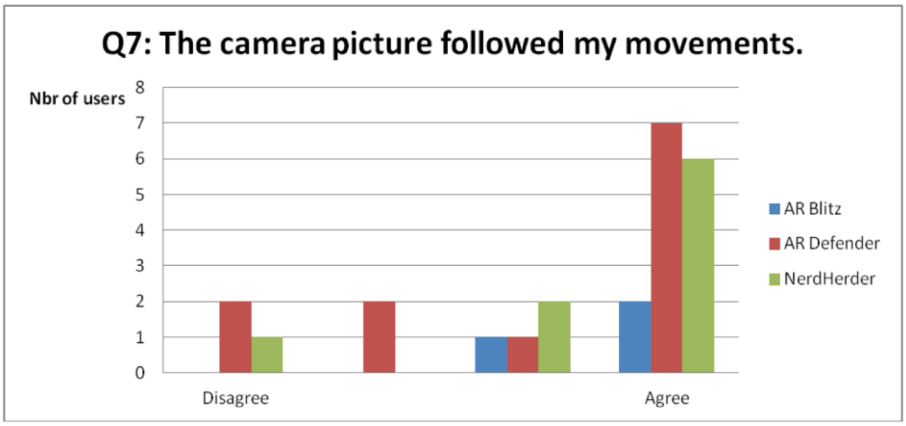


FIGURE 7.14 RESULTS FOR Q7: THE CAMERA PICTURE FOLLOWED MY MOVEMENTS

7.8 Answers from questionnaires tablet form factor

Below tables are presenting the results for each game that participants answered for the phone form factor.

TABLE 7.4 SHOWS PARTICIPANTS ANSWERS FOR AR BLITZ TABLET

AR Blitz Tablet				
	Disagree			Agree
I easily understood how to play the game.	0	1	0	2
I became physically tired of playing the game.	1	0	2	0
Technical instabilities irritated me during the game.	0	0	3	0
The detection of the game board was stable enough for me.	1	2	0	0
The location of the virtual objects was stable enough for me.	0	2	0	1
The game responded fast enough on my input.	1	0	1	1
The camera picture followed my movements.	0	0	1	2

TABLE 7.5 SHOWS PARTICIPANTS ANSWERS FOR AR DEFENDER TABLET

AR Defender Tablet				
	Disagree			Agree
I easily understood how to play the game.	1	0	3	8
I became physically tired of playing the game.	8	1	2	1
Technical instabilities irritated me during the game.	4	1	0	7
The detection of the game board was stable enough for me.	4	1	0	7
The location of the virtual objects was stable enough for me.	7	1	0	4
The game responded fast enough on my input.	3	1	2	6
The camera picture followed my movements.	2	2	1	7

TABLE 7.6 SHOWS PARTICIPANTS ANSWERS FOR NERDHERDER TABLET

NerdHerder Tablet				
	Disagree			Agree
I easily understood how to play the game.	2	0	2	5
I became physically tired of playing the game.	2	1	3	3
Technical instabilities irritated me during the game.	6	2	0	1
The detection of the game board was stable enough for me.	1	1	0	7
The location of the virtual objects was stable enough for me.	2	0	0	7
The game responded fast enough on my input.	1	1	1	6
The camera picture followed my movements.	1	0	2	6

7.9 Graphs from NASA-TLX phone form factor

Below graphs are presenting the results for each question that participants answered for the phone form factor, with the NASA-TLX.

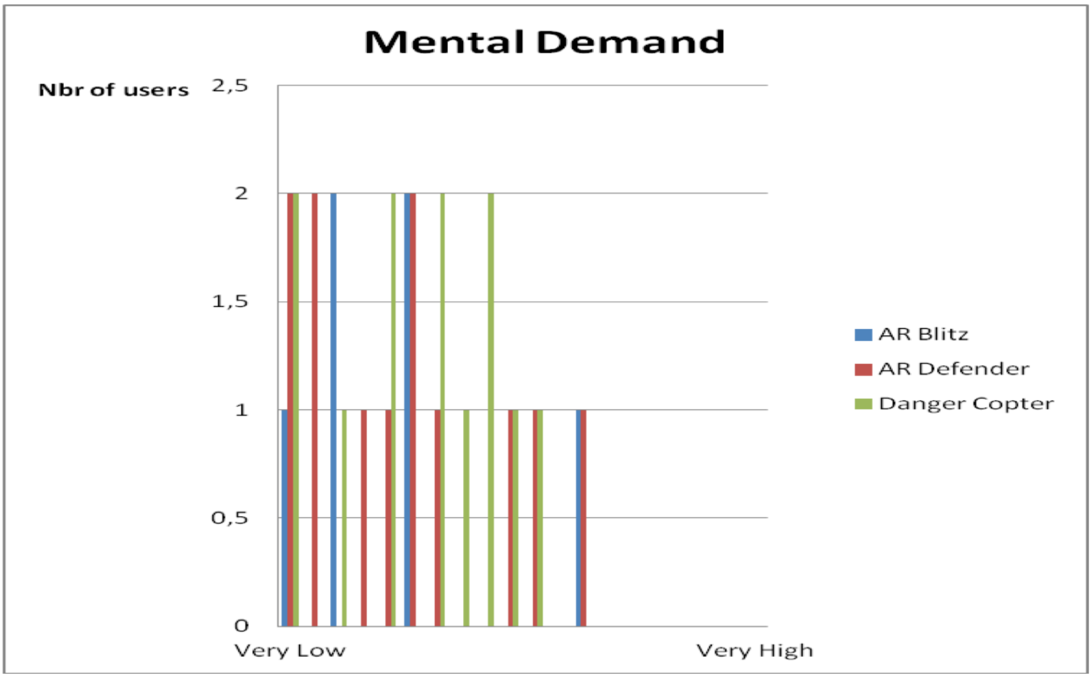


FIGURE 7.15 RESULTS FOR MENTAL DEMAND

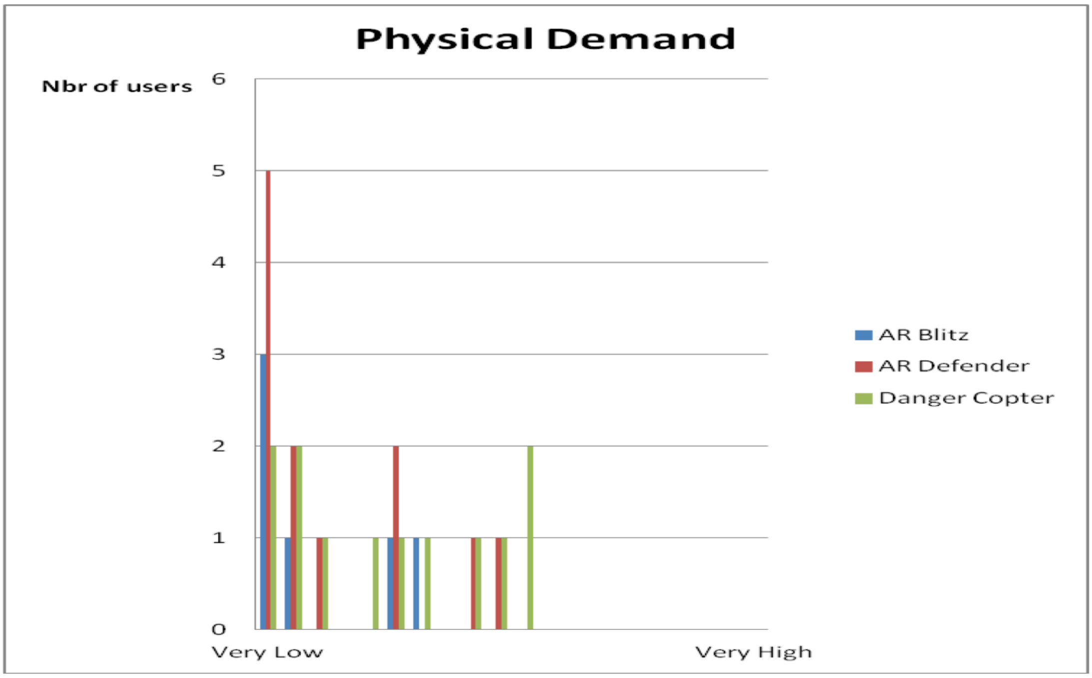


FIGURE 7.16 RESULTS FOR PHYSICAL DEMAND

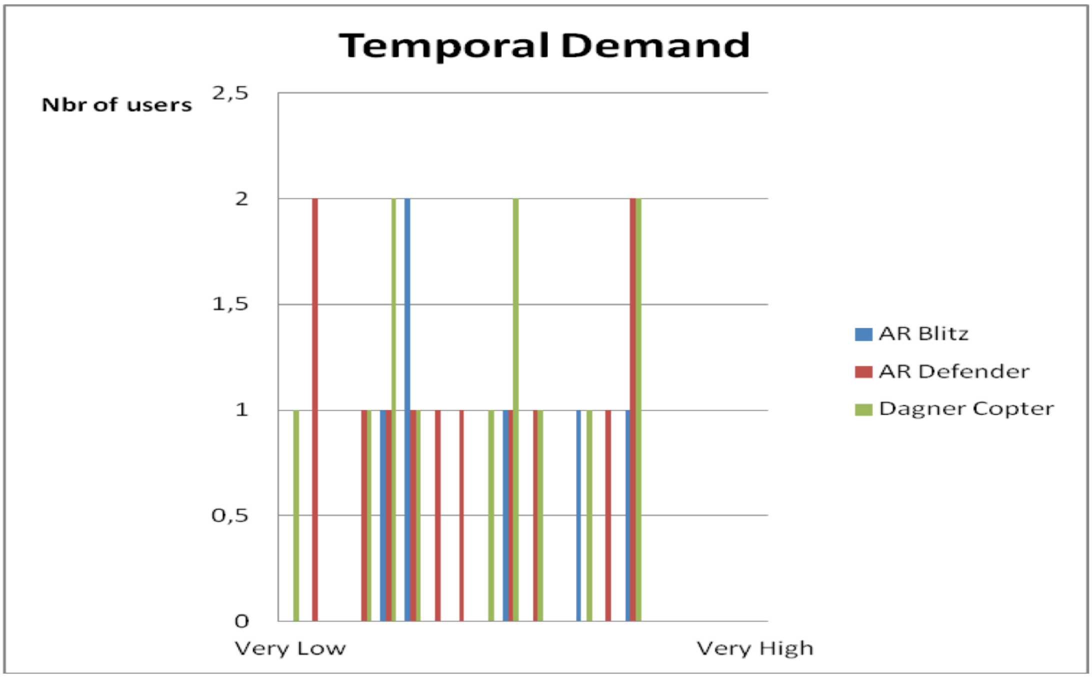


FIGURE 7.17 RESULTS FOR TEMPORAL DEMAND

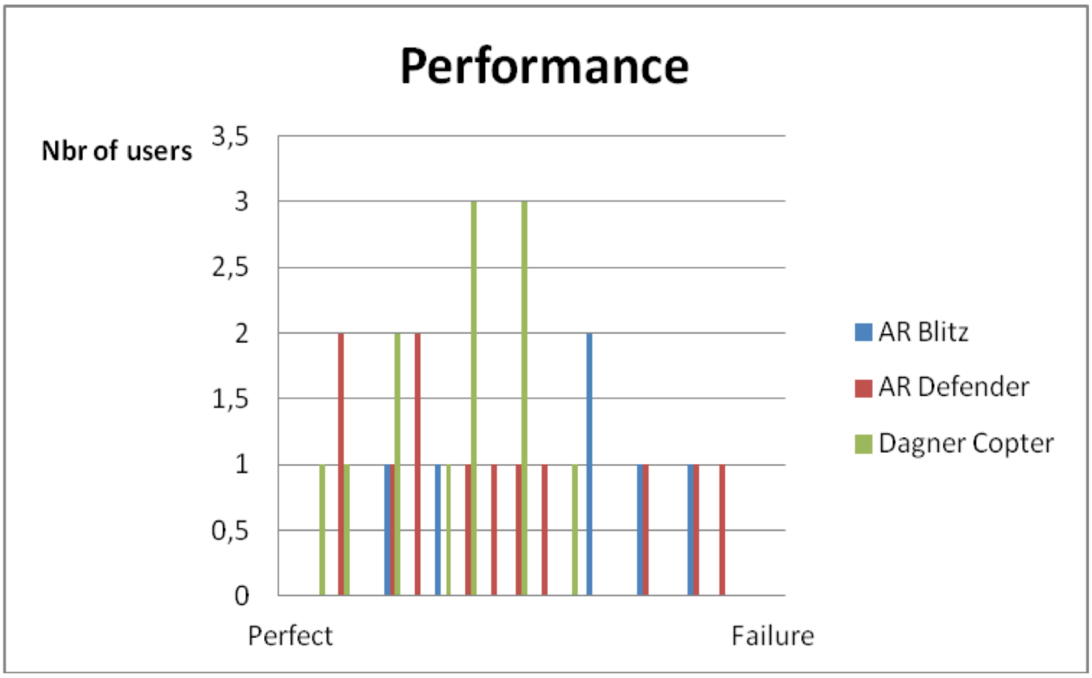


FIGURE 7.18 RESULTS FOR PERFORMANCE

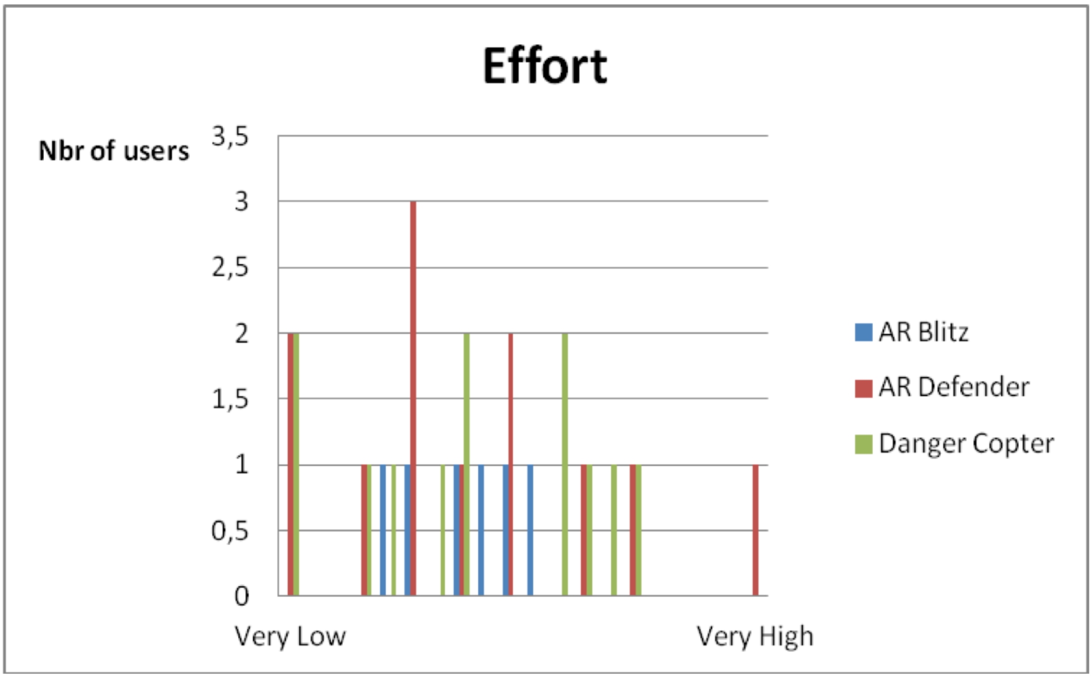


FIGURE 7.19 RESULTS FOR EFFORT

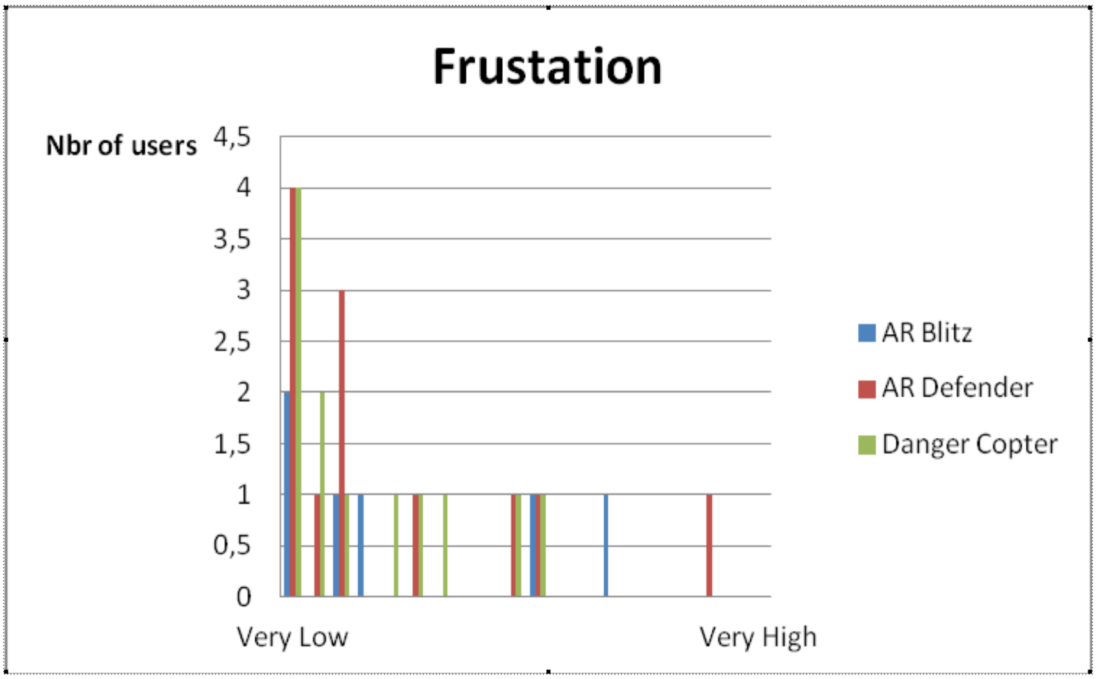


FIGURE 7.20 RESULTS FOR FRUSTRATION

7.10 Answers from NASA-TLX phone form factor

Below tables are presenting the results for each game that participants answered for the phone form factor.

TABLE 7.7 SHOWS PARTICIPANTS ANSWERS FOR AR BLITZ

AR Blitz Phone																				
	VL	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	VH
Mental Demand (How mentally demanding was the task?)	1	0	2	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Physical Demand (How physically demanding was the task?)	3	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Temporal Demand (How hurried or rushed was the pace of the task?)	0	0	0	0	1	2	0	0	0	1	0	0	1	0	1	0	0	0	0	0
Performance (How successful were you in accomplishing what your were asked to do?)	0	0	0	0	1	0	1	0	0	0	0	0	2	0	1	0	1	0	0	0
Effort (How hard did you have to work to accomplish your level of performance?)	0	0	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0
Frustration (How insecure, discouraged, irritated, stressed, and annoyed were you?)	2	0	1	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0

TABLE 7.8 SHOWS PARTICIPANTS ANSWERS FOR AR DEFENDER

AR Defender Phone																				
	VL	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	VH
Mental Demand (How mentally demanding was the task?)	2	2	0	1	1	2	1	0	0	1	1	0	1	0	0	0	0	0	0	0
Physical Demand (How physically demanding was the task?)	5	2	1	0	0	2	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Temporal Demand (How hurried or rushed was the pace of the task?)	0	2	0	1	1	1	1	1	0	1	1	0	0	1	2	0	0	0	0	0
Performance (How successful were you in accomplishing what your were asked to do?)	0	0	2	0	1	2	0	1	1	1	1	0	0	0	1	0	1	1	0	0
Effort (How hard did you have to work to accomplish your level of performance?)	2	0	0	1	0	3	0	1	0	2	0	0	1	0	1	0	0	0	0	1
Frustration (How insecure, discouraged, irritated, stressed, and annoyed were you?)	4	1	3	0	0	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0

TABLE 7.9 SHOWS PARTICIPANTS ANSWERS FOR DANGER COPTER

Danger copter Phone	VL	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	VH
Mental Demand (How mentally demanding was the task?)	2	0	1	0	2	0	2	1	2	1	1	0	0	0	0	0	0	0	0	0
Physical Demand (How physically demanding was the task?)	2	2	1	0	1	1	1	0	1	1	2	0	0	0	0	0	0	0	0	0
Temporal Demand (How hurried or rushed was the pace of the task?)	1	0	0	1	2	1	0	0	1	2	1	0	1	0	2	0	0	0	0	0
Performance (How successful were you in accomplishing what your were asked to do?)	0	1	1	0	2	0	1	3	0	3	0	1	0	0	0	0	0	0	0	0
Effort (How hard did you have to work to accomplish your level of performance?)	2	0	0	1	1	0	1	2	0	0	0	2	1	1	1	0	0	0	0	0
Frustration (How insecure, discouraged, irritated, stressed, and annoyed were you?)	4	2	1	0	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0

7.11 Graphs from NASA-TLX tablet form factor

Below graphs are presenting the results for each question that participants answered for the tablet form factor, with the NASA-TLX.

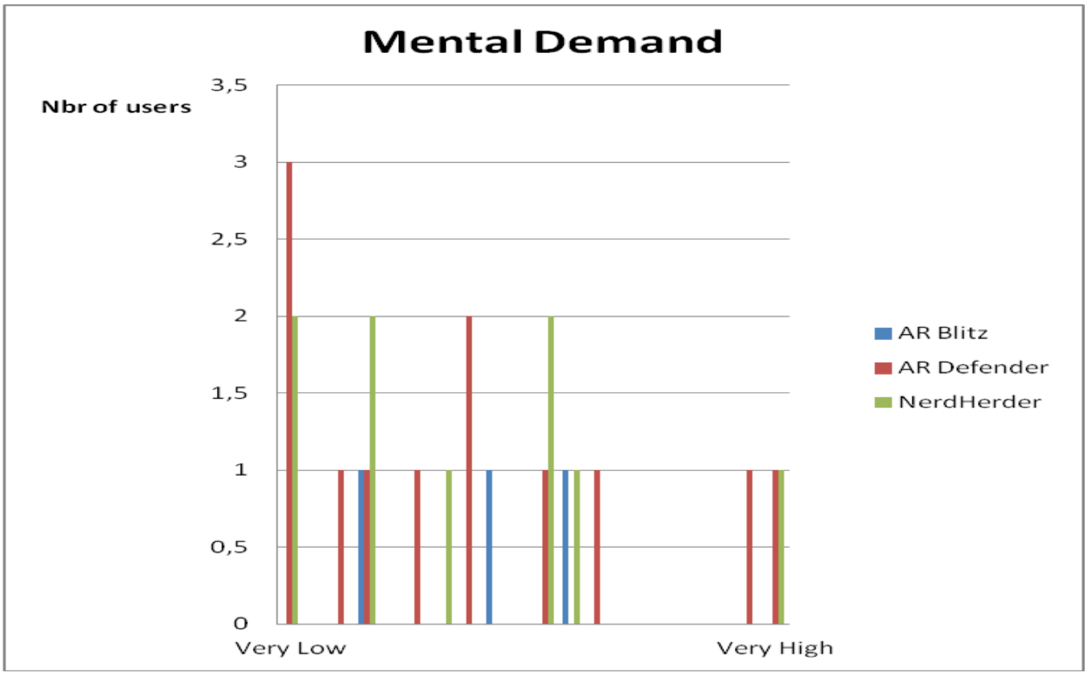


FIGURE 7.21 RESULTS FOR MENTAL DEMAND

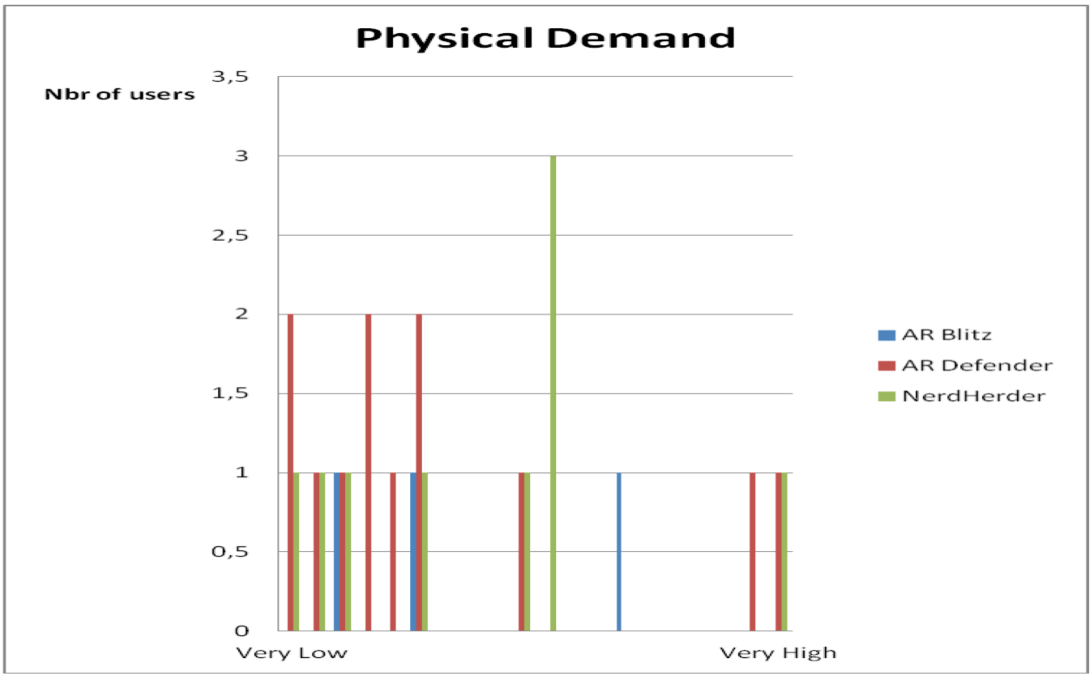


FIGURE 7.22 RESULTS FOR PHYSICAL DEMAND

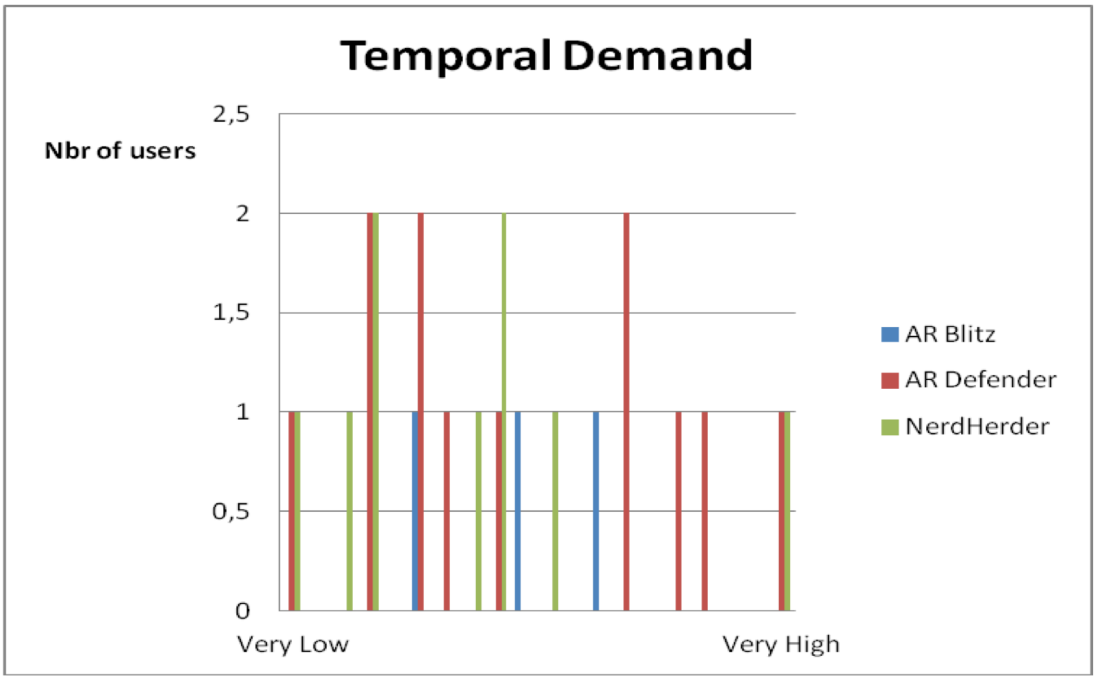


FIGURE 7.23 RESULTS FOR TEMPORAL DEMAND

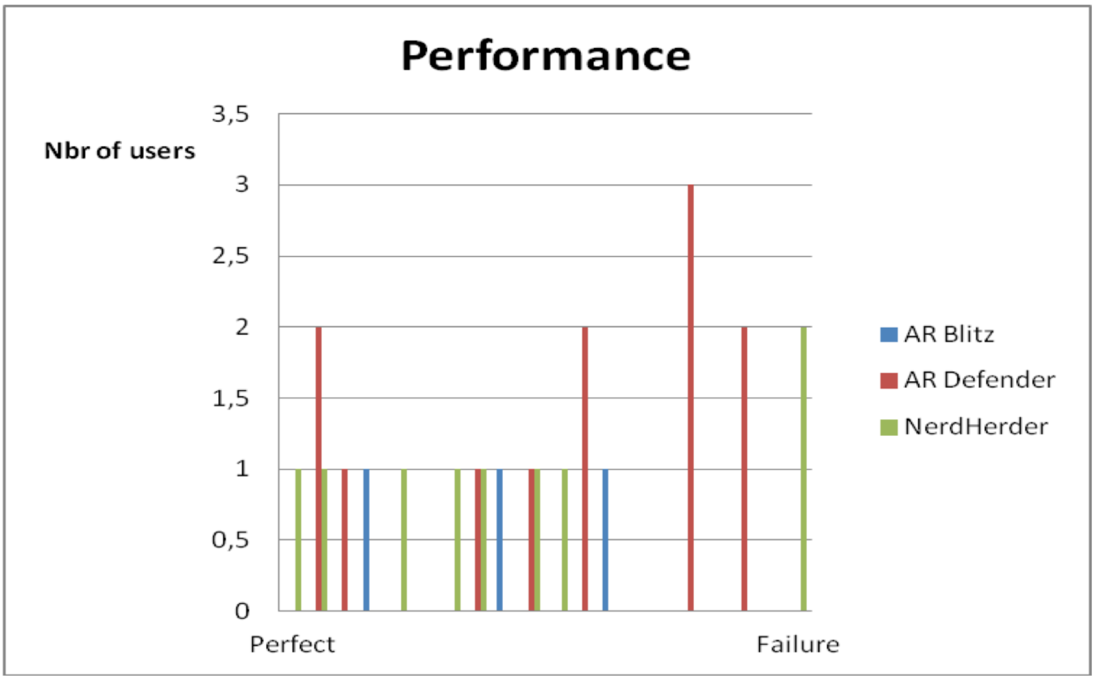


FIGURE 7.24 RESULTS FOR PERFORMANCE

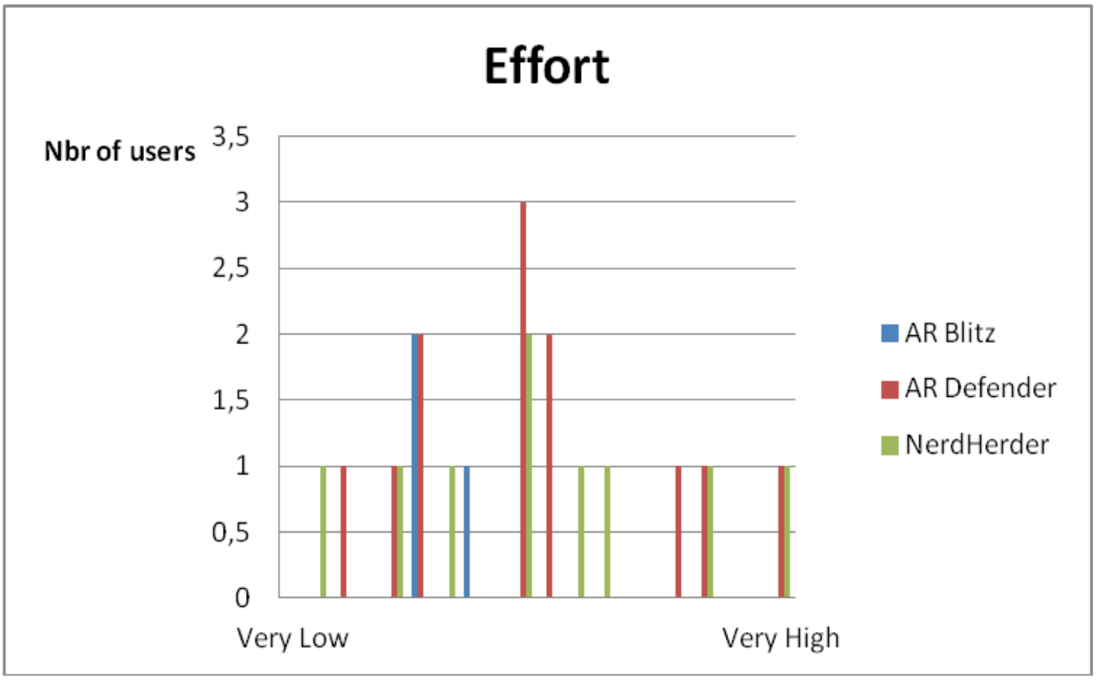


FIGURE 7.25 RESULTS FOR EFFORT

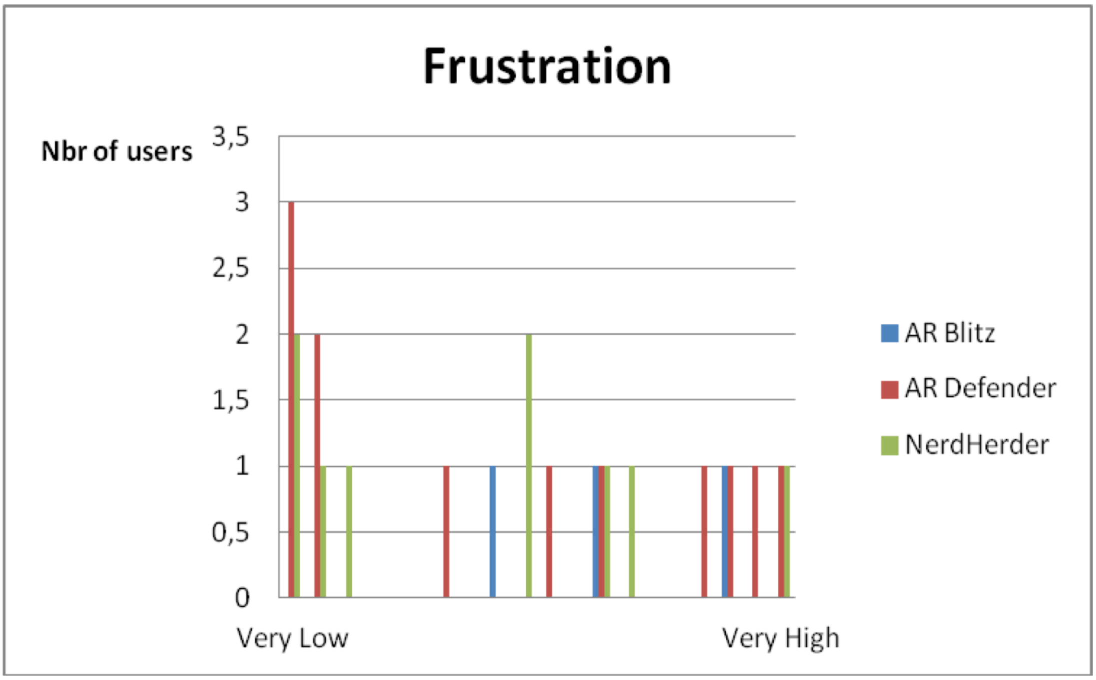


FIGURE 7.26 RESULTS FOR FRUSTRATION

7.12 Answers from NASA-TLX tablet form factor

Below tables are presenting the results for each game that participants answered for the tablet form factor.

TABLE 7.10 SHOWS PARTICIPANTS ANSWERS FOR AR BLITZ TABLET

AR Blitz Tablet	VL	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	VH
Mental Demand (How mentally demanding was the task?)	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
Physical Demand (How physically demanding was the task?)	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Temporal Demand (How hurried or rushed was the pace of the task?)	0	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0
Performance (How successful were you in accomplishing what your were asked to do?)	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
Effort (How hard did you have to work to accomplish your level of performance?)	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Frustration (How insecure, discouraged, irritated, stressed, and annoyed were you?)	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0

TABLE 7.11 SHOWS PARTICIPANTS ANSWERS FOR AR DEFENDER TABLET

AR Defender Tablet	VL	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	VH
Mental Demand (How mentally demanding was the task?)	3	0	1	1	0	1	0	2	0	0	1	0	1	0	0	0	0	0	1	1
Physical Demand (How physically demanding was the task?)	2	1	1	2	1	2	0	0	0	1	0	0	0	0	0	0	0	0	1	1
Temporal Demand (How hurried or rushed was the pace of the task?)	1	0	0	2	0	2	1	0	1	0	0	0	0	2	0	1	1	0	0	1
Performance (How successful were you in accomplishing what your were asked to do?)	0	2	1	0	0	0	0	1	0	1	0	2	0	0	0	3	0	2	0	0
Effort (How hard did you have to work to accomplish your level of performance?)	0	0	1	0	1	2	0	0	0	3	2	0	0	0	0	1	1	0	0	1
Frustration (How insecure, discouraged, irritated, stressed, and annoyed were you?)	3	2	0	0	0	0	1	0	0	0	1	0	1	0	0	0	1	1	1	1

TABLE 7.12 SHOWS PARTICIPANTS ANSWERS FOR NERDHERDER TABLET

NerdHerder Tablet																				
	VL	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	VH
Mental Demand (How mentally demanding was the task?)	2	0	0	2	0	0	1	0	0	0	2	1	0	0	0	0	0	0	0	1
Physical Demand (How physically demanding was the task?)	1	1	1	0	0	1	0	0	0	1	3	0	0	0	0	0	0	0	0	1
Temporal Demand (How hurried or rushed was the pace of the task?)	1	0	1	2	0	0	0	1	2	0	1	0	0	0	0	0	0	0	0	1
Performance (How successful were you in accomplishing what your were asked to do?)	1	1	0	0	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	2
Effort (How hard did you have to work to accomplish your level of performance?)	0	1	0	0	1	0	1	0	0	2	0	1	1	0	0	0	1	0	0	1
Frustration (How insecure, dis-couraged, irritated, stressed, and annoyed were you?)	2	1	1	0	0	0	0	0	0	2	0	0	1	1	0	0	0	0	0	1