Auditory Icons vs Earcons in Games

*What Makes the Most Efficient Non-verbal Sound Cue?*

Gabriel Johannesson
2016

Bachelor of Arts
Audio Engineering

Luleå University of Technology
Department of Arts, Communication and Education
Auditory Icons vs Earcons in Games: What Makes the Most Efficient Non-verbal Sound Cue?

By
Gabriel Johannesson

Luleå University of Technology 2016
Abstract
Non-verbal sound cues can be used to convey different kinds of information and are often used for both simple and more complex types of feedback in computer games. Non-verbal sound cues can be divided into two different categories; auditory icons, that are sounds that represent real world events and earcons, that are abstract synthetic or musical sounds. In order to see if there was a difference between the player response times yielded by auditory icons and earcons in a video game setting, an experiment game was created. A game where the response times of subjects were recorded when reacting to auditory icons and earcons, while playing through four different treasure hunt like game scenarios. The results from the experiment seem to indicate that there was no significant difference between the combined response times produced by the auditory icons compared to the combined response times produced by the earcons.
# Table of Contents

Abstract ................................................................................................................................................. 2  
Introduction ............................................................................................................................................... 5  
Background ................................................................................................................................................ 5  
Gameplay and interaction ....................................................................................................................... 6  
Non-verbal sound cues ............................................................................................................................. 7  
Player efficiency ......................................................................................................................................... 8  
Sound cue efficiency ................................................................................................................................. 8  
Sounds conveying information .................................................................................................................. 9  
Psychoacoustics ........................................................................................................................................ 9  
Navigation ................................................................................................................................................ 11  
Within-vehicle auditory display .............................................................................................................. 11  
What does this mean for sound in games? ............................................................................................... 12  
Method ...................................................................................................................................................... 13  
Experiment game creation ....................................................................................................................... 13  
Stimuli ...................................................................................................................................................... 15  
Subjects ................................................................................................................................................... 16  
Test set-up ................................................................................................................................................. 17  
Results ...................................................................................................................................................... 18  
Analysis .................................................................................................................................................... 21  
Experiment game ...................................................................................................................................... 21  
Demographic information ......................................................................................................................... 23  
1. Place of origin ..................................................................................................................................... 23  
2. Game time per week ............................................................................................................................ 24  
3. Game genres ....................................................................................................................................... 25  
4. Gaming platform ............................................................................................................................... 25  
5. Previous musical experience ............................................................................................................. 26  
6. Duration of musical experience ........................................................................................................ 27  
7. Personal electronic devices (sound/no sound) ................................................................................ 27  
8. Thoughts on notification type sounds ............................................................................................ 28  
Discussion ................................................................................................................................................ 30  
Experiment findings ............................................................................................................................... 30  
Future game sound design and development ......................................................................................... 31  
Experimental design .............................................................................................................................. 32  
Conclusion ............................................................................................................................................... 33  
Further research ...................................................................................................................................... 33
References ........................................................................................................................................... 34
Appendix A ........................................................................................................................................ 35
  Experiment game screenshots ......................................................................................................... 35
Appendix B ........................................................................................................................................ 38
  Questionnaire raw data .................................................................................................................. 38
    Question 8 full responses .............................................................................................................. 41
Introduction
Sound is an important part of modern video games that contributes to the whole gaming experience, from making the game world more believable to being an essential part of the gameplay mechanics. A good sound design can be important for sustaining immersion or providing the player with necessary information and feedback. One way of using sound to give the player information is by using non-verbal sound cues. In other words, sound cues that aren’t dialog or voiceover. It is these non-verbal sound cues that will be the focus of this paper. There are two different categories that you can divide non-verbal sound cues into; auditory icons, that are sounds that represent real world events (Brazil and Fernström, 2011) and earcons, that are abstract synthetic or musical sounds (McGookin and Brewster, 2011).

The purpose of this essay is to see if there is a difference between the player response times yielded by auditory icons and earcons in a video game setting. This, based on the results of a study by McKeown and Isherwood (2007) that showed auditory icons as more efficient, yielding both faster response times and higher accuracy ratings, when mapping different audio stimuli to their referent driving events in an auditory display for use in vehicles.

For this purpose, an experiment was conducted where, the response times of players were recorded, when reacting to auditory icons and earcons in different treasure hunt like game scenarios. Before the experiment is explained, the area of non-verbal sound cues in games will be elaborated upon and examples of non-verbal sound cues conveying information will be presented. The concepts of player efficiency and sound cue efficiency will be presented and explained as well.

Background
Non-verbal sound cues can be very important for gameplay. They can serve as a confirmation that the player has performed some kind of action, for example, an objective has been completed. The footsteps of an enemy in a competitive first person shooter can give the player information about the enemy’s whereabouts. The second example with the enemy footsteps is an example of a non-verbal sound cue that directly affects the player’s decision making and actions. If, for example, the player isn’t able to hear the enemy’s footsteps clearly or if they aren’t distinctive as enemy footsteps in particular, the player won’t be able to track and predict the enemy’s movements correctly or to perform appropriate actions as a response to the footsteps. In this way, what information the sound provides to the player and how it delivers this information is connected to the player’s actions, in particularly the efficiency of those actions for successful gameplay. The information the non-verbal sound cue gives the player helps him or her, to make a decision and to perform an appropriate action as a response to the information. The non-verbal sound cue in this example scenario enables the player to be more efficient and to perform on a higher overall level. A player’s performance is based on, for example, his or her perception, skill, strategies and responses, things that affect what he or she does. E.g. if a football player makes a play, the play will be influenced by these things. This play is then representative of the player’s performance, if it was a successful play, he or she performed well in that particular situation.

The idea that non-verbal sound cues can be used to convey information and the capabilities and/or limitations of our auditory system are important things to take into consideration when doing sound design of any kind, especially video game sound design. In our everyday life we often rely on sound to make decisions and perform actions to, for example, avoid a dangerous situation. Games can be considered to be virtual worlds, worlds where we also must rely on sounds to make decisions and perform these kinds of actions. How these non-verbal sound cues and the information they can convey then affect a player’s efficiency is an area that is relevant for further research if we want to intentionally control the challenges presented to the player, encourage certain actions or for the player to develop a particular set of skills in order to enhance the player’s ability to respond efficiently.
Gameplay and interaction
There’s a lot of different types of games that people enjoy and choose to spend their time playing; all with different types of interaction between the player and the game, requiring different sets of skill. Everything from games with minimal input and minimal eye-hand coordination, like simple “point and click” games, to for example Massively Multiplayer Online Role Playing Games (MMORPG) or competitive First Person Shooter (FPS) games. Games that requires the player to do a lot of things at the same time, including complex strategic thinking at the same time as moving and attacking while tracking multiple moving targets. This paper will not try to categorize these interactions, but instead consider the connection between interaction and non-verbal sound cues. Therefore, only some types of interactions will be mentioned and considered here, but they will be considered in depth.

Interaction is essential to any type of game (Salen and Zimmerman 2004). To sustain interaction, the game needs to give the player feedback. The player needs to know they’ve had an impact on the game. A common sort of feedback is to simply confirm that the interaction/input they made was registered by the game by, for example, a text message being displayed or a sound being played. Otherwise the player might not trust that the interaction or input they made to the game was accepted or even happened at all and proceed to repeat the input multiple times, which can be very frustrating and often breaks immersion.

Salen and Zimmerman devised a model of interactivity with four modes of interactivity, or as the authors describe it “four different levels of engagement, that a person might have with an interactive system” (Salen and Zimmerman, 2004, p. 55).

Salen and Zimmerman (2004, p. 55-56) describe these four modes as the following:

“Mode 1: Cognitive interaction; or interpretive participation”

“Mode 2: Functional interactivity; or utilitarian participation”

“Mode 3: Explicit interactivity; or participation with designed choices and procedures”

“Mode 4: Beyond-the-object-interactivity; or participation within the culture of the object”

When it comes to sound in video games, modes 2 and 3 are where sound design can have the largest impact on a player’s experience. The sound design will also have an impact on modes 1 and 4, but since these modes are mostly based on the player’s imagination and own thoughts, what is done with sound design in modes 2 and 3 will impact modes 1 and 4.

Mode 2: Functional interactivity can be described as if sounds are being played at the right times/situations and if they are audible or if they are being masked by other less important sounds. Also the overall sound quality of all the sounds. For example, if the game’s visual rendering and game mechanics are working as intended, but all sounds are being played out of sync, the game won’t feel very functional or responsive to the player. On the other hand, if all the sounds in the game are functioning as intended, it will likely make the game feel even more responsive to the player’s inputs and functional than without any sounds at all. This, because of all the feedback it gives to the player.
Mode 3: Explicit interactivity can be described as when the player is actually playing the game and reacting to all of the sounds in the game. At this point, it is not only a matter of whether a sound is playing at the correct time or not, but also if the sound being played enhances the interaction, if it’s representative of the visual aspect it might represent, if it’s believable in the games world etc. Examples of all this might be: players reacting and adapting their gameplay to enemy footsteps in a competitive FPS game or players being able to solve a difficult puzzle or maze with the help of a well-designed sound cue. This is also where a believable sound design can help keep the player immersed in the game world during the gameplay.

As we can see from this, the feedback and information that sounds can give to the player is tied in to the interaction between the player and the game. Sound design can facilitate and control this interaction, which in turn makes sound an important part of the gaming experience.

Non-verbal sound cues

As mentioned above, a common sort of feedback a game can give the player is to simply confirm that an input the player made was registered by the game. It is common to use non-verbal sound cues for this kind of confirmation, for example, with a simple sound like a beep, boop or ding. Using a non-verbal sound cue as a confirmation in these kinds of scenarios is often preferable compared to using dialog as a confirmation saying “nice”, “confirmed” or “ok”. The reason for this being, mainly, “localization”, i.e. translating a game into multiple different languages including dialog and menu text (Lundqvist, J., Game Sound Lecture, 09/10-2015). It is also the fact that if a piece of dialog saying “nice” is played every time the player, for example, picks up an item; it can quite quickly become very repetitive, which in return can break immersion. Hearing this “nice” over and over is fatiguing and it might also get confused with narrative dialog or other sound effects.

Non-verbal sound cues are used for far more complex types of feedback as well, for example, something more advanced that is tied in to the gameplay; e.g. in Counter Strike: Global Offensive (Valve, 2012), a team based competitive first person shooter, if the player shoots an opponent in the head, if that opponent has head armor, a specific sound will be played. If the enemy doesn’t die, it means that they are low on health and is an “easy kill” which might change the player’s approach when continuing to fight the enemy. It also gives the player the information that the enemy has bought head armor, this is important in higher level play were the economical aspect of the game is important and you want to keep track of your enemies’ economic situation from a strategical stand point. Counter Strike: Global Offensive (Valve, 2012) has an economical system wherein players must buy their equipment at the start of each round and are rewarded different amounts of in-game money based on, for example, kills and winning or losing a round. The sound cue in this scenario helps the player create an efficient strategy. The fact that these non-verbal sound cues can be very simple/contain little information and complex/contain a lot of information depending on how you use them, make them a very versatile tool for sound designers and game designers alike. It all comes down to what sound you use, what situation you use it in and if you treat is as a simple sound effect or if you actually build game mechanics around the sound; which, if done right, can make room for really interesting gameplay.

Non-verbal sound cues can be divided into two main categories, auditory icons and earcons. Auditory icons are recorded non-speech sounds that the player can be familiar with from experiences of the real world, for example, the sound of crumpling paper playing when you delete a digital document (Brazil and Fernström, 2011). Earcons are short and structured synthetic or musical sounds (McGookin and Brewster, 2011), for example, a beep when a player presses a button in-game. The main difference between these two categories is that auditory icons can be considered to require almost no learning in the game when it is obvious what they represent. An earcon on the other hand would require the player to learn and associate that sound to a specific event in the game.
**Player efficiency**

In order to look further into how these non-verbal sound cues relate to player efficiency, the concept of player efficiency must first be defined. How a player responds to the information provided by the game, what actions he or she takes, can be considered to be more or less efficient; if the player is efficient or not, or in other words player efficiency. Efficient meaning that the player uses a minimum amount of time, effort and focus to successfully achieve a desired goal. In this paper the meaning of player efficiency is related to the conscious competence model that represents the four stages of competence a person goes through when learning something new (Maslow, 1987). This model can be applied for video games, because when playing a game for the first time you have to learn the mechanics and the rules of the game. It is a learning process where you develop a set of skills, hence the conscious competence model can be applied.

The four stages of competence, in order, are unconscious incompetence, conscious incompetence, conscious competence and lastly unconscious competence (Maslow, 1987). The stage of unconscious incompetence is where a person isn’t aware of the fact that he or she doesn’t know something or is incompetent at something; the person is unconsciously incompetent (Maslow, 1987). A player in, for example *Counter Strike: Global Offensive* (Valve, 2012), is unconsciously making mistakes, leading to he or she losing matches. The stage of conscious incompetence is where a person is aware that he or she doesn’t know something or needs to improve at something and is willing to improve (Maslow, 1987). A player of *Counter Strike: Global Offensive* is aware of the fact that the reason he or she is losing matches is because the decisions he or she is making in the game do not have the desired outcome most of the time or that he or she is lacking skill. The stage of conscious competence is where a person has required a certain skill or skills and can perform them reliably, but they need to concentrate and actively think about the skill in order to perform it reliably (Maslow, 1987). A player of *Counter Strike: Global Offensive* is making decisions that have the highest or a high chance of a successful outcome, but has to be thinking about the decisions before making them; the player has to be conscious of their actions. The final stage of unconscious competence is where a person has required a certain skill or skills and can perform them reliably. The difference from the conscious competence stage is that the skill is practiced enough for the person to be able to perform it unconsciously. The person doesn’t have to think about their actions, but instead performs them instinctively (Maslow, 1987). A player of *Counter Strike: Global Offensive* is making decisions that have the highest or a high chance of a successful outcome without having to think about them beforehand. He or she makes them unconsciously. The way the conscious competence model relates to player efficiency is that when a player performs an action, if the action is efficient or not depends not only on if it is done quickly or with a small amount of effort, but also if the action is successful in the specific situation. These are all things that can be described by the different stages of the conscious competence model. E.g. if a player performs an action quickly, but the action isn’t successful in the situation and he or she isn’t aware of it, the player can be placed in the unconscious incompetence stage and the action is therefore not efficient. In this paper, the concept of player efficiency is defined as how the player responds to the information provided by the game, related to the conscious competence model.

**Sound cue efficiency**

Non-verbal sound cues themselves can be considered to be efficient or not in the way they facilitate interactions with the player. More precisely the sound qualities of the non-verbal sound cues affect how efficiently they convey information to the player (e.g. are the non-verbal sound cues clearly audible? Are they distinctive and not easily mistaken for another sound?). This concept, in the context of this paper, will be called *sound cue efficiency* and is an important aspect of sound design that can contribute to, for example, a game being easily accessible, intuitive and enjoyable. A game where important sound cues are easily misunderstood or mistaken for environmental sounds, can turn an otherwise enjoyable game in to a frustrating experience.
For a sound cue to be efficient, it needs to convey the desired information (is the sound cue effective or not?) and it needs to be easily understood by the player, it is clearly audible and distinctive. Another thing that is connected to how easily the sound cue is understood by the player is how distracting it is, or how much effort it takes to understand. Does it require a lot of focus to understand, focus that has to be taken from other aspects of the game? For example, does it in Counter Strike: Global Offensive (Valve, 2012) require the player to take focus away from moving, aiming or shooting to understand? In this paper, sound cue efficiency is defined as how efficiently a non-verbal sound cue conveys the desired information to the player in relation to the above stated conditions.

In order to be efficient, the player needs information about what is going on in the game. However, if the player receives too much information, he or she can’t process it in time to be efficient. On the other hand, if the player receives too little information, he or she might make poor decisions or act ineffectively based on the situation. From a high level viewpoint, sound cue efficiency and its connection to player efficiency can be thought of as a balancing act of too much and too little information.

**Sounds conveying information**

In the real world we often rely on sound to give us a lot of information that is necessary for survival in our everyday life. Some of this sound we hear is communication through speech, but other sounds are important as well. Non-verbal sounds of opening/closing doors or approaching cars are examples of this, i.e. sounds not originating from other humans. These kinds of sounds give us vital information about our surroundings and aids us in our everyday life. This concept of non-verbal sounds, specifically non-human sounds, being able to convey information is something very useful in sound design where speech isn’t a suitable option. If there is dialog in a game, a confirmation saying “nice” being played every time the player, for example, picks up an item, this could be confusing or mask other important dialog.

In the following section, examples of non-verbal sound conveying information in different ways will be presented. This in order to show that non-verbal sound is capable of conveying different kinds of information and to be able to further the discussion about sound cue efficiency and player efficiency by connecting them to actual examples of non-verbal sound conveying information.

**Psychoacoustics**

Carlile (2011) describes how advanced the human auditory system is and how small differences between different sounds we are able to perceive. There is a lot of information about a sound’s characteristics that we are able to detect. This is something that affects how we hear everything and is therefore relevant information for all areas of audio engineering. Sound design is a related field of audio engineering where this also applies. If we want to optimize and make sure that the sounds we create for e.g. auditory displays or video games actually have their intended functions, this needs to be taken into consideration. If we can’t hear the sound or perceive the information it tries to convey, we can’t act or respond to that sound the way the sound designer intended. Carlile (2011) considers six topics to be important to the area of psychoacoustics that are things that affect how we perceive sound as a whole. How we perceive sound might also be a factor that affects sound cue efficiency and how it relates to player efficiency.

**The human auditory system**

The anatomy of the human auditory system consists of the outer, middle and inner ear. This is where the entire process of converting acoustic sound waves to biological signals takes place. It goes from acoustic sound waves in the outer ear to mechanical vibrations in the middle ear and finally to biological signals in the inner ear. All of the sounds we hear are affected by the characteristics of this system, for example, some aspects of sounds might be filtered out by this process (Carlile, 2011). In the end, all sounds have to go through this conversion process, including the sounds of a video game. Hence this final stage of conversion is important to consider when doing video game sound design.
Loudness
We perceive loudness in a logarithmic scale known as the decibel scale, abbreviated dBSPL. SPL stands for sound pressure level. The reference level of this scale is the lowest intensity of sound pressure we are able to detect, 20 µP (micro pascal). We also perceive loudness differently at different frequencies, this is called the equal loudness contours or the Fletcher-Munson curves (Carlile, 2011). This is important to take into consideration in video game sound design. E.g. to know about the equal loudness contours and that we are less sensitive to lower frequencies, lets the sound designer know to leave room for more low frequencies to make a big explosion more impactful.

Pitch
Our perception of musical pitch in complex sounds is based on harmonically related frequencies in the complex sound and not which frequency is the loudest. This means that we can perceive the difference of the timbre between two different complex sounds with the same pitch. In relation to this we seem to lose melodic sense for sequences of tone above 5 kHz. We also have the ability to perceive individual frequencies with very high resolution, this is also affected by the specific frequency area. This is called the frequency difference limen (Carlile, 2011). Information about our perception of pitch can be useful when doing video game sound design. E.g. because of our high resolution ability to perceive different frequencies, we can detect very small changes to a sound making it possible to make some variations of a footstep sound from a single source footstep.

Temporal variation
Our perception of temporal variation (rapid variation in spectral content over time) can be broken down into two different components. Carlile (2011) describes these as “a slowly varying envelope and a rapidly varying fine structure” (p. 49). We can, for example, detect gaps in broadband noise no longer than 2-3 ms, this gap increases with lower stimulus levels. Our sensitivity to modulation of a sound’s envelope is the greatest for rates below 50-60 Hz, above which sensitivity falls off and is undetectable for rates above 1 kHz (Carlile, 2011). An example of this is that for a tremolo effect we can’t hear the actual modulation if the rate is too high. This topic is relevant for video game sound design since, for example, because of our ability to detect small gaps in noise, an ambient environmental loop must be looping very precisely. If there is any small gap in the continuous loop, the player will be able to hear it. This in turn can brake immersion.

Grouping
We tend to group single spectral components into auditory groups and streams that we perceive as cohesive auditory events. This is based on different factors, for example, the Gestalt principles. We can also segregate sounds of interest from a complex background of other sounds, related to the same principles (Carlile, 2011). This is something a game sound designer needs to consider when wanting a set of sounds to either be perceived as a single cohesive sound or as multiple separate sounds.

Perception of space
How we perceive different spaces and the source location of sounds is based on the reflections the sound creates in the space, or as Carlile (2011) names them, the “acoustic cues” (p. 53) and also head related transfer functions (Carlile, 2011). This is important in sound design if we, for example, want a sound’s source to feel like it’s placed behind you. It might also be important to know what reverb actually is when working with one, meaning what is early reflections, bass ratio etc.
**Navigation**

Previous research on sound design and player efficiency (in terms of response efficiency) was done in the field of navigation by Dahlstrøm et al. (2013). To define navigation, it can be broken down into the two concepts of “wayfinding and locomotion” (Dahlstrøm et al. 2013, p. 1), where locomotion is the, as Dahlstrøm et al. (2013) describes it, “physical translation through space” (p. 1) or in other words, moving through the environment and wayfinding is the cognitive aspect of navigation (Dahlstrøm et al. 2013). Dahlstrøm et al. (2013) investigated “whether sound design guidelines previously adopted to navigate through a train station can improve the navigation of a maze game” (p. 1). For this purpose, a desktop maze game was designed in which the player had to navigate the maze whilst completing a few tasks in order to complete the game. Using sound design guidelines, similar to that of a train station, an auditory feedback navigation system was created that utilized a continuous beacon sound (a low frequency rumble that the players could navigate to) and a non-continuous confirmation sound (a short high frequency sound which played when the player reached a beacon and a new one started). The beacon and confirmation sounds were accompanied by an ambient soundscape and by player character footsteps.

Dahlstrøm et al. (2013) conducted an experiment where test subjects were divided into two groups, one that was exposed to the maze with the auditory feedback and one that was exposed to the maze without the auditory feedback. The test subjects were not told of the existence or function of the auditory feedback navigation system. The completion time of each test subject was measured and all test subjects were asked a set of questions after they had played the game. The results show that an auditory feedback navigation system, like the one used in this experiment, can be successful in aiding the player of a maze game with significantly faster completion times achieved by the subjects exposed to the auditory feedback navigation system. Worth noting is that some of the test subjects found the chosen sound for the beacon (a low frequency rumble) to be confusing instead of helpful, for example, they thought that the sound was meant to trick them. This study is an example of non-verbal sound cues being used to give a player information about where to go, it helps the player to navigate (in this case) a maze.

This is relevant for game sound designers, since a system like this can be implemented to help players navigate an eventual maze section. They can use sound to help the player navigate a maze.

**Within-vehicle auditory display**

McKeown and Isherwood (2007) investigated speech, environmental sounds, auditory icons and abstract synthetic warnings as “candidates for within-vehicle auditory display interfaces” (p. 1). They did this to examine what types of sounds are the best candidates for within-vehicle displays, in terms of fastest response times and highest accuracy. For this purpose, as pilot studies, driving situations (e.g. low fuel level and car driving in blind spot) and all stimuli, except speech, were placed in a three level urgency scale (from 1 = very urgent warning to 3 = relaxed advisory). A main study was then preformed where test subjects were exposed to a computer task where they matched candidates (speech, auditory icons etc. mentioned above) to referent driving events. Response times and accuracy were measured in the computer task and lastly pleasantness and perceived urgency were assessed separately (McKeown and Isherwood, 2007). The results showed that speech and auditory icons produced the fastest response times and the highest accuracy. Abstract and environmental sounds produced slower response times, but environmental sounds produced accuracy ratings between speech/auditory icons and abstract sounds (McKeown and Isherwood, 2007).

This is an example of non-verbal sound being able to convey information to a driver about potentially dangerous situations that might occur. Worth noting is that auditory icons (that has a natural relationship to the driving event they refer to) performed as well as speech and better than other sounds. This shows the importance of how intuitive or how much prior learning a sound requires, when looking at the efficiency of the sound. This can relate to the concept of sound cue efficiency, described previously.
What does this mean for sound in games?
The main similarity between the literature presented above is that they all seem to show that non-verbal sound can be used to convey different kinds of information. This is really important, not just for sound design in general, but also for the connection between non-verbal sound cues and player efficiency. If non-verbal sound could not convey information to the player, a connection between the two wouldn’t be able to exist. The literature above reveals connections between sound and action, including sense of direction/navigation, urgency, response time, and accuracy. The actions connected to these sounds can happen more or less efficiently. E.g. the beacon sound of the maze (a low rumbling continuous sound) might have been affected by the ambience in terms of frequency masking or possibly that the subjects grouped it with the ambience, believing that it was a part of it. In turn, making it less efficient as a navigational cue.

The literature above also seem to indicate that auditory icons are easier to understand and to associate with specific events compared to other non-verbal sounds. Auditory icons also seem to produce the fastest response times and accuracy in these kinds of scenarios. This means that a connection between if a sound is intuitive or not and player efficiency, is likely to exist. This in turn indicates that there could be a connection between player efficiency and sound cue efficiency.

The fact that the indistinctive low rumble sound as a beacon was confusing to some players of the maze in the study performed by Dahlstrøm et al. (2013), but not the short high frequency confirmation sound, also seem to indicate this. The fact that the rumble was confusing is not very surprising when looking at it from a psychoacoustic point of view, it might have been grouped with the ambience (mentioned above) and due to the fact that lower frequencies tend to be less directional than higher ones, meaning that it is harder to distinguish the exact source location of lower frequency sounds. The confirmation sound wasn’t confusing because of its short length, high frequency and synthetic nature, it was very distinctive and easy to interpret.

Since non-verbal sound cues seem to be able to convey a lot of different types of information, they should be a useful and versatile tool for sound designers. Both auditory icons and earcons can be useful in sound design, although, according to the study by McKeown and Isherwood (2007), auditory icons seem to have the upper hand when it comes to yielding fast player responses and in turn, a higher player efficiency. Auditory icons seem to have a higher sound cue efficiency than earcons. The focus of the study by McKeown and Isherwood (2007) was within-vehicle auditory displays and not video games. This makes one wonder if the same holds true in the context of video games. This was the main motivation for the study presented in this paper. This is why the purpose of this paper is to see if there is a difference between the player response times yielded by auditory icons and earcons in a video game setting.

As mentioned above, to fulfill this purpose, an experiment was conducted; an experiment described in the following section.
Method

Experiment game creation

The experiment game was created using Unreal Engine 4 (Epic Games, 2014) and was built from the “First person” template and used the “Starter content”, both included within the engine. The game level graphics was three-dimensional in which the subjects could walk in all directions (using W, A, S and D keys to move) and look around (using a mouse). Four game scenarios were created for the subjects to play through. The time from when the subject started the timed navigation section of each scenario to when he or she completed it, would be measured. The task of each scenario was for the subject to, as fast as possible, go to the correct location, or sub-area. The correct sub-area was announced by a representative audio stimulus. When the subject arrived at this correct sub-area, the scenario would end. Only one subject played through the experiment at a time.

In each scenario, the subject would spawn in the middle of a square room with one distinct location, or sub-area, in each corner. The sub-areas were represented by big pictures covering the two walls in each corner accompanied by signs with the names of the sub-areas, for example, a photo of an airport with a sign in front of it saying “Airport” (see figure 1).

![Figure 1 Scenario 1](image)

After spawning, the subject wouldn’t be able to move yet, but he or she had the ability to look around, at this point. The subjects would themselves start the timed navigation section by pressing “mouse 1” or “left click”. This in turn gave the subject the ability to move. Clicking the mouse also started a timer and played a stimulus. The idea of this was to have the subject look around themselves to see what different sub-areas were represented in each corner, then when the subject was ready, he or she would themselves start the timed navigation section. This was done so that the subjects knew what different sub-areas surrounded them before they started the measured interaction. The stimulus that played when the subject started the timed navigation section was either an auditory icon or an earcon designed to represent one of the sub-areas in the scenario. If the subjects would walk to the wrong sub-area, nothing would happen and they would just have to keep going until they navigated to the correct sub-area. The stimulus would be repeated every 10 seconds until the subject completed the scenario. This, so that if the subject were to forget what the stimulus was, he or she would be reminded.
If the stimulus was an auditory icon or an earcon in each scenario was not randomized, but instead two different experimental set-ups were designed. One with an auditory icon in scenario 1, an earcon in scenario 2, an auditory icon in scenario 3 and an earcon in scenario 4; the second experimental set-up was the opposite. This was done because in *Unreal Engine 4* (Epic Games, 2014) it wasn’t possible for the author to both completely randomize the presentation of scenarios and guarantee that each subject and scenario experienced a balanced number of auditory icons and earcons. Subject 1 would then play experimental set-up 1, subject 2 would play experimental set-up 2, subject 3 would play experimental set-up 1, subject 4 would play experimental set-up 2 and so on. To remove any systematical errors, the order that the scenarios were presented in each experimental set-up, was randomized every time a new instance of an experimental set-up was launched (the order was randomized on every play-through). Another two versions of the experiment were created in which the signs were in either English or Swedish. This was done so that if the subject’s native or preferred language was Swedish, he or she could play the Swedish version and if the subjects native or preferred language wasn’t Swedish, he or she could play the English version. All subjects were either native Swedish or English speakers.

At the beginning of the experiment and in between each scenario, the subjects spawned/were taken to a virtual training level. The training level included a portal that took the subjects to each scenario. The purpose of this was to let the subjects familiarize themselves with the controls of the game before they started playing through any of the scenarios. After having played through all of the scenarios the subject would walk into the portal one last time, this would take away the subject’s movement and display the temporarily stored completion times of each scenario; which until that point had been hidden to the subject. The timing data was then recorded by manually copying the completion times of each scenario into a separate document.

*Figure 2 Training area with portal*

(Pictures of the rest of the scenarios can be found in Appendix A)
The sub-areas in each scenario followed a theme so that each scenario had sub-areas similar to each other. This was done to make the scenarios feel different from one another. Another important thing was that no sub-area was present in more than one scenario. This to try and further distinguish each scenario from one another.

The different sub-areas in the four scenarios were as follows:

- **Scenario 1:** *Airport*, cinema, circus and zoo.
- **Scenario 2:** *Winter*, spring, summer and fall.
- **Scenario 3:** *Desert*, lake, forest and beach.
- **Scenario 4:** *Hospital*, school, bank and church.

The first sub-area for every scenario, in italics, in the list above was the correct sub-area that the subjects were supposed to go to.

**Stimuli**

The sub-areas were represented by both an auditory icon and an earcon, this in order to compare the two types of non-verbal sounds against each other. When choosing stimuli to represent each scenario, there had to be more criteria than just the sound being an auditory icon or an earcon. Because of this, a lot of care was taken to have both the auditory icon and the earcon being *equivalent representatives* of each sub-area. This was the most important aspect that was considered when choosing stimuli. Since if, for example, the auditory icon and the earcon representing the airport sub-area were nowhere near equivalent, let’s say a recording of an airplane taking off (a good auditory icon representation of an airport) and a random jingle (a bad earcon representation with no connection to an airport at all), there would be no point in comparing them to each other.

The other important aspect taken into consideration when choosing stimuli was that no sounds were to be spatialized (placed as sound sources in the game world) and to have all of the sounds playing in mono. This in order to remove any influence that panning and having a sound in the game world that you can hear the location of, would have had on the results. The idea was to let the sounds speak for themselves and have the subjects reacting to only the information that the sounds themselves conveyed. Lastly the sounds were to be approximately the same length, this to remove any advantage or disadvantage that a lot longer sounds might have.

The stimuli for each correct sub-area were the following:

**Airport:** Airplane passing over head at a close distance (auditory icon) and an announcement chime (earcon). An airplane passing over head is a good representation of an airport because an airport is where airplanes take off and land. You probably will hear this sound at an airport or for example, have heard it in a movie when a scene switches to an airport setting. An announcement chime is something that is very often heard at airports (or other means of transportation like trains). These types of chimes are seldom heard at any other places than airports and places like trains. This makes it a good representation of an airport.
**Winter:** Heavy snow footsteps (auditory icon) and sleigh bells (earcon). The picture representing this sub-area is of a forest path covered by a thick layer of snow. This means that if you were to be in a location like that in real life, heavy snow footsteps is a sound you most certainly will hear. This makes it a good representation of the winter portrayed by the sub-area. The sound of sleigh bells is something that isn’t necessarily representative of winter itself, but more the tradition of celebrating Christmas. Sleigh bells are heard in almost every Christmas song, they’re present at some point in most Christmas movies and on some occasions experienced when they’re attached to actual sleighs. This makes them a good representative of winter.

**Desert:** Hawk cry/shriek (auditory icon) and a western/cowboy slide guitar lick (earcon). Hawk shrieks are often associated with deserts. They are heard in most western/cowboy movies taking place in a desert and also heard when being in real deserts. This makes it a good representative of a desert. The slide guitar lick is something heard in a lot of western/cowboy movies. Since most of these movies take place in deserts, the slide guitar lick is a good representative of a desert.

**Hospital:** A car crash (auditory icon) and a heart monitor beeping more rapidly until a continuous tone is heard (earcon). The sound of a car crash isn’t something you will generally hear in a hospital, but in car crashes people tend to get injured. When someone’s injured, they most likely have to go to a hospital. This is why a car crash is a good representative of a hospital. The sound of a heart monitor is a sound you might actually hear in a hospital. Though a very common way of hearing this sound of a heart monitor beeping faster and faster and finally flat lining, is something very often heard in movies or television-shows taking place in a hospital or when a character dies in a hospital. This makes it a good representative of a hospital.

Besides the audio stimuli, the experiment game featured footstep sounds for the subject’s character. This in order to make the subjects feel as they were not just floating around and to make the experience a bit more immersive. Care was taken to make sure that the character footsteps did not mask any potential information provided by the stimuli. This was done by making them significantly lower in amplitude than the stimuli.

The stimuli were designed using *Reaper* (Cuckos, 2015) and sounds from various free online sound libraries.

**Subjects**

A total of 23 subjects participated in the experiment, all with some kind of previous gaming experience. The reason for the criteria that all subjects had to have some sort of previous gaming experience was in order to gain more ecological validity in the experiment and also to make sure that a subject’s response times weren’t majorly affected by their inexperience with controlling a character in a game. All subjects were students at Luleå University of Technology in Piteå, this is also where the experiment was conducted.

After the subjects had played through the experiment, they were asked to fill out a questionnaire. The questionnaire was designed to gather demographic information about the subjects’ previous experiences, habits and preferences in areas deemed relevant for further analysis of the data gathered by the experiment. These areas were:

- *Where they lived during their upbringing*, this can affect how they relate to different sounds and what sounds they relate to.
- *Their gaming habits*, information about how many hours per week they play games etc.
- *Previous musical experience*, this can affect how they relate to musical earcons.
- *Preferences for turning off the sound of their personal electronic devices*, this can affect how they relate to different auditory icons and earcons.
The first area (where they lived during their upbringing) was included based on the ideas of Jägerskogh (2000), who mentions that we learn how to interpret and associate different sounds from a very young age. This means that what associations we have and how we react to different sounds, can be influenced by what we have experienced in the past. This is in turn affected by where one lived when growing up and what experiences one might have had because of this. That is the reason for retrieving this information from the subjects.

The questionnaire consisted of the following questions:

1. **Where did you live during the majority of your childhood/upbringing? Please include, country and city or part of country.**

2. **Approximately how many hours per week do you spend playing games?**

3. **What types of games do you usually play?**

4. **What platform/s do you use to play your games?**

5. **Do you play any instruments and/or are you a vocalist?**

6. **How long have you been a musician or when did you start learning/studying music?**

7. **What sound-making, personal electronic devices do you have/use frequently and do you leave the sound on when you use them? For example, phones, laptops, tablets, etc.**

8. **If you leave the sound on, do you think the sounds help you use the devices more efficiently? If so, how?**

As mentioned above, separate English and Swedish version of the experimental set-ups were made, the same was done for the questionnaire. The experiments were conducted in Piteå located in the northern part of Sweden, because of this chances were that subjects would most likely be Swedish. The English versions were made as a precaution in the event of subjects not being Swedish and their first or second language being English rather than Swedish. The university does have foreign students. The subject could choose which language he or she played through the experiment and answered the questionnaire in.

**Test set-up**

The experiment was conducted in two different locations. On average it took about 5-15 minutes for the subjects to complete the experiment, depending on how long it took them to answer the questionnaire following having played through the experiment game.

Location 1 was the author’s apartment (very close to Luleå University of Technology in Piteå) where the experiment game was run on a custom built desktop PC running Windows 10 on an Intel i5 3570K 3.4GHz quad-core processor and an Asus GeForce GTX 680 DC2 graphics card. The subjects played using a Full-HD 24” monitor and mouse and keyboard. A home environment was chosen as a primary location since it is a more realistic gaming environment than, for example, a studio or a control room. Since the focus of the experiment wasn’t for subjects to listen for and compare fine details in different sounds, noise usually present in an apartment (such as noise from neighboring apartments or from outside the apartment) was deemed to have no significant impact on the results. This is also why it was deemed acceptable to use a Edifier M3200 Blue Band 2.1 speaker set-up. Another reason speakers were used instead of headphones was in order to ease the communication between the subject and the author during the experiment. The instructions were given to the subjects orally and the subjects had the possibility to ask questions during the experiment when they were taken to the training level between
every scenario. This was done in order make the subjects more comfortable and also to make it seem less like an experiment and more like a normal gaming experience. Take note that no clues about the correct sub-locations or upcoming audio stimuli were given to the subjects.

Location 2 was a computer lab at Luleå University of Technology in Piteå. The reason this secondary location was used was in order to get more subjects to participate in the experiment. The reason the computer lab was chosen as the location in Luleå University of Technology in Piteå was because the computers in that computer lab had good enough hardware to reliably run the experiment game. The experiment game in this location was run on a desktop PC running Windows 7 on an Intel i5 4570 3.2 GHz quad-core processor and a GeForce GTX 750 graphics card. The subjects played using a Full-HD 24” monitor and mouse and keyboard. At this location a pair of Steel Series Siberia V2 headphones were used instead of a speaker set-up. This was because of the computers lack of a proper speaker set-up, but also because of the fact that the computer lab is free for anyone to use. This means that the possibility of interfering noise (from for example others using the computer lab) was a lot higher than in location 1 and for this reason headphones were used to eliminate as much interfering noise as possible.

Approximately 50% of the subjects who participated, did so in each location.

Results

The results of the experiment will be presented in the following tables and graphs:

Table 1 Response times of subjects and scenarios.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Scenario 1 (seconds)</th>
<th>Scenario 2 (seconds)</th>
<th>Scenario 3 (seconds)</th>
<th>Scenario 4 (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4,57</td>
<td>4,47</td>
<td>21,70</td>
<td>7,64</td>
</tr>
<tr>
<td>2</td>
<td>4,10</td>
<td>3,63</td>
<td>10,62</td>
<td>3,85</td>
</tr>
<tr>
<td>3</td>
<td>3,38</td>
<td>3,83</td>
<td>29,16</td>
<td>6,36</td>
</tr>
<tr>
<td>4</td>
<td>5,43</td>
<td>4,30</td>
<td>4,71</td>
<td>3,80</td>
</tr>
<tr>
<td>5</td>
<td>3,95</td>
<td>3,44</td>
<td>5,97</td>
<td>3,62</td>
</tr>
<tr>
<td>6</td>
<td>4,57</td>
<td>4,04</td>
<td>5,24</td>
<td>10,12</td>
</tr>
<tr>
<td>7</td>
<td>6,54</td>
<td>4,18</td>
<td>5,91</td>
<td>4,47</td>
</tr>
<tr>
<td>8</td>
<td>5,38</td>
<td>6,73</td>
<td>8,55</td>
<td>15,44</td>
</tr>
<tr>
<td>9</td>
<td>3,05</td>
<td>2,92</td>
<td>3,52</td>
<td>4,76</td>
</tr>
<tr>
<td>10</td>
<td>4,26</td>
<td>3,51</td>
<td>6,08</td>
<td>4,76</td>
</tr>
<tr>
<td>11</td>
<td>3,34</td>
<td>2,76</td>
<td>3,27</td>
<td>4,6</td>
</tr>
<tr>
<td>12</td>
<td>5,45</td>
<td>4,16</td>
<td>10,39</td>
<td>5,06</td>
</tr>
<tr>
<td>13</td>
<td>5,80</td>
<td>5,17</td>
<td>6,12</td>
<td>5,42</td>
</tr>
<tr>
<td>14</td>
<td>3,15</td>
<td>4,93</td>
<td>3,79</td>
<td>8,80</td>
</tr>
<tr>
<td>15</td>
<td>2,88</td>
<td>4,51</td>
<td>3,66</td>
<td>4,54</td>
</tr>
<tr>
<td>16</td>
<td>4,51</td>
<td>4,43</td>
<td>15,60</td>
<td>18,15</td>
</tr>
<tr>
<td>17</td>
<td>2,57</td>
<td>2,96</td>
<td>3,58</td>
<td>2,62</td>
</tr>
<tr>
<td>18</td>
<td>6,79</td>
<td>5,52</td>
<td>6,96</td>
<td>24,47</td>
</tr>
<tr>
<td>19</td>
<td>5,81</td>
<td>3,33</td>
<td>10,53</td>
<td>4,38</td>
</tr>
<tr>
<td>20</td>
<td>4,86</td>
<td>4,85</td>
<td>14,67</td>
<td>20,83</td>
</tr>
<tr>
<td>21</td>
<td>3,36</td>
<td>3,77</td>
<td>3,73</td>
<td>2,69</td>
</tr>
<tr>
<td>22</td>
<td>7,08</td>
<td>5,98</td>
<td>17,40</td>
<td>18,22</td>
</tr>
<tr>
<td>23</td>
<td>2,95</td>
<td>3,40</td>
<td>5,12</td>
<td>4,19</td>
</tr>
</tbody>
</table>
Figure 3 Box plot of combined response times of auditory icons and earcons for scenario 1

Figure 4 Box plot of combined response times of auditory icons and earcons for scenario 2
Figure 5 Box plot of combined response times of auditory icons and earcons for scenario 3

Figure 6 Box plot of combined response times of auditory icons and earcons for scenario 4
Since the data from the experiment game was the main focus of the experiment, the data from the questionnaires will not be mentioned in the main results section. They will however, be summarized in the “Demographic information” section of the analysis section (the raw data of the questionnaires can be found in Appendix B).

**Analysis**

**Experiment game**

Subject 3 had to be excluded from the analysis since an error occurred during the experiment where the subject walked to the correct sub-area in scenario 3, but failed to walk far enough into the sub-area to trigger the end of the scenario. This made the subject believe that the correct sub-area wasn’t the correct one, and continued to search among the other incorrect sub-areas. This in turn lead to a very misrepresentative result. For this reason, the subject was excluded from the analysis.

From the results we see that the earcons on average produced similar but slightly faster player responses than the auditory icons. We can also see that the response times produced by auditory icons had a bigger spread than those of the earcons.

In order to analyze the results to see if the difference between the combined response times of the auditory icons and the earcons for all scenarios was significant, a paired two tailed t-test was conducted. The t-test used the mean values of the two auditory icon times and the two earcon times for each subject. The significance level was set to 0,05. The t-test showed that there was no significant difference between the response times for the auditory icons compared to the response times of the earcons ($p = 0,0624$).

Since half of the subjects took part in the experiment in two different locations with different monitoring conditions, unpaired two tailed t-tests were performed to see if there was a significant difference in the player response times between the two different locations for both the auditory icons and the earcons. Although the response times recorded at location 2 (the computer lab) were slightly slower than the ones in location 1 (the author’s apartment), the t-tests showed that there was no significant difference between the two locations (auditory icons $p = 0,6343$ and earcons $p = 0,5257$).
Due to the fact that the results of the experiment had a big spread and were a bit skewed, a log-transformation was performed. This was done in order to try and make the distribution of the results less skewed. A log-transformation was done for the results of each individual scenario and for the combined data of all scenarios. Unpaired two tailed t-tests were then performed for each scenario and an additional paired two tailed t-test was performed for the combined data for all scenarios. The reason for this was to see if any further analysis would yield any additional interesting results. Why was a paired t-test used on the combined data of all scenarios, but unpaired t-tests used on the data of the individual scenarios? The reason for this is because for all scenarios combined, all subjects were exposed to both auditory icons and earcons, making it the same group of subjects that were exposed to the two different types of stimuli. Because of how the experiment was constructed, for each individual scenario the subjects were only exposed to either an auditory icon or an earcon. For example, in scenario 1 subject 1 was exposed to an auditory icon, subject 2 was exposed to an earcon and so on. This makes it so that for each individual scenario, it was two different groups of subjects that were exposed to the two different types of stimuli. This is the reason for using both paired and unpaired t-tests in this part of the analysis. The significance level for the log-transformed t-tests was also set to 0.05.

The results of the paired two tailed t-test for the log-transformed results for all scenarios combined, also showed that there was no significant difference between the response times of the auditory icons and earcons ($p = 0.1321$). More interesting are the results of the individual scenarios, the $p$-values of each scenario are as follows:

- Scenario 1, $p = 0.0589$
- Scenario 2, $p = 0.0116$
- Scenario 3, $p = 0.0843$
- Scenario 4, $p = 0.0020$

This shows that for the log-transformed results of the individual scenarios, the difference between the auditory icons and earcons was significant for scenario 2 and scenario 4, but not for scenario 1 and scenario 3. An important aspect of these results is that the results of the individual scenarios should be seen more as a comparison of the specific stimuli for each scenario and not a general comparison between auditory icons and earcons. That is instead the purpose of the paired t-test of the combined log-transformed results for all of the scenarios.

If we look at the box plots of the scenarios individually, we can see that in scenario 1 and 2 the auditory icons and earcons produced similar response times and are very tightly grouped. It is scenario 3 and especially scenario 4 that produced the slowest response times for both auditory icons and for earcons. The auditory icons in scenario 3 (the hawk shriek representing the desert) had two major outliers, one of which is subject 3 who will be excluded from the rest of the analysis. Worth noting is that the other outlier was a subject who (in informal discussion with the author) said that he or she had no previous experience with the stimulus and because of that, had no idea what sub-area the stimulus was supposed to represent. The fact that the earcons in scenario 3 (the western/cowboy slide guitar lick) the produced the slowest and most spread out response times of the earcons was probably because 3 subjects mistook the guitar lick for a “surf-rock” guitar lick and therefore thought it represented the beach sub-area also present in scenario 3; this was mentioned in informal conversations between the subjects and the author after they had played the experiment game.
We can also see that in scenario 4, the auditory icons and earcons produced the least similar response times. This can also be seen in the results of the t-test for scenario 4. The response times of the auditory icon stimulus in scenario 4 (the car crash representing the hospital) were quite slower and also more spread out. This shows that the auditory icon stimulus in scenario 4 was a worse representative of a hospital than the earcon stimulus (a heart monitor). This might be due to its more segmented structure. It starts with the sound of a car loosing grip and skidding, this was then followed by the sound of a crash with metallic characteristics. This means that it’s a possibility that before the subjects could realize that it was a car crash and in turn connect that to the hospital, the stimulus had to play for a certain amount of time to get to the crash part of the stimulus. Some subjects initially mistook the auditory icon stimulus as a representative of the bank sub-area in the same scenario; this was also mentioned in informal conversations between the subjects and the author after they had played the experiment game. This because they thought that the sound of wheels loosing grip and skidding in the beginning of the stimulus, was that of a “getaway-car” trying to get away from a bank robbery. This could be reasons for the poor performance of the stimulus.

Another interesting thing to note is that in scenario 2, the earcon (sleigh bells representing winter) produced faster response times than the auditory icon (heavy snow footsteps). The results of the t-test for scenario 2 also show this. This is an indication of how strong associations the subjects had and how strongly they relate to the tradition of celebrating a religious holiday and all of the traditions and conventions that comes with it. The subjects were almost exclusively Swedish and in Sweden, Christmas is a very big holiday for the majority of the population that is celebrated every year. This in turn enables people to form associations with it from a very young age, which can be the reason for the strong connection that the results show the subjects had to the stimulus.

Demographic information
The questionnaire that the subjects were tasked with after playing through the experiment game, provides demographic information about the subjects relevant for further analysis. The core meaning of the questions present in the questionnaire will be presented and analyzed.

The way the questions were analyzed, was that the subjects were grouped based on how they answered the different questions. The author then manually looked at and compared the test results of the different groups of each question, to see if any patterns were present between how the subjects answered the different questions and how they performed in the experiment game. The kind of patterns looked for was, for example, if subjects that play games more than 10 hours per week performed better in the experiment game than the subjects that spend less than 10 hours per week playing games. More examples of the kind of patterns looked for will also be given for each question of the questionnaire below.

1. Place of origin
A subject’s place or origin can affect how they relate to different sounds and what sounds they relate to (Jägerskogh, 2000). This means that subjects’ places of origin could have had an effect on their results in the experiment game. Due to the fact that the subjects were almost exclusively from Sweden, and because Sweden is a very long and narrow country, the subjects were categorized as from the norther part of Sweden, from the southern part of Sweden and from any other country. For example, since winter is very different in the northern part of Sweden compared to the southern part of Sweden, this might have had an influence on the subjects’ response times when reacting to the auditory icon stimulus in scenario 2.
Patterns like those observed by Jägerskogh (2000) and the one mentioned in the example above were looked for, but no patterns of this nature were found. The subjects’ places of origin did not seem to have had an influence on their results in the experiment game.

2. Game time per week
How many hours per week a subject spends playing games can be a good indication of how used to playing games the subject is and how skilled they might be at playing games. For this reason, how many hours per week a subject spends playing games, might have had an effect on their results in the experiment game. For example, a subject who spends more than 10 hours per week playing games might have an easier time controlling the character in the experiment game than a subject who spends less than 5 hours per week playing games and therefore, might have gotten faster response times.

Figure 9 Game time per week (the 5-10 bracket includes answers of 6-10 and so on)
No patterns related to how many hours per week subjects spend playing games were found. The subjects’ game time per week did therefore not seem to have had an influence on their results in the experiment game.

3. Game genres
What different genres of games the subjects tend to play can influence what different sets of skills they might have. This might have had an effect on the subjects’ results in the experiment game. For example, a subject who plays primarily FPS (first person shooter) games might be better at controlling the character in the experiment game compared to a subject who primarily plays 2D platforming games. Also a subject who plays both FPS games and other types of games might be better at understanding the treasure hunt task in the experiment game than a subject who only plays FPS games. Since the experiment game was in a first person perspective and the majority of the subjects had some previous experience with FPS games, the subjects were categorized by if they played only FPS games, no FPS games or FPS and other types of games.

![Game genres](image)

*Figure 10 Game genres*
No patterns related to the subjects’ preference in game genre were found. Because of this, the subjects’ preference of game genres did not seem to have had an influence on how fast their response times were in the experiment game.

4. Gaming platform
What platform the subjects use to play their games might have had an influence on their results in the experiment game. For example, a subject who plays games on a PC (personal computer) might be more familiar with using a mouse and keyboard to control their character in the experiment game than a subject who does not use a PC to play their games. This might influence their results. Since the experiment game was played on a PC using mouse and keyboard the subjects were categorized as using, only a PC, PC and gaming consoles/handheld devices and only gaming consoles/handheld devices, to play their games.
No patterns related to the subjects’ preference of gaming platform, like the one mentioned in the example above, were found. The subjects’ preference of gaming platform did not seem to have had an influence on their results in the experiment game. In fact, the training area of the experiment game was partly designed to eliminate this kind of influence on the results, this seems to have been a success.

5. Previous musical experience
If the subjects had previous musical experience or not could have had an influence on their results in the experiment game. For example, a subject with previous musical experience might relate more strongly to earcons of musical nature (like the slide guitar lick representing the desert in scenario 3) than a subject with no previous musical experience, in turn making them respond quicker to this type of stimulus.
The subjects previous musical experience did not seem to have had an influence on their results in the experiment game. Since only two out of the 22 subjects included in the analysis had no previous musical experience, even if a pattern like the one mentioned in the example above had been found, it would not have been enough to confirm its validity as a pattern due to the group of subjects with no previous musical experience being too small.

6. **Duration of musical experience**

How many years of musical experience the subjects have, could also have had an influence their results in the experiment game. For example, a subject with 10 years or more of musical experience might relate more strongly to earcons of musical nature than a subject with 5 or less years of musical experience, making them respond quicker to this type of stimulus.

![Figure 13 Duration of previous musical experience](image)

*Figure 13 Duration of previous musical experience (the 5-10 bracket includes answers of 6-10 and so on)*

The duration of the subjects’ previous musical experience, did not seem to have had an influence on their results in the experiment game. No patterns, between how many years of previous musical experience the subjects had and their results in the experiment game, were found.

7. **Personal electronic devices (sound/no sound)**

Subjects’ experience and habits regarding personal electronic devices, in this case if they leave the sounds of notifications/prompts/events on or if they turn them off, might have had an influence on their results in the experiment game. For example, a subject who is used to hearing notification sounds (and is used to reacting/responding to those types of sounds), might be better at interpreting the information provided in the different stimuli used in the experiment game, than a subject who is not used to hearing those types of sounds. This could enable them to respond faster to the stimulus.
patterns like the one mentioned in the example above were looked for, but no patterns of this nature were found. What different personal electronic devices the subjects had and if they leave the sound of things like notifications on or off, did not seem to have had an influence on their results in the experiment game.

8. Thoughts on notification type sounds

The final question on the questionnaire was analyzed in order to find themes in the subjects’ answers about their thoughts on how notification type sounds aided them in their everyday use of their personal electronic devices. Three themes were found:

- **Theme 1. Easier to notice notifications**, this means that the subjects believed that the notification type sounds made it easier for them to notice notifications. “Yes, I notice, for example, notifications easier.” (Subject 7).

- **Theme 2. Instant info about type of notification**, this means that the subjects believed that, because different applications on their personal electronic devices have different notification sounds, they could instantly know what type of notification they had gotten as soon as they received the notification (for example, a text message, an E-mail or a phone call). This without even looking at their personal electronic device. “It’s useful that on my smartphone, different applications have different sounds. That way I can know what type of notification I’ve gotten before I even look at my smartphone.” (Subject 2).

- **Theme 3. Simple feedback increases usability**, this means that the subjects believed that notification sounds, like the simple non-verbal sound cues giving feedback to the player/user mentioned previously in the paper, helped them use the device more easily or more efficiently. For example, a program on a laptop providing the user with an error message when trying to do something not possible in the program or a confirmation sound when sending an E-mail. “The sounds on my laptop usually notify me when something happens. For example, error messages, when a download has completed etc. The sounds help to redirect my attention when I’m focusing on something else.” (Subject 4).

(The subjects’ full translated answers for question 8 can be found in Appendix B)
Which of these categories/themes subjects were placed in, might have had an impact on their results in the experiment game, similar to the example given for question 7 above, but for each of these three categories/themes. An example of a comparison made between which theme subjects had been placed in and their results in the experiment game, would be: A subject who believes that notification type sounds give him or her useful feedback aiding them in the use of their personal electronic device, might be more used to listening for these types of confirmations than a subject who doesn’t feel the same way. This could increase their ability to decode the information given to them by the stimuli in the experiment game.

![Subjects placed in themes](image)

*Figure 15 Amount of subjects placed in the three themes found in question 8*

What themes subjects’ answers of question 8 in the questionnaire could be placed in, did not seem to have had an influence on their results in the experiment game.
Discussion

Experiment findings

The purpose of this paper was to see if there is a difference between the player response times yielded by auditory icons and earcons in a video game setting. From the analysis of the results from the experiment game, we can see that there does not seem to be a significant difference between the combined response times of the auditory icons compared to the combined response times of the earcons for all four scenarios; although for the individual scenarios, the earcon stimuli in scenario 2 and scenario 4 were significantly faster than the auditory icon stimuli. This contradicts the results of the study by McKeown and Isherwood (2007) about auditory displays in vehicles, that showed that auditory icons seem to produce faster response times than earcons. What could be the reason for this? Are games really that different from auditory displays?

The reason for this might be the difference in the process used to choose and match stimuli to their referents in the two studies. In the study by McKeown and Isherwood (2007), stimuli were matched to their referents (in this case driving events) by first being rated on a perceived urgency scale of 1-3 and then matched with referents with the same score on the scale. In the experiment described in this study, stimuli were matched to their referents (in this case the different sub-areas of the four scenarios), or rather created/chosen to match their referents, by first looking at the referents and then trying to figure out what sound of each type (earcons and auditory icons) would be a good and equivalent representative of the referent.

An example of the pairing of stimuli to their referents in the study by McKeown and Isherwood (2007) is, for the driving event “The car is drifting off the road” (p. 420), the sound of driving over “rumble strips” as the auditory icon stimulus and a “high-rate, high-pitched tone alarm” (p. 420) as the earcon stimulus. The sound of driving over “rumble strips” is something that can be commonly associated with a car drifting off the road. Since many roads are equipped with these “rumble strips” it is not unusual that a person has experienced this pairing in the real world. On the other hand, the “high-rate, high-pitched tone alarm” is something that is probably very hard to associate with the driving event without some prior information about the pairing, which the results of the study by McKeown and Isherwood (2007) seem to indicate. If you were to compare these pairings to those of scenario 1 in the experiment game, an airplane passing over head at a close distance as the auditory icon and an announcement chime as the earcon, representing an airport. We can see that the pairings of scenario 1 in the experiment game are more equivalent representatives of their referent than those of the example from the study by McKeown and Isherwood (2007).

Because of this, the reason for there not being a difference between the combined response times of the two types of nonverbal sounds, is possibly because a lot of time was spent on trying to make sure that all sounds were equally good representatives of the referent. Maybe the stimuli in the experiment game were too well matched or maybe neither of them was a particularly intuitive representation. The point is that, when combining the results of all four scenarios in the experiment game, the auditory icon and earcon stimuli seem to have been more equal representatives of their referents than in the study by McKeown and Isherwood (2007). Any advantages that auditory icons might have had in the study by McKeown and Isherwood, like perhaps being more easily recognized and remembered than their more abstract and synthetic earcon counterparts, were possibly completely removed by the process of matching stimuli to their referents used in this study. This because both the earcons and auditory icons were chosen because they were thought to share these different advantages, compared to other sounds not chosen to be used as stimuli.
One could say that the stimuli in this experiment were chosen based on how intuitive and easy to interpret they were; an aspect of sound cue efficiency discussed previously in this paper. This is another example of how important the aspect of intuitiveness can be for sound cue efficiency. This aspect is something completely reliant on what different associations people might have with different sounds. The importance of which could be seen in the results of scenario 2, discussed in the analysis section. This could mean that, for efficiency, the intuitiveness of a non-verbal sound might be more important than if it’s an auditory icon or an earcon. This also serves as an example of sound cue efficiency affecting player efficiency, which in turn indicates a connection between the two. This is something very interesting, making the area more relevant for further research.

Another interesting thing was found in the analysis of scenario 4, this being that simple sounds seem to be more efficient as sound cues, than more complex sounds with multiple components at different points in time. The auditory icon of the car crash (representing the hospital) performed poorly compared to its earcon counter part of the heart monitor. In more complex sounds of this nature, there are more things for players to process before the desired information can be retrieved. It is also the fact that the sound has to play through all of its different sections before the player knows what is actually going on. This in turn could make it a less efficient sound cue. This could mean that simpler sounds have a higher sound cue efficiency. No really inefficient earcons or auditory icons were observed in this experiment, this makes it another area relevant for further research.

Something worth discussing is whether or not the earcon stimuli of the heart monitor and the airport chime are to be considered as earcons or as auditory icons. The reason for this being that, these two sounds are generally experienced in the real world and this is where associations to these sounds are made. This means that they can be considered to be auditory icons. They are however considered to be earcons in the context of this paper. The reason for this being that, if you look at the core definition of earcons, they are short and structured synthetic or musical sounds and this is exactly what the heart monitor and airport chime stimuli used in the experiment are. This is reinforced by the fact that the heart monitor and airport chime stimuli used in the experiment, are not recorded real world sounds, but instead completely dry, synthetic recreations of these sounds, not accompanied by appropriate reverb or other ambient noises usually heard in combination with these sounds. Another difference between them and their real world counterparts, is that in the real world a heart monitor is continuous and an airport chime is directly followed by an announcement. Meanwhile, the stimuli used in the experiment does not share these traits. For these reasons, they are considered to be earcons and not auditory icons.

**Future game sound design and development**

The fact that non-verbal sound cues seem to be able to convey different kinds of information and serve as both simple and complex types of feedback to players, makes them a really interesting aspect to take into consideration when doing sound design and when designing a game. This makes it possible for game designers to implement game mechanics that revolve around non-verbal sound cues and to make sound a really big part of the gameplay.

Since it seems as though there is no significant difference in the efficiency of auditory icons and earcons, which can be seen in the results of the two paired t-tests for the combined results for all four scenarios, a sound designer has more freedom when it comes to choosing what types of non-verbal sounds they incorporate into their sound design. This means that non-verbal sound cues are a very versatile tool for a sound designer and what function in the game they have or what information they convey in a particular situation is almost completely up to the sound designer. They can become what you make them out to be.
Something worth considering is that the results and findings of this experiment and previous research mentioned in this paper, were derived from experiments or scenarios with a predominant absence of multitasking and with quite big time windows in which to complete the different tasks. For example, driving a car in real life or playing a video game like an FPS, are situations where you have to process a lot of different kinds of audio and visual information and you often also have to do this quite quickly. This is an important aspect worth considering. How well do the findings presented in this paper actually hold up in “real world” situations like the ones mentioned above? What are the limits of this study’s applicability? This is something in need of answering by future research on these areas.

Put shortly, non-verbal sound cues, and sound in general, have the potential to become a much bigger part of game design and game sound design than what they have been in the past. It is all up to the game designers and sound designers to use them to their full potential.

**Experimental design**

The experimental design in this study was not perfect and some choices and compromises had to be made. There are a few things that could have been done differently, that might have had an impact on the results of the experiment.

The type of scenario, a treasure hunt like scenario were the subjects had to find the correct sub-area based on the stimulus, used in the experiment game and the parameter that was being measured, could have been different. In this case it was how long it took the subjects to navigate to the correct sub-area from when they heard the stimulus that was measured. In other words, their response times. Something else, like accuracy for example, could have been the focus instead of response time. This means that a completely different type of scenario had to have been created. A different scenario and experiment like this might have had differing results compared to the ones of the experiment design for the purpose of this study.

The fact that the experiment game did not include a test scenario that the subjects could play through to get more familiar with the conditions of the experiment game, could have had an impact on the results. No data was gathered on what order each subject played through the four scenarios, but it is possible that some learning occurred during the subjects’ play through of the experiment game, making them perform better on the scenarios they played last. A test scenario could have decreased the influence of the learning that might have occurred during the experiment. A test scenario would have also helped to eliminate errors such as that of subject 3, who didn’t walk far enough into the sub-area to trigger the completion of the scenario, in one of the scenarios.

Something that possibly had an impact on the results were what stimuli was chosen to represent the different sub-areas. It is likely that having the subjects exposed to other stimuli would have had some kind of impact on their results. It is hard to say how things would have turned out with different stimuli, without actually performing another experiment using different stimuli to represent the different sub-areas. This is something to consider for further research on the topic.

Something that might have produced some interesting information about the results is if questions about what subjects thought of and how they reacted to the different stimuli, would have been included in the questionnaire. This could have provided greater insight as to why the results were what they were.

In hindsight, something that would have been a good addition to experiment game, is to have had implemented some sort of confirmation (either by a confirmation sound or preferably a text message) that the correct sub-area had been reached. This in combination with the inclusion of a test scenario might have helped the subjects understand the conditions of the experiment game better and possibly to eliminate errors such as that of subject 3.
Another thing that might have had an impact on the results was the selection of test subjects. It would be interesting to see what results a different selection of subjects would have produced. It is unlikely that the results would have differed significantly from the ones of this experiment, depending on what specific group of subjects were to be tested, but nonetheless it would have been interesting to see what results different groups of subjects would have produced.

Something else that could have impacted the results were what locations and under what conditions the experiment was conducted. As we saw during the analysis, results of the subjects that took part in the experiment on the two different locations did not differ significantly from each other. It is however a possibility that other locations and conditions than the ones mentioned in the paper, could have had an impact on the results.

**Conclusion**

In order to see if there was a difference between the player response times yielded by auditory icons and earcons in a video game setting, an experiment game was created. A game where the response times of subjects were recorded when reacting to auditory icons and earcons, while playing through four different treasure hunt like game scenarios. The results from the experiment seem to indicate that there was no significant difference between the combined response times produced by the auditory icons compared to the combined response times produced by the earcons. However, for the individual scenarios the earcon stimuli in scenario 2 and scenario 4 were significantly faster than their auditory icon counterparts, but this should be seen more as a comparison of the specific stimuli of those scenarios and not a comparison of earcons and auditory icons in general. From the results we can see that when it comes to determining the efficiency of a non-verbal sound cue, whether or not it is an auditory icon or an earcon seems to be less important than its intuitiveness and what associations people might have with it. We can also see that a connection between sound cue efficiency and player efficiency, is likely to exist.

**Further research**

The area of non-verbal sound cues in games is relevant for further research. Research that could focus on, for example, more similarities and dissimilarities between auditory icons or maybe the current role of non-verbal sound in games and more in depth about what possibilities non-verbal sound have to create enjoyable gameplay experiences.

The two topics sound cue efficiency and player efficiency seem like really interesting topics for further research. Especially the connection between the two and how they might affect each other. But also a more in depth investigation about what sound cue efficiency actually is and what it actually means for a sound cue to be efficient.
References


Appendix A
Experiment game screenshots

Figure 16 Training area with portal

Figure 17 Scenario 1
Figure 18 Scenario 2

Figure 19 Scenario 3
Figure 20 Scenario 4
### Appendix B

#### Questionnaire raw data

*Table 2 Questions 1-3 Subject 3 (in bold) excluded from analysis*

<table>
<thead>
<tr>
<th>Subject</th>
<th>1. Place of origin</th>
<th>2. Game time/week (Hours)</th>
<th>3. Game genres</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>North SWE</td>
<td>10</td>
<td>FPS, Racing, Adventure, Platforming, Party</td>
</tr>
<tr>
<td>2</td>
<td>South SWE</td>
<td>7</td>
<td>FPS</td>
</tr>
<tr>
<td>3</td>
<td><strong>South SWE</strong></td>
<td>2</td>
<td><strong>Mobile</strong></td>
</tr>
<tr>
<td>4</td>
<td>North SWE</td>
<td>2</td>
<td>Platform, Racing</td>
</tr>
<tr>
<td>5</td>
<td>South SWE</td>
<td>5</td>
<td>FPS, RPG</td>
</tr>
<tr>
<td>6</td>
<td>South SWE</td>
<td>6</td>
<td>FPS, RPG 2D-Platformers</td>
</tr>
<tr>
<td>7</td>
<td>South SWE</td>
<td>16</td>
<td>FPS, MMO, MOBA, Strategy, Survival</td>
</tr>
<tr>
<td>8</td>
<td>North SWE</td>
<td>6</td>
<td>Puzzle, FPS, Open World/Sandbox</td>
</tr>
<tr>
<td>9</td>
<td>North SWE</td>
<td>10</td>
<td>FPS, RTS</td>
</tr>
<tr>
<td>10</td>
<td>South SWE</td>
<td>6</td>
<td>FPS, Adventure, Sci-fi/survival horror</td>
</tr>
<tr>
<td>11</td>
<td>North SWE</td>
<td>17</td>
<td>FPS, RPG, MMORPG, MOBA</td>
</tr>
<tr>
<td>12</td>
<td>South SWE</td>
<td>4</td>
<td>FPS, MMORPG, Open World/Sandbox</td>
</tr>
<tr>
<td>13</td>
<td>Finland</td>
<td>30</td>
<td>FPS, MOBA, Strategy</td>
</tr>
<tr>
<td>14</td>
<td>South SWE</td>
<td>3</td>
<td>FPS, Single player story driven, Puzzle</td>
</tr>
<tr>
<td>15</td>
<td>North SWE</td>
<td>18</td>
<td>FPS</td>
</tr>
<tr>
<td>16</td>
<td>South SWE</td>
<td>8</td>
<td>Single player story driven</td>
</tr>
<tr>
<td>17</td>
<td>South SWE</td>
<td>1</td>
<td>FPS</td>
</tr>
<tr>
<td>18</td>
<td>Russia</td>
<td>6</td>
<td>FPS, Indie</td>
</tr>
<tr>
<td>19</td>
<td>North SWE</td>
<td>5</td>
<td>Mobile</td>
</tr>
<tr>
<td>20</td>
<td>South SWE</td>
<td>4</td>
<td>FPS, Action RPG, Action</td>
</tr>
<tr>
<td>21</td>
<td>South SWE</td>
<td>1</td>
<td>FPS, Strategy</td>
</tr>
<tr>
<td>22</td>
<td>South SWE</td>
<td>1</td>
<td>Racing, RPG</td>
</tr>
<tr>
<td>23</td>
<td>South SWE</td>
<td>4</td>
<td>FPS, RTS</td>
</tr>
</tbody>
</table>
Table 3 Questions 4-6 Subject 3 (in bold) excluded from analysis

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Wii U, 3DS, Xbox One</td>
<td>No</td>
<td>0 years</td>
</tr>
<tr>
<td>2</td>
<td>PC, Mobile, N64</td>
<td>Yes</td>
<td>10 years</td>
</tr>
<tr>
<td>3 Mobile</td>
<td>Yes</td>
<td>12 years</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Wii, PS2</td>
<td>Yes</td>
<td>11 years</td>
</tr>
<tr>
<td>5</td>
<td>PC, PS4</td>
<td>Yes</td>
<td>&gt;10 years</td>
</tr>
<tr>
<td>6</td>
<td>PC, PS2, Gameboy Color/Advance SP</td>
<td>Yes</td>
<td>2-3 years (Bass), 8 years (Vocals)</td>
</tr>
<tr>
<td>7</td>
<td>PC</td>
<td>Yes</td>
<td>8 years</td>
</tr>
<tr>
<td>8</td>
<td>3DS, Gamecube, Xbox 360, PS4</td>
<td>No</td>
<td>0 years</td>
</tr>
<tr>
<td>9</td>
<td>PC</td>
<td>Yes</td>
<td>20 years</td>
</tr>
<tr>
<td>10</td>
<td>PS4, PS Vita</td>
<td>Yes</td>
<td>13 years</td>
</tr>
<tr>
<td>11</td>
<td>PC, OSX</td>
<td>Yes</td>
<td>17 years</td>
</tr>
<tr>
<td>12</td>
<td>PC, OSX, PS3</td>
<td>Yes</td>
<td>10 years</td>
</tr>
<tr>
<td>13</td>
<td>PC, OSX</td>
<td>Yes</td>
<td>16 years</td>
</tr>
<tr>
<td>14</td>
<td>PC, Xbox</td>
<td>Yes</td>
<td>15 years</td>
</tr>
<tr>
<td>15</td>
<td>PC</td>
<td>Yes</td>
<td>5 years</td>
</tr>
<tr>
<td>16</td>
<td>PS4</td>
<td>Yes</td>
<td>10 years</td>
</tr>
<tr>
<td>17</td>
<td>PS4, Xbox</td>
<td>Yes</td>
<td>7 years</td>
</tr>
<tr>
<td>18</td>
<td>PC</td>
<td>Yes</td>
<td>14 years</td>
</tr>
<tr>
<td>19</td>
<td>Mobile</td>
<td>Yes</td>
<td>21 years</td>
</tr>
<tr>
<td>20</td>
<td>PS4, PS3, NES, SNES</td>
<td>Yes</td>
<td>10-12 years</td>
</tr>
<tr>
<td>21</td>
<td>PC, Gamecube</td>
<td>Yes</td>
<td>8 years</td>
</tr>
<tr>
<td>22</td>
<td>PC</td>
<td>Yes</td>
<td>12 years</td>
</tr>
<tr>
<td>23</td>
<td>PC</td>
<td>Yes</td>
<td>12 years</td>
</tr>
</tbody>
</table>
Table 4 Question 7 Subject 3 (in bold) excluded from analysis

<table>
<thead>
<tr>
<th>Subject</th>
<th>7. Personal electronic devises (sound/no sound)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>2</td>
<td>Smartphone (no sound), laptop (no sound)</td>
</tr>
<tr>
<td>3</td>
<td>Smartphone (no sound), laptop (no sound)</td>
</tr>
<tr>
<td>4</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>5</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>6</td>
<td>Smartphone (sound), laptop (sound)</td>
</tr>
<tr>
<td>7</td>
<td>Smartphone (sound), laptop (sound)</td>
</tr>
<tr>
<td>8</td>
<td>Smartphone (Sound)</td>
</tr>
<tr>
<td>9</td>
<td>Smartphone (no sound)</td>
</tr>
<tr>
<td>10</td>
<td>Smartphone (no sound), tablet (sound), laptop (sound)</td>
</tr>
<tr>
<td>11</td>
<td>Laptop (no sound)</td>
</tr>
<tr>
<td>12</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>13</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>14</td>
<td>Smartphone (sound), laptop (sound)</td>
</tr>
<tr>
<td>15</td>
<td>Smartphone (no sound), laptop (sound), Smartwatch (no sound)</td>
</tr>
<tr>
<td>16</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>17</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>18</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>19</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>20</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>21</td>
<td>Smartphone (sound), laptop (sound)</td>
</tr>
<tr>
<td>22</td>
<td>Smartphone (no sound), laptop (sound)</td>
</tr>
<tr>
<td>23</td>
<td>Smartphone (no sound), laptop (no sound), tablet (no sound)</td>
</tr>
</tbody>
</table>
Question 8 full responses
All of the answers were translated from Swedish to English by the author.

Subject 1
“I turn off the sounds if they are annoying or unnecessary, but on my laptop I think that sounds telling me what’s going on, can be helpful.”

Subject 2
“It’s useful that on my smartphone, different applications have different sounds. That way I can know what type of notification I’ve gotten before I even look at my smartphone.”

Subject 3
Not applicable, turns off all notification type sounds.

Subject 4
“The sounds on my laptop usually notify me when something happens. For example, error messages, when a download has completed etc. The sounds help to redirect my attention when I’m focusing on something else.”

Subject 5
“My laptop gives me clear/distinct notifications for E-mail and chat messages, so that I quickly check what is going on. I feel that I get a better and quicker communication with the sound turned on.”

Subject 6
“It is easier to notice text messages and phone calls if the smartphones notification type sounds are on. Computer, TV and Gameboy are use way better with the sound on (video/film without sound is just moving pictures which makes it so that the context/information is lost, notification sounds/messages and warning sounds/messages are more easily noticeable with the sound on, it’s possible to play games without sound but they can be harder to complete). The sounds of the devices, with the exception of the Gameboy, rarely disturb people around you, which decreases the incentive to turn them off. Gameboy games can manage without sound, since they are very simple, but this adds an extra dimension to the games.”

Subject 7
“Yes, I notice, for example, notifications easier.”

Subject 8
“Guiding, they help make the sorting of information easier, for example Aftonbladet (Swedish newspaper/application) has one sound while Twitter has another.”

Subject 9
Not applicable, turns off all notification type sounds.

Subject 10
“Yes, you get feedback that you’ve done something.”

Subject 11
Not applicable, turns off all notification type sounds.

Subject 12
“Absolutely, you get a confirmation that an E-mail has been sent through sound, error messages, you understand that things happen because you get a confirmation through the sound.”

Subject 13
Not applicable, turns off all notification type sounds.
Subject 14
“Yes, I feel connected and my ability to, for example, answer E-mails or text messages is faster. Besides that, notifications remind me about things I’ve forgotten.”

Subject 15
“Notifications are easier to notice.”

Subject 16
“On my computer, there was previously a sound that indicated how loud the volume from the computer was, this made it easy to keep track of approximately how loud a sound was going to be before you played it. After they removed this sound, I now have to lower the volume before I play something to make sure that I don’t blow my ears out.”

Subject 17
“Partly helps to inform the user about what type of notification you have received, for example E-mail or text message.”

Subject 18
“For example on a laptop (windows) it’s easier to understand if you click the wrong thing or on a specific area.”

Subject 19
Not applicable, turns off all notification type sounds.

Subject 20
“Some specific system sounds (low battery level, phone call) helps me. More “information-less” sounds like button clicks doesn’t help me that much.”

Subject 21
Not applicable, turns off all notification type sounds.

Subject 22
“Don’t like background sounds like clicks (when pressing buttons), so I tend to avoid those.”

Subject 23
Not applicable, turns off all notification type sounds.
Table 5 Question 8 themes Subject 3 (in bold) excluded from analysis

<table>
<thead>
<tr>
<th>Subject</th>
<th>Easier to notice notifications</th>
<th>Instant info about type of notification</th>
<th>Simple feedback increases usability/ease of use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>6</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>7</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>