Which compound-earcon’s attributes may improve a player’s performance in a search-oriented gameplay: rhythm vs timbre?

Anton Savvateev

Ljudteknik, kandidat
2018

Luleå tekniska universitet
Institutionen för konst, kommunikation och lärande
Abstract

Earcons are commonly used by sound designers in order to support visual cues in a game and to make a gaming experience more enjoyable. This study covers two earcons’ attributes: rhythm and timbre. Rhythm and timbre attributes were chosen according to the earcon sound design guidelines from the previous studies. An experiment in a form of a video game was conducted in order to research whether one of the conditions can increase a players’ performance. A subject had to choose the correct key to the door in order to go to the next location. There were 3 different locations and there were totally 5 different own-designed earcons: 1 incorrect earcon in the both conditions and 2 different correct earcons in each condition. 20 subjects with various gaming experience from the Luleå University of Technology participated in the experiment. The subjects were randomly divided into two groups with different conditions: rhythm and timbre. The amount of wrong trials and completion time were analyzed for each condition and the results were given with the help of Mann-Whitney U-test and t-test calculations. The results of U-test showed that there was a significant difference between two groups in terms of the wrong trials amount. Group with rhythm condition showed better performance in terms of the wrong trials amount. The t-test showed a significant difference between the two groups in terms of completion time. Group with timbre condition showed better timing performance, although considering the analysis it did not increase their performance in terms of making correct choices. Further research might be recommended on comparing various earcon attributes in different ecologically valid scenarios.
Acknowledgements

I wish to thank my teachers at Luleå University of Technology and especially my supervisor Nyssim Lefford for her invaluable guidance and great support.

I also wish to thank all the participants of this work.
Table of Contents

Introduction ......................................................................................................................... 4
  Background ......................................................................................................................... 4
  Auditory attributes of earcons .......................................................................................... 6
  Earcons in digital games ................................................................................................. 7
  Survey of earcons in digital games .................................................................................... 7
  Earcon functions in digital games .................................................................................... 9
  Analysis of earcons in digital games ................................................................................ 10
  Motivation of the research question ................................................................................. 11

Method ............................................................................................................................... 12
  Pilot experiments .............................................................................................................. 12
  Game mechanics description ............................................................................................ 13
  Subjects ............................................................................................................................. 17
  Stimuli ............................................................................................................................... 17
  Equipment ........................................................................................................................ 20
  Experiment setup ............................................................................................................. 21
  Collecting data .................................................................................................................. 21

Results ................................................................................................................................ 23
  Amount of wrong trials, U-test ....................................................................................... 24
  Completion time, t-test ..................................................................................................... 25
  Results from the questionnaires ...................................................................................... 26

Discussion .......................................................................................................................... 30
  Amount of wrong trials ................................................................................................... 30
  Completion time .............................................................................................................. 30
  Experienced game difficulty ......................................................................................... 31
  Gaming experience ......................................................................................................... 31
  Studying program ............................................................................................................ 32
  The game strategies and player behavior ...................................................................... 32
  Experiment findings ........................................................................................................ 33
  Limitations of the experiment ....................................................................................... 34
  Further research ............................................................................................................. 35

References .......................................................................................................................... 36

Appendix ............................................................................................................................. 37
Introduction

The interactive game industry is growing rapidly, and more independent game studios and companies appear across the world. The audio is an important part of game development and gaming experience, therefore demand for great sound design and sound designers will grow as well. In order to create a great game audio, it requires to have comprehension on how the audio works in different game scenarios. One of the key-elements in a game audio is sound effects which quite often have the function to help players, lead them to a goal, focus their attention on important game features. Unfortunately, this function of sound effects is quite often underestimated in a game scenario, that is why it is essential to make more researches on the supportive sound design functions. One solution is to create one of the auditory display representations – earcons, which might increase a player’s performance. In this study two different earcon attributes were studied: rhythm and timbre. It is based on the earcon design guidelines, which were reviewed in this section of the study.

Background

Auditory display usually utilizes three different representations: auditory icons (which are natural sounds, for example a sound of paper, which is triggered by deleting a file in computer operating system), earcons (synthetic and abstract sounds) and speech. Blattner defined Earcons as: “non-verbal audio messages used in the user-computer interface to provide information to the user about some computer object, operation, or interaction”. (McGookin & Brewster, 2011, p.339). Brewster further refined this definition as: “abstract, synthetic tones that can be used in structured combinations to create auditory messages” (as cited in McGookin & Brewster, 2011). McGookin and Brewster (2011) described earcons as “they can be thought of as short, structured musical messages, where different musical properties of sound are associated with different parameters of the data being communicated” (p.339).

Blattner, Sumikawa and Greenberg proposed different ways to form families of earcons: one-element earcons, compound earcons, transformational earcons and hierarchical earcons.

1) **One-element earcons** usually are simple-pitch sounds, which cannot be decomposed, unlike the other types of earcons.

2) **Compound earcons** are formed by bind of one-element earcons together in order to create more advanced communication. “In many ways they are analogous to forming
a phrase out of words, where one-element earcons represent words and compound earcons represent phrases.” (McGookin & Brewster, 2011, p.341).

3) **Transformational earcons** “are constructed around a specific set of rules where there exists a consistent set of structured symbolic mappings from individual data parameters (such as file type) to individual sound attributes (such as timbre). Specific values of data parameters (e.g., a paint file) are then mapped to specific values of the corresponding auditory attribute (e.g., a piano timbre)” (McGookin & Brewster, 2011, p.341).

4) **Hierarchical earcons** consist of several levels with different nodes. As well as transformational earcons, hierarchical earcons constructed around a set of rules. “Each earcon is a node in a tree and each node inherits all of the properties of the nodes above it in the tree.” (McGookin & Brewster, 2011, p.342).

It might seem that the level of earcon complexity is a main factor of learnability and recognition. However, it does not necessarily mean that one-element earcons are much easier to learn than more complex types of earcons. McGookin and Brewster concluded that well designed earcons might be learned during a quite short period of time (5-10 minutes), for example during a pilot experiment or a training phase. Such short training phase could rise earcon recognition up to 80% (McGookin & Brewster, 2011). The similar conclusion was made in another study by Dingler, Lindsay and Walker (2008). Evidently, to conduct a training phase is not everything that is needed. In order to understand and to design well-working earcons it requires to study earcons’ attributes. Brewster, Wright and Edwards described seven different earcons’ auditory attributes with a guideline on their use in earcon design: timbre, register, pitch, rhythm, intensity, spatial location, timing (McGookin & Brewster, 2011). According to the guideline the best working attributes together (within an earcon) are rhythm, pitch/timbre and melody. The recommendation is to combine rhythm and timbre/pitch in order to provide greater difference between earcons. The greater difference might help players recognize different earcons.

Following this guideline, it was decided to focus on compound earcons due to their common use in popular games, as well if one of the earcon attributes (rhythm or timbre) can increase a player’s performance. It was decided not to include melody as researched stimuli due to study’s focus on sound design. It is out of the study’s scope as researching melody aims rather to musicology than to sound effects.
Auditory attributes of earcons

In order to have better understanding of an earcons` construction, further characteristics of earcons have been overviewed. Brewster, Wright and Edwards described seven different earcons` auditory attributes with a guideline on their use (as cited in McGookin & Brewster, 2011).

Seven auditory attributes and guidelines on their use in earcon design (McGookin & Brewster, 2011):

- **Timbre.** According to Brewster, musical timbre is preferable than simple sinusoidal tones. In case two earcons concurrently present, it is recommended to use different instruments from the same music family.

- **Register.** In order to achieve absolute judgements from listener, it is recommended not to use register on its own to differentiate earcons. It is better to use large differences such as two or three in octaves within the used register.

- **Pitch.** Due to the factor that it is hard to distinguish two different earcons that have only pitch difference, it is recommended to combine complex pitch structure with other parameters. Suggested maximum pitch should not be higher than 5 kHz and no lower than 125 Hz - 150 Hz.

- **Rhythm.** Rhythms should be as different as possible in order to achieve better perceived difference between earcons. It is suggested to combine rhythm with pitch for better distinguishes between earcons.

- **Intensity.** It is recommended to avoid the use of intensity on its own as it is hard for listeners to make judgements which are based only on this parameter.

- **Spatial Location.** The difference between spatial location is recommended when earcons from different families are used. Spatialisation can also be used to improve identification when earcons from the same family group are concurrently presented.

- **Timing.** It is suggested to use short gaps between earcons and at least 0.1 second should be used between the end and the start, when two or more compound earcons are used. It is suggested to use at least 300 ms in case of concurrently presented earcons.
Earcons in digital games

Earcons are widely used in digital games and they can be used as a support tool in order to complete different tasks in games. Earcons can provide a player with various useful information, give some hints or simply alert about danger or reducing health points. Earcons’ learnability and recognition can vary depending on conditions, i.e. depending on a game genre, game task and information that an earcon might provide to a player.

In relation to informative sound design in games, earcons allow game developers the ability to convey complex sound patterns. However, the interpretations should be learned, and the encodings may rely somewhat on both the designer and players’ musical listening skills. In video games, earcons can be used to provide detailed information to players, for example in the form of warning signals (Ng & Nesbitt, 2013).

In order to describe earcons in games and their use, it is necessary to give an explanation of earcons within the game context. It is possible to define earcons in games based on the definitions, which were previously presented in this study. Earcons in games are short, non-verbal structuralized audio messages, which are used to provide information to a player about different events and interactions in a game.

Survey of earcons in digital games

Different practical examples of earcons in different game genres are described in this part of the study. It might be reasonable to begin earcon observation from one of the most popular game genre – FPS (first-person shooter). Half-Life 2 (2004) is taken as an example of a popular FPS-game. There is an example of earcon use in Half-Life 2, when a player presses the action button (default setting is the “E” button on keyboard). If a player aims on an object, which is usable and occurs some type of interaction, and presses the interaction “E”-button, the earcon would be played. The earcon in this case can be described as a very short and fast combination of identically clicking signals. Another earcon plays in case a player aims on an unusable object or somewhere else in the game and presses the interaction “E”-button. The earcon in this case can be described as a single short beeping sound. In the Half-Life 2 example the described earcons provide information to a player whether an object is interactive and important for completing the game or not. The earcons can be easily learned
by a player and therefore they are additional and important helping aids throughout the whole game.

DOOM (id Software, 2016) and Quake 3 Arena (id Software, 1999) are other examples of popular FPS games. In DOOM, notification sounds occur when there are few seconds left of a time limited ability, which was picked up by player. It was presented the similar way in Quake 3 Arena. Sounds are synthetic-type in both examples and synchronized to seconds timing. In Quake 3 Arena it is also synchronized to a visual timer. Both examples give to a player information about the amount of time that left of an ability, which was picked up. Often a player does not have much time to focus on the visual interface during a gameplay, for example during an intense battle. The earcons help to inform a player without dragging his attention from a battle.

Ng and Nesbitt (2013) describe an example from Battlefield 3 game (EA Dice, 2011). In this example they describe different sounds of alarm in the game, which informs player about condition of the tank armor. Increasing rate and frequency of alarm indicates how much armor is left. This example can be supplemented further: another type of alarm sounds with different rate and frequency play in case an enemy aims a missile towards a player in the tank.

Some examples can be taken from strategy games. One of the most popular RTS game is Starcraft 2: Legacy of the Void (Blizzard Entertainment, 2015). An example of earcon in this game is a short beep type of sound that plays in case if it is not possible to target a unit or a building. In this example the earcon has a function to inform a player that an action is impossible to perform. Starcraft 2 franchise is a very popular online competitive type of a game where the factor of APM (performance of actions per minute) is extremely important. Considering this, the additional support provided by described earcon is a great help for players.

Earcons can have other functions in strategy games. As an example, a popular turn-based strategy Heroes of Might and Magic 3 (New World Computing, 1999) can be taken. In this game different sound cues play depending on which type of resource a player collects. Different minerals have different short sounds, but all of them are quite similar to each other. Another type of a sound cue occurs if a player collects a chest, in which might be gold or experience points. In this case triggers a non-verbal audio message which consists of a string
instrument passage. The examples in Heroes 3 are mainly focused to help a player to distinguish different types of objects with the help of provided earcons.

Murdered: Soul Suspect (Airtight Games, 2014) is a third-person view action-adventure game. There are many different types of investigations during the whole gameplay. To complete an investigation a player has to collect important clues and evidences. Sometimes a player has to choose a correct clue from various presented ones. If a player chooses the right one, then a short violin synth-type audio message would be played. There is a short drum sound cue, which is played simultaneously with synth-pad and a violin. In case a player chooses the wrong one, then the earcon consisted of a short single synth sound would be played.

However, not all information collected by a player is considered as main clues. A player might collect some additional information, which is not aimed on completing the game, but gives additional details about the story of the world in the game. If a player collects such type of information, a short sound with a synth and a drum cue would be played. This earcon is similar in some way to the “right clue” earcon due to consistence of synths and the similar drum sound cue. But there is a noticeable difference between these two earcons mainly in a type of melody played by a synth cue. The sonic difference between these two earcons will be described more detailed further in this study. The earcons in Murdered: Soul Suspect helps a player to distinguish different types of clues, their importance level for completing the game, as well as if the choice in a puzzle was correct or not. These earcons are easy to learn and they are consistent during the whole gameplay, therefore, they have function to provide a player with additional help.

**Earcon functions in digital games**

Earcons can be classified based on their function which they carry. Earcons from the observed examples in games have mainly informative functions, but the importance of information varies. In some cases, earcons help a player to make a fast decision, which might be crucial in a matter of a characters’ life in the game (Battlefield 3, DOOM, Quake 3 Arena). In case of Starcraft 2 earcons help to avoid some possible mistakes, which can be taken in multi-tasking APM-based scenarios. In Half-Life 2 the described earcon supports a player by indicating if an object is usable or not, which helps in search-type assignments in the game.
In a similar way act the earcon examples in Murdered: Soul Suspect. They provide information whether something was done right or wrong as well as if something belongs to the main game-quest. Finally, the earcon examples in Heroes of Might and Magic 3 help a player to distinguish different types of items collected in the game world. In this case earcons support the visual part and don’t impact neither on players decision nor a character’s life and abilities. They act as an additional non-crucial indicator of an occurred event.

**Analysis of earcons in digital games**

It is necessary to consider sonic parameters and auditory attributes in earcons in order to analyze examples described previously in this study. According to Brewster, Wright and Edwards the auditory attributes are timbre, register, pitch, rhythm, intensity, spatial location and timing (as cited in McGookin & Brewster, 2011). Besides, the Blattner, Sumikawa and Greenberg’s classification of earcon families is considered during analysis of the earcon examples. The classification includes one-element earcons, compound earcons, transformational earcons and hierarchical earcons (as cited in McGookin & Brewster, 2011). It is important to mention that no parameters and auditory attributes were measured in this study and therefore further studies with the data of measurements required for a complete analysis.

Presented earcon examples in Half-Life 2 and Starcraft 2: Legacy of the Void are one-element earcons. However, the earcon which is triggered by interacting with a usable object in Half-Life 2 consists of several identical fast played sounds within very short amount of time. The whole earcon is less than 1 second long. Such short gaps make the perception of the earcon identification as a one-element earcon. The main difference between the two various earcons in Half-Life 2 are in pitch and rhythm. If an object is non-interactive a single beep type of a sound would be played. The gaps-timing between sounds might be experienced, however, due to the short amount of time between them it is still perceived as a single sound, as it was mentioned above. The examples of earcons in DOOM, Quake 3 Arena and Battlefield 3 are similar in their structure. The earcons in these examples are one-element earcons, that signal about an armor or an ability which is about to end or alert about an incoming missile as in the case of Battlefield 3. With the increasing level of danger increases the rhythm and the timing between these earcons decreases. The same principle is used with
the decreasing level of armor, except for the time synchronized earcons with the visual timer in Quake 3 Arena.

The described earcons in Heroes of Might and Magic 3 have similar frequencies (except for the chest example), therefore they might be classified as more complex and structured. However, more examples from this game should be investigated in order to define them in a correct way. The frequencies changes within the same register. The earcon which is triggered by opening a chest consists of different sound cues, which might be classified as a compound earcon. The difference in sound cues within this earcon are in timbre, pitch and rhythm. In this particular example the earcon provides a non-crucial type of gameplay interaction.

The two “success” earcons in Murdered: Soul Suspect can be classified as hierarchical ones due to the same drum sound cue. The synth sounds change pitch, as well as changes the intensity of drum sound cue. However, the drum sound cue exists in both examples, which can be an argument to classify these earcons as hierarchical. The example of earcon which triggers by a wrong choice decision can be classified as a one-element earcon. It is similar to a Half-Life 2 issue of experiencing more than one sound may occur, though it can still be perceived as a one-element earcon due to the short timing between different frequencies. The sound cue consists of two single synth-pad sounds, which slightly changes in pitch and timbre during playback of the earcon (approximately 2 seconds long).

Based on the conducted analysis it might be concluded that pitch, rhythm and timing are the most usable auditory attributes. It is also partly described by Brewster, Wright and Edward’s guideline on auditory attributes and how they should be used (as cited in McGookin & Brewster, 2011).

**Motivation of the research question**

The aim of the bachelor is to research on which type of earcon attributes (rhythm vs timbre) might increase a player’s performance in an ecologically valid game scenario. The purpose of the research is to provide possible help to sound designers in their work with digital games. Moreover, it might be useful even for sound designers who work with various auditory displays, for example different types of software, or even for those who work with industrial sound design. It makes sound design processing more careful and therefore deeper knowledge about earcons and earcon attributes design can be essential for sound designers.
Method

The experiment was decided to make in a form of a digital game where a subject had to complete the level by listening and choosing different earcons. The experiment game was created by using Unreal Engine 4.17.2 (Epic Games, 2017). The engine’s first-person template was used as well as free distributed packages: Infinity Blade: Fire Lands assets (Epic Games, 2015) and Game Textures Material Pack (Epic Games, 2014).

The game level consisted of three rooms with keys, platforms and doors in each room. It was required to choose a correct key in each room to be able to pass the doors. The keys looked alike therefore it was only an earcon cue which differed between them. Each earcon was connected to a key, which a player could hear by interacting with a platform on the ground. Each room had several keys: 4 in the first one and 8 in the second and the third rooms. First room had 3 incorrect earcons/keys and 1 correct, the second and the third rooms had 7 incorrect and 1 correct key/earcon. One of the incorrect earcons in the second room was the same earcon which was correct in the first room. The same principle was in the third room, where one of the incorrect earcons was the one which was correct in the second room. It can be described by a simple scheme:

- **The First room**: 3 incorrect and 1 correct earcon (E1 – earcon 1) = 4 earcons.
- **The Second room**: 7 incorrect (one of them is E1, which became incorrect in the second room) and 1 correct (E2 – earcon 2) = 8 earcons.
- **The Third room**: 7 incorrect (one of them is E2, which became incorrect in the third room) and 1 correct (E1 – earcon 1) = 8 earcons.

The first and the third doors looked the same while the second was noticeably different which was made in order a player could understand the game logic (more detailed description of the doors and game mechanics will be covered further below in this section). The game was finished after subject passed the third door. Fading out screen and a drum sound cue helped a player to notice that the game was over.

Pilot experiments

Seven experiments were conducted with 7 subjects: four in rhythm and three in timbre condition. The pilots showed that subjects in both conditions could distinguish different
earcons. However, the main problem was to provide some hints to understand the door logic. The second and the first/third doors had different size and visual pattern during the pilot experiments. Besides, the correct earcon played at the moment when a player picked up the correct key and the same earcon played at the moment a player passed the correct door.

To help the players to understand the logic it was decided to implement different material design on the oval figures which were on the doors. The first and the third doors’ ovals were changed to golden material and the second door oval was changed to silver material. Besides, the green/red light was removed from the second door to make the doors look even more different. Additionally, the spotlights were implemented to spot on all doors in order to give player a hint that visual cues could help.

Another comment which was received after the pilots was to adjust the letter font so that it would be easier to read it and not to miss important descriptions. Font adjustments can be seen at Figure 2.1.

**Game mechanics description**

Game controls were attached to the following buttons: “W”, “A”, “S”, “D” to walk, “E” to interact with platforms, keys and doors. Mouse was used to look around. By holding “I” button a subject could read the letter, which had an introduction to the game and descriptions how to interact with the objects in the game.

![Figure 2.1 The letter with instructions](image)
Room 1

Figure 2.2 Room 1. Game start position

A subject spawned in the first room, which had four keys, four platforms and a door. The room was quite dark and looked dungeon alike. Spotlights were added to the platforms, keys and doors in order to help a player to navigate and focus attention on the interactive objects. There were two different earcons: a correct and an incorrect one. Three keys and platforms had incorrect earcon and one key had the correct one. In order to listen to an earcon a subject could trigger it by pressing “E” button while standing on a platform near a stairway which led to a key. The amount of interactions with platforms was unlimited.

Figure 2.3 An interactive platform

Respawn was triggered in case a subject collected the wrong key, and a short text “Wrong!” popped up on the screen. Player had to start the game from the beginning in the first room, no matter in which room the wrong key was collected.
The correct earcon sound was triggered when a player picked up the correct key. The same earcon played when the first door opened. After passing the first door a short letter appeared to the player and it was impossible to go back to the first room. “Now it is getting a bit more complicated… Think carefully before you make your choice!”

Figure 2.4 Room 1, the first door

Room 2

The second room was larger than the first room and there were 8 keys and platforms. The second door was also noticeably larger than the first one and had different type of oval material – silver, and another type of pattern on the door.

Figure 2.5 Room 2, eight keys
The third earcon was added in the second room. This earcon was the correct one in the second room. There were three types of earcons: incorrect one, the correct earcon from the first room (became incorrect one as well as there was a different type of door) and the new correct earcon.

The correct earcon played when the second door opened. After passing the second door a short letter appeared to the player and it was impossible to go back to the previous room. “You are on the right way! Now you have to think more carefully…”

**Room 3**

The third room had the same principle as the second one with the only difference that the correct earcon was the same as in the first room. The doors in the first and in the third rooms were similar as well. Player activated the game end-trigger after passing the third door.

The difference between doors (as well as the fact that door 1 and door 3 were similar) was made in order to help player with an additional guideline.
Subjects

The subjects were 20 students from the Luleå University of Technology. 15 students were from audio engineering program, 4 students were from media program and the other one from musician program. In order to participate in the experiment a subject required to have some experience in video gaming and to have normal hearing. Subjects were randomly divided into two groups. Group A participated in the rhythm condition and Group B in the timbre condition. One subject could participate only in one condition.

Stimuli

Five earcons were designed for the experiment. Each earcon consisted of the same processed melodic harp and processed percussive shaker audio cues. It was important to make fair balance between different earcons’ attributes in both groups. It was made in order to avoid possible larger difference between earcons in one group which might lead to unfair comparison due to larger difference in one of the groups. The shaker was recorded with Shure SM58 microphone. The melodic cue remained the same for all the keys in each condition. It was made in order to make it more realistic to the game scenario as well as to make earcons
sound more interesting in terms of sound design. Besides, the goal was to not make earcons very easy to distinguish between each other. The earcons can be divided here according to the experiment conditions.

**Group A (Rhythm) stimuli**

Incorrect earcon: melodic and percussive cues.

*Figure 2.8 Screenshot from Pro Tools 12.17.12 DAW. Processed melodic cue*

*Figure 2.9 Screenshot from Pro Tools 12.17.12 DAW. Processed percussive cue*

Correct earcon in the 1st room: the percussive cue was slower than the same cue in the incorrect earcon.

*Figure 2.10 Screenshot from Pro Tools 12.17.12 DAW. Processed percussive cue, 5 accented sounds*

Correct earcon in the 2nd room: the percussive cue was much slower than previous two.

*Figure 2.11 Screenshot from Pro Tools 12.17.12 DAW. Processed percussive cue, 3 accented sounds*

Correct earcon in the 3rd room: the earcon was the same as in the 1st room.
Group B (Timbre) stimuli

**Incorrect earcon:** melodic and percussive cues (same as in the Group A, see Figure 2.8 and 2.9).

![Figure 2.12 Screenshot from Pro Tools 12.17.12, Waves PAZ-Frequency plugin. Processed percussive cue, around 500 – 8000 Hz.](image)

**Correct earcon in the 1st room:** the percussive cue had higher frequencies than the cue in the incorrect earcon.

![Figure 2.13 Screenshot from Pro Tools 12.17.12, Waves PAZ-Frequency plugin. Processed percussive cue, around 2300 – 16000 Hz.](image)

**Correct earcon in the 2nd room:** the percussive cue was much slower than the previous two.
Correct earcon in the 3rd room: the earcon was the same as in the 1st room.

Other type of audio in the game: ambience sounds, foley, burning fire from the torches and opening/closing door sounds were added to the experiment level in order to make the game more ecological valid. A drum sound cue played when player passed the final door.

Equipment

- Microsoft Windows 10, 64-bit Edu OS. Intel Core i5-3470, 3.20 GHz CPU, 8 GB RAM
- Focusrite Scarlett 18i20 audio interface
- AKG K240 studio headphones
- Dell U2412M, 24” screen
- Asus AMD Radeon R9 200 series graphic card
- Apowersoft Ace thinker screen recorder
- Shure SM58 microphone
- Unreal Engine 4.17.2
- Epic Games plugins. Infinity Blade: Fire Lands assets, Game Textures Material Pack
- Pro Tools 12.17.12 DAW
- Avid plugins: Pitch II
- Waves plugins: Butch Vig Vocals, CLA Drums, Sound Shifter, PAZ-frequency
- Sound Toys plugins: Little plate, Crystallizer

Figure 2.14 Screenshot from Pro Tools 12.17.12, Waves PAZ-Frequency plugin. Processed percussive cue, around 1200 – 12000 Hz
**Experiment setup**

Author’s apartment in Piteå was used as location for experiments. Headphones were chosen in order to eliminate possible disturbing sounds in the building. It is also common that people use headphones when they play games, besides the choice of using headphones might help subjects to focus on the game and especially on the game audio.

All 20 participants had the same audio volume set up at -15 dB in the Focusrite control application. All other applications were disabled during experiments to eliminate risk of possible disturbing notifications.

**Collecting data**

All subjects were randomly divided into two groups, A (rhythm) and B (timbre). Key positions were changed manually in the game engine. There were totally two setups for key positions, therefore ten subjects participated in one setup and ten in another one. Important to notice that the key positions did not change during same experiment, i.e. in case a subject made an incorrect choice and he or she was respawned to the beginning of the game.

**Quantitative data**

The amount of wrong trials (wrong key choices) and the total game completion time were collected. It was made with the help of Apowersoft Ace thinker screen recorder, which was used to record gameplay of each subject for further analysis. Completion time was calculated afterwards by video files analysis. The start point was set to a short intro sound, which played directly after the game started. The end point was set to the moment when the light on the third door changed to green.

The subjects did **not** know that gameplay was recorded, as well as the time completion data was collected. It was decided that way in order to avoid potential risks of random guessing.

After experiment a subject was asked to answer anonymously several questions. One of the questions was to rate the game difficulty from 1 (very easy) to 5 (very difficult).
Qualitative data

In the same questionnaire it was asked to answer the following questions:

- Describe your strategy choosing keys
- Describe any differences you heard in the game sounds
- How much did the sounds help you to solve the tasks?

Demographic data

The following questions were included in the questionnaire to collect demographic data:

- How many hours do you usually play games per week (in average)?
- Which type of games do you usually play?
- What program do you study?

There were several options to answer these questions, except for the question about studying program (see Appendix). It is important to notice that it was decided not to collect data about age and gender, as gaming experience and potential experience in music/audio engineering was in the focus.
Results

One subject was eliminated from the experiment due to lack of the game mechanics comprehension and random guessing during almost the whole game without listening to the sounds. The subject was replaced with another one later on. Table 3.1 shows the data for both groups.

<table>
<thead>
<tr>
<th>Subjects Group A (Rhythm)</th>
<th>Wrong trials</th>
<th>Time (s)</th>
<th>Subjects Group B (Timbre)</th>
<th>Wrong trials</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1</td>
<td>501</td>
<td>S11</td>
<td>2</td>
<td>506</td>
</tr>
<tr>
<td>S2</td>
<td>7</td>
<td>689</td>
<td>S12</td>
<td>2</td>
<td>438</td>
</tr>
<tr>
<td>S3</td>
<td>1</td>
<td>371</td>
<td>S13</td>
<td>9</td>
<td>563</td>
</tr>
<tr>
<td>S4</td>
<td>1</td>
<td>380</td>
<td>S14</td>
<td>1</td>
<td>427</td>
</tr>
<tr>
<td>S5</td>
<td>1</td>
<td>484</td>
<td>S15</td>
<td>0</td>
<td>305</td>
</tr>
<tr>
<td>S6</td>
<td>2</td>
<td>502</td>
<td>S16</td>
<td>3</td>
<td>574</td>
</tr>
<tr>
<td>S7</td>
<td>1</td>
<td>500</td>
<td>S17</td>
<td>2</td>
<td>385</td>
</tr>
<tr>
<td>S8</td>
<td>1</td>
<td>558</td>
<td>S18</td>
<td>5</td>
<td>324</td>
</tr>
<tr>
<td>S9</td>
<td>1</td>
<td>622</td>
<td>S19</td>
<td>2</td>
<td>365</td>
</tr>
<tr>
<td>S10</td>
<td>1</td>
<td>503</td>
<td>S20</td>
<td>2</td>
<td>248</td>
</tr>
</tbody>
</table>

Table 3.1 Results: completion time and wrong trials in the both groups
Amount of wrong trials, U-test

Due to the difference in variance and skewed distribution (Fig. 3.1) it was decided to use a non-parametric test. Mann-Whitney U-test was used in order to check significant difference between two groups in terms of the wrong trials. It was calculated without the two outliers (S2 and S13). The results are presented in the box plot (Fig. 3.2). The Mann-Whitney test’s results showed the critical value of U at p < .05 is 17 (the U-value = 16.5), therefore there was a significant difference at p < .05 (Social Science Statistics 2018).

Figure 3.1 Box plot: amount of wrong trials in the both groups

Figure 3.2 Box plot (without outliers S2 and S13): amount of wrong trials in the both groups
Completion time, t-test

Data of completion time follow the criteria for a parametric test, therefore an unpaired two-tailed t-test was used in order to study whether one of the groups showed better performance in terms of the completion time. The calculations were made manually (without outliers S2 and S13) and then they were checked with the GraphPad Software (2018). The t-test’s results for the completion times showed $p = 0.0425$ ($t = 2.2047$, $df = 16$, $\alpha = 0.05$), which means that there was a significant difference between the two groups. The results are presented in the box plot (Fig. 3.3).

![Box plot](https://via.placeholder.com/150)

*Figure 3.3 Box plot (without outliers S2 and S13): completion time in the both groups*
Results from the questionnaires

The quantitative results from the questionnaires were put into graphics and charts in order to have better view on the collected data.

Group A (Rhythm)

Figure 3.4 presents the Group A answers on the question “Was the game easy/difficult?”

![Bar chart showing Group A answers on “Was the game easy/difficult?”](image)

Figure 3.4, Group A. Experienced game difficulty

Figure 3.5 presents the Group A answers to the question “How many hours do you usually play games per week (in average)?”

![Pie chart showing average gameplay hours per week](image)

Figure 3.5, Group A. Average amount of gameplay hours per week
Figure 3.6 presents the Group A answers to the question “Which type of games do you usually play?” It is important to notice that this question gave possibility to choose more than one option.

![Bar chart showing game genre preferences](chart1.png)

**Figure 3.6, Group A. Game genre preferences/experience**

Figure 3.7 presents the Group A subjects according to their studying program.

![Pie chart showing studying programs](chart2.png)

**Figure 3.7, Group A. The studying programs of the subjects**
Group B (Timbre)

*Figure 3.8* presents the Group B answers to the question *“Was the game easy/difficult?”*

![Figure 3.8, Group B. Experienced game difficulty](image)

*Figure 3.9* presents the Group B answers to the question *“How many hours do you usually play games per week (in average)?”*

![Figure 3.9, Group B. Average amount of gameplay hours per week](image)
Figure 3.10 presents the Group B answers to the question “Which type of games do you usually play?” It is important to notice that this question gave possibility to choose more than one option.

Figure 3.10, Group B. Game genre preferences/experience

Figure 3.11 presents the Group B subjects according to their studying program.

Figure 3.11, Group B. The studying programs of the subjects
Discussion

Amount of wrong trials

Mann-Whitney U-test was used in order to check significant difference between two groups in terms of the wrong trials (Fig 3.2). It was made without the results of S2 and S13, which were the outliers. S2 had 7 wrong trials and S13 had 9 wrong trials (Table 3.1). The Mann-Whitney test’s results showed the critical value of U at \( p < .05 \) is 17 (the U-value = 16.5), therefore there was a significant difference at \( p < .05 \) (Social Science Statistics 2018).

Why are the results of the Group A (Rhythm) showed better performance than Group B (Timbre)? There could be various reasons for such results. The stimuli’s design can be a reason for better performance results, i.e. the difference between the earcons in the Group A (Rhythm) was larger than the difference between the earcons in the Group B (Timbre). The demographic data did not show any critical difference between the two groups (demographic data will be discussed further in the text). It can be stated that in this experiment the rhythm group showed better performance.

Completion time

The t-test’s results for the completion times showed \( p = 0.0425 \) (\( t = 2.2047, \text{df} = 16, \alpha = 0.05 \)), which means that there was a significant difference between two groups. Group B (Timbre) showed better performance in terms of the completion time.

It is important to notice that the subjects did not know that the completion time was measured and nothing in the game indicated that it was time-based as well. It was made in order to avoid possible risks of random guessing. As the results show (Fig. 3.3), Group B (Timbre) showed better completion time results, although the completion time was not the main indicator as it was possible to complete the game with any possible completion time result. Moreover, the better completion time results might show that there was potential random guessing strategy. In the scenario with two different earcons (room 2 and 3) a subject might randomly choose one of the earcons and in case it was the wrong one he/she might choose the other one after respawning. Subjects’ description of their strategies as well as analysis of the recorded gameplay can be supportive evidence of such behavior.
On the other hand, a subject might spend more time listening to different earcons and make a correct choice from the first trial. Therefore, it can be stated that better completion time results do not indicate better performance in terms of the game logic and performance.

**Experienced game difficulty**

Another observation was made after analyzing the subjects’ experienced game difficulty. As it shows in the *Fig. 3.4* and *Fig. 3.8*, none of the subjects graded the game difficulty as “very difficult”. It can be explained by limited amount of different earcons, i.e. there were only two possible choices if a subject heard two earcons that differs from all the others. Additionally, there was no provided information to the subjects that the completion game time was recorded which potentially could decrease the expected performance demand from the subjects. As a result, it might be an explanation why the subjects did not rate the game as very difficult.

Five subjects in the Group A (Rhythm) rated game difficulty by value of “3”, which is the middle value between 1 (very easy) and 5 (very difficult). Five subjects in the Group B (Timbre) rated the game difficulty by value of “2”. An unpaired two-tailed t-test was used to check if there was a significant difference between the two groups in terms of their perceived game difficulty. The results showed that there is **no** significant difference with $p = 1.0000$ ($t = 0.0000$, df = 18, $\alpha = 0.05$)

The limited amount of different earcons (a subject could choose another different earcon after respawning) and the fact of no completion timing stressor might be the explanation of such results.

**Gaming experience**

Gaming experience was one of the selecting criteria to be able to participate in the experiment. It required to have at least some digital gaming experience.

Group A (Rhythm) had 5 (4 without the outlier) subjects who play 5-10 hours per week in average, 2 subjects who play 10-20 hours per week in average and 3 subjects who play less than 5 hours per week in average (*Fig. 3.5*).
Group B (Timbre) had 6 (5 without the outlier) subjects who play less than 5 hours per week in average and 4 subjects who play 10-20 hours per week in average (Fig. 3.9).

The more time a player spends on gaming per week – the more potential for a better performance. However, it cannot be stated that analyzed gaming experience data gave any noticeable pattern between the subjects’ results.

The same results showed the preferable game genres (Fig. 3.6 and Fig. 3.10), i.e. it cannot be stated that there was a pattern that showed increased performance results depending on the preferable game genres.

**Studying program**

The question about the current studying program might help to see whether a subject has deeper experience in audio (Fig. 3.7 and Fig. 3.11). Most of the subjects were audio engineers, 8 in the Group A (Rhythm) and 7 in the Group B (Timbre). There was one subject in the Group B who studies music, which can be a marker of more experience in audio. Therefore, there should not be any influence on the performance results depending on the studying program in the experiment.

**Qualitative data**

**The game strategies and player behavior**

The most common game strategy in both groups was to find the new sound which differs from all the others. The main difficulty was in the room 2 and 3, where there was more than one different earcon. Some subjects randomly guessed and some reported that they searched the sound that they heard when they passed the previous door. One subject (S15) in the Group B (Timbre) had 0 wrong trials, although it is unclear how he/she made the choices.

“It was only possible to pick up 1 key, so I searched for the key which sounded different from the others. So, I continued with this strategy.” (Subject 15).

The recorded S15 gameplay did not provide any helpful information about subject’s strategy.
The main observation of this data is that subjects did not completely understand the game logic and why they had to choose one or another earcon in the room 2 and 3.

**Group A (Rhythm) and Group B (Timbre)**

The qualitative results from the questionnaires were summarized according to the patterns which subjects in both groups had in common. The game strategy descriptions and perceived earcon difference did not differ between the two groups. Most of the subjects could hear the difference between three different earcons (three different earcons in one condition), which is the main and important observation.

**In common patterns in both groups**

- Describe your strategy using keys

The most common strategy was to choose earcons that sounded differently from the others. None of the subjects explained which strategy he/she had in the second and the third room. One of the subjects wrote that he/she randomly guessed when it was more than one different sound.

- Describe any differences you heard in the key-sounds

Five subjects described that they heard difference in the rhythm/speed. Three subjects wrote that they heard difference, although they did not explain which one they noticed. Two subjects stated that some of the sounds had more bass frequencies than the others.

- How much did the sounds help you to solve the tasks?

All subjects stated that it was only the sounds which led them. Two subjects described further that in the second and in the third room they randomly guessed between the two different sounds.

**Experiment findings**

According to the experiment results it is possible to conclude that rhythm condition showed better performance than the timbre condition in this particular game scenario setup. The subjects in the rhythm group could identify and recognize different earcons more easily than the subjects within the timbre condition. An interesting observation was made on players’
comprehension of the game logic. None of the subjects in both scenarios reported as he/she understood the logic in choosing correct earcons in case there were more than two different earcons.

Although, the fact that a subject could start from the beginning (in case of the wrong choice) leads to another interesting observation. As it was concluded by McGookin and Brewster (2011), the training phase could increase earcon recognition up to 80% (McGookin & Brewster, 2011). The similar conclusion was made in another study by Dingler, Lindsay and Walker (2008). There was no separated training phase in the experiment, however, the earcon presentations in the first room and the potential respawns during the whole game might have a role of a “training”. The subjects in both conditions showed increased performance after respawns.

**Limitations of the experiment**

The earcon design has a great influence on the players’ behavior. It could have been possible to design earcons with larger difference within the timbre condition, which might lead to other results. It was decided to design earcons with balanced difference in this experiment, i.e. the difference between the earcons in rhythm condition should be not greater than difference between the earcons in timbre condition. It was made in order to have a fair comparison between two conditions. However, it was not so easy to find well balance, therefore a potential larger difference within the rhythm condition might be the reason for such results.

The lack of game logic comprehension can be explained by limited amount of supportive visual cues. It was an attempt to achieve the balance between the game difficulties. It should not be very easy or very difficult. A possible solution might be to increase the amount of supportive visual cues and make it more understandable to the players.

Another observation which was made is that subjects chose the second different earcon, in case they did a wrong choice previously. The fact that there were only three different earcons (room 2 and 3) limited the possible choices made by listening and not by random guessing.
Further research

The more complex research on different earcon attributes might be recommended in order to study players behavior in various conditions. Various conditions in this context mean different game genres and different game tasks. It is important to notice that the game scenario and conditions have influence on the sound designers’ choice. Different game scenarios might require different approaches for a sound design and different choices of earcon attributes. For example, some attributes might be more proper for scenarios when a player needs to make a fast decision, and other attributes might be a better choice in more puzzle-oriented type of games. There is a potential in further research in terms of different game genres and game tasks, as well as in a more complex compound earcon research. For instance, different changing rhythm and timbre parameters can be studied within the same earcon. Despite the conditions and attributes choice, it can be recommended to create an ecologically valid experiment in order to achieve better results for a further practical use. It is also important to design fair-balanced earcons, so that possible larger attribute difference between compared earcons will not have big impact on results.
References


GraphPad Software. (2018). T-test online calculator [computer software]. La Jolla, California: GraphPad Software


Appendix

The questionnaire

You are NOT allowed to tell any details about this experiment before the final results reveal.

1) Describe your strategy choosing keys
   Long answer text

2) Describe any differences you heard in the key-sounds *
   Long answer text

3) How much did the sounds help you to solve the tasks? *
   Long answer text

4) Was the game easy/difficult? *

   Very easy

   1 2 3 4 5
   Very difficult

5) How many hours do you usually play games per week (in average) *

   ☐ 0-5
   ☐ 5-10
   ☐ 10-20
   ☐ More than 20

6) Which type of games do you usually play? *

   ☐ FPS
   ☐ Adventure
   ☐ RPG
   ☐ Strategy
   ☐ MOBA
   ☐ Other...

7) What program do you study? *
   Short answer text