Songs of Transistor

A study of sound design in video games

By: Ingeborg Asplund

Supervisor: Petri Lankoski

Södertörn University | School of Natural Sciences, Technology and Environmental Studies

One Year Master Thesis, 30 hp

Media technology | Spring semester 2017

SÖDERTÖRNS HÖGSKOLA | STOCKHOLM
sh.se

Songs of Transistor a study of sound design in video games

Ingeborg Asplund
Sodertorn University
Stockholm, Sweden
Ingeborg.asplund@gmail.com

ABSTRACT
While there is a lot of research about other aspects of game design, there are fairly few studies about music and sound in video games. Since music and sound are components of next to all games, it is interesting to investigate how this aspect affects the perceived immersion of gamers. The aim of this study is to investigate how sound and music affect player sense of presence in a video game, *Transistor* [19], which was chosen due to its distinct and strongly emotional music and sound. Five video prototypes were made using gameplay and sound from the game. The videos presented different variations of the soundscape. These were tested by a web survey with questions from the PENS questionnaire [15], providing the users a seven point Likert scale by which they could rate their experience. The answers were analyzed with a mixed model regression and compared with an estimated image of which degree of immersion would be experience for each of the videos. The result of the study showed that the complete soundscape was significantly more immersive than all the other soundscapes, while silence was significantly less immersive than the other soundscapes. The conclusions were the more complete the soundscape is, the more immersive it is, and that even a small part of the total soundscape is more immersive than complete silence.

Author Keywords
Sound design; Video games; Interaction design; Prototypes

ACM Classification Keywords
Computing~*User studies* • Human-centered Computing~*Usability testing*

INTRODUCTION
Immersion in video games is, according to the definition by Jennet, Cox and Cairnes [9], a mental state perceived as being somewhere else than in the real world. It was described by their interview participants as "Being in the game". The most common way to measure *Immersion* is *Presence*, which can be described as the sense of being, and is a crucial part of the experience of video games. Immersion and presence are sometimes used interchangeably as they describe similar concepts.

There are several factors that may contribute to the experience of presence as well as immersion perceived by the player. Those are the mechanics, the narrative, the art style and the sound design, given that they all contribute to the overall experience in a way that encourages the player to enter the game's world and embrace it. Immersion in a game can also be affected by the perspective the game is played from, as well as more complicated relations such as how the overall sound collaborates with the rest of the game and how the art and mechanics collaborate.

However there are also the aspects of sound and music. A question worth asking is what role these have regarding the immersion in video games. While there is much research about how story, gameplay and art affect the player experience in different ways the discussion about sound and music in video games is considerably less prominent. Still sound and music can have great impact on the sense of presence, which means that they can have a prominent impact on the immersion in the video game. Especially when the music is well designed and contributes to the themes and story of the video game in question, it can have a major effect on the feeling of presence perceived by the player.

While there is already much research concerning how mechanics, art and narrative contribute to the immersion perceived by players, notably fewer have investigated the role of music and sound design in immersion. Investigating the soundscape of games therefore offers a possibility to expand a relatively unexplored field. This in turn can provide interesting information for industry professionals and customers alike, as music and sound are indicated to have an effect on how games are percieved. In order to investigate how sound design and music is affecting the immersion in games this study focuses on the following.

- What role does the music play in the matter of immersion?
- How will changes in the sound of a game affect the player's immersion?
- Does the presence of music affect the level of immersion and how?

BACKGROUND

Immersion in games
Immersion in games can be described as a phenomenon by which the player is drawn into the world of a fictional

universe presented by a piece of media. This is often described and measured through the level of presence perceived while exposed to the piece. Presence can be described as the feeling of being somewhere else, as in stepping into a fictional character's mind [2].

Jennet, Cox and Cairnes [9] describe immersion as a gradual experience that is caused by presence building up over time. In their interview study 20 gamers were asked about their gaming experience. The researchers found that players who played for a longer time was more immersed in the game, compared to players who played for a shorter time.

The authors [9] describe the experience as a split between body and mind. It contributes to make the player perceive the game from the character's perspective [2]. This is according to Jennet, Cox and Cairnes [9] more prevalent in first person shooters, where the players do not see their avatar, but step into the character's shoes and view the game through its eyes. The investigation in question concludes that a definition of presence could be phrased as being "in the game". In these interviews they found that the gamers often referred to the sense of presence as if they had been in the game.

Music and emotion

Another factor that contributes to immersion of players is the evoking of emotions related to the game experience.

There are several studies on how music impacts and evokes emotion. Among them is the 1936 study by Hevner [8], which established a model where seven different emotions were linked to different musical pieces. The emotions were derived from adjectives used by the participants to describe what they felt when they were listening to the music. Hevner [8] categorized the adjectives as groups representing the seven different emotions. Then she used the adjectives to classify which emotion corresponded with which musical piece. For instance one of the emotions could be derived from the adjectives, *merry, joyous, gay, happy, cheerful* and *bright,* while another emotion could be derived from, *dreamy, yielding, tender, sentimental, longing, yearning, pleading* and *plaintive.*

A study similar to the study by Hevner [8], was carried out by Juslin and Lindström [12]. They exposed participants to soundtracks which they linked to five basic emotions based on the model Hevner [8] defined. They took into account different aspects of how the music sounded, for instance *Pitch, Rhythm,* and *Timbre,* when they measured the emotions participants perceived while listening to the music. A main difference compared to the study by Hevner [8], is that Juslin and Lindström [12] used music synthesis. This meant that the combinations of clips the participants were exposed to, were put together by a computer. This was combined with the element of interaction, and participants could rate the musical pieces using a digital system, while being exposed to the music.

Sloboda and Juslin [18] present and discuss methods for how to measure emotion in regard to music, as a diagram over different emotions that can be felt by an individual. Among the discussed topics is the use of musical stimuli which the participants were exposed to and rated, similar to the setup described by Hevner [8]. In relation to this they bring up the idea that music do not only strengthen already felt emotions, but also evoke completely new emotions in the listener.

The theme of music and emotion can also be related to Moore's [14] investigation of film music. He investigated the impact music could have on emotion with the focus on film music. Similar to Juslin and Lindström [9] as well as Hevner [8] Moore asked participants to rate how the music impacted them emotionally. However he also took visuals into account. Moore compared the emotions felt by participants when the same music accompanied different visuals. The participants' reactions were compared in order to see how the music affected the perception of the different visuals. Moore [14] also used a control group which was only exposed to the music and no visuals. This way he could compare the results of the two groups in order to see if swapping the visuals impacted the immersion significantly different from when there was only the music without visuals. The result showed that there were significant differences between how participants perceived the visuals, depending on which music was played.

Music and emotion in games

Based on what Sloboda and Juslin [18] write about how music evokes new emotion in listener, Lankoski [13] defines sound as one of three major factors that affects the way players feel when playing a game. These aspects are *Empathy, Beauty* and *Sound. Empathy* is mostly related to the narrative of the game, *Beauty* to the aesthetics and graphics while *Sound* refers to both music and sound effects in the game. Together these aspects work like puzzle pieces that build the overall emotional experience.

Sound is used to reinforce the message given by Empathy and Beauty and strengthen the message they are conveying. If the player achieves a goal in the game, the sound will reinforce the emotion the character is supposed to feel, move the emotion to the player and increase his/her emphatizing with the avatar. Similarly, sound can work together with the art and game mechanics and for instance communicate the tense feeling of a long difficult battle, or enhance the sorrow of a great loss the player character goes through. Lankoski's [13] reasoning is in this regard similar to Moore's [11] reasoning concerning film music.

However there is an important difference between games and films because games are interactive. The *Empathy* part is for games deriving from the gameplay, which means that the music and sounds of games enhances and has to work well with the mechanics of said games. The interaction aspect is what sets games apart from movies and therefore the sound in games may work in a different way compared to movies. This can be related to Bridget [2] who argues that game music works differently from music in films, due to that the music in games reflects and responds to the player's actions, while film music does not. Although visuals in games have become more and more cinematic in recent times, the music still has to fit with the player's actions in the game. Discussing what he refers to as interactive mixing, Bridget [2] brings up how game music are both more limited and different, in comparison to film music. For instance he reasons that while film music has a complete control over where and when certain music is played, the music played in a game depends on the player's actions and the situations the player is in. The music in games has to be responsive to player actions, which film music does not. This leads Bridget [2] to suggest that in the future games may give the player control over the dynamics in the music.

Further discussion on the impact of music in games is brought up by Ekman and Lankoski [5], who investigated the effect of the music in two horror games, *Silent hill 2* and *Fatal frame*. The focus of the study is on sound effects, as these games rely heavily on sound to convey the intended feeling of the environment. Ekman and Lankoski [5] conclude the sound effects and ambient noise are significant parts of the soundscape. The sound design in these horror games use silence and ambient noise as a contrast to build tension, similar to how music in other games may build emotional arousal through intense or emotional music. Although different from investigations such as that by Hevner [8] or Juslin and Lindström [9] the focus is still on the sound and its impact on emotion in the listener.

Music and its impact on immersion.

Another interesting aspect is how the emotions conveyed by the music affects immersion in games. Gasseleder [7] investigated the impact of game music on immersion. In relation to this he discussed *Expressive fidelity* in music and what it means for video game music. Expressive fidelity is according to Gasseleder [7] the way the music in the game collaborates with gameplay mechanics. It is how accurately a sound effect is played when the player picks something up, or the intensity of the music during a hard boss battle. Gasseleder [7] defines expressive fidelity as a scale with a level that can be higher or lower, dependent on how well the sound and music in the game are percieved to reflect the situation the player is put in.

To build further on the theory regarding the expressive fidelity, Gasseleder [6] made a study where he manipulated the sound of a video game, (*Batman Arkham City*) while participants were playing. He replaced the original music with other soundtracks and compared how the different soundtracks affected the players' immersion, dependent on the level of *emotional arousal* and the degree of *dynamics* in the music. In order to gather information from his respondents Gasseleder [6] employed a method similar to the one used by Juslin and Lindström [9]. Participants rated their experience using a digital system. He found that the highest immersion effects were created by soundtracks that had high dynamics, and in some situations by soundtracks with low degree of dynamics and low emotional arousal level. According to Gasseleder [6] this indicated that the players' immersion increased when the degree of dynamics of the music increased.

Another investigation in the same field was made by Allouche et al. [1] who investigated how different music could change the experience of a simple open source video game, which they modified in order to allow players to choose from several predefined genres of music. Each genre also had a different color scheme which changed the appearance of the game depending on which genre was chosen. Allouche et al. [1] still put a heavy focus on how the change in sound would impact the gameplay experience. One main idea behind the project was to allow players to customize the sound of the game and measure how this affected their experience of the game. This was a different approach from that of Gasseleder [6], as the players themselves were manipulating the sound, thus changing aspects of the game and how it felt to them. Although not investigating the aspect of immersion as much as Gasseleder [6], Allouche et al. [1] present an interesting investigation about music and the gameplay experience.

Jorgensen [11] writes about a previous investigation of how players were affected when the sound was turned off while they were playing a game. She made players play two games (*Warcraft III* and *Hitman contracts*) and turned off the sound halfway into the gaming session. Afterwards she interviewed the participants regarding their experience of playing the games with and without sound respectively. She identified several differences in how players reacted to having the sound turned off. As the title of her article "Left in the dark" suggests, losing the sound made players partially lose control over what happened in the game world, as they missed out on crucial information. For instance indications about where enemies were located and information about upcoming events were lost. This led to a decreased performance among players, as they found themselves surprised by the game events and therefore had less control over what happened to them in the game world. The emotion and the sense of

presence in the game were also negatively impacted by the lack of sound.

In relation to Jorgensen's previous study [10] she discussed the importance of functionality of the sound in games [11]. She defines this as the relationship between the player actions and the responsive sound in games. According to Jorgensen [11] audio in games communicate to the player where they are and what is happening around them, serving both an immersive and a functional purpose. By investigating how two different games (*Warcraft III* and *Hitman Contracts*) used sound, she found five different categories of functions regarding the use of sound in games. These are defined as *Action oriented functions, Atmospheric functions, Orienting functions, Controlling functions* and *Identifying functions.* The *Action oriented functions* are sounds related to the combat situations while the *Atmospheric functions* are related the emotions conveyed in different settings and situations. The *Orienting*, *Controlling* and *Identifying* functions of sound is to make the players aware about different events in the game world, point out directions of incoming danger and help them to identify valuable objects and noteworthy locations. As Jorgensen [10] investigated what happened when she changed the soundscape in two different games. Her investigation is somewhat similar to Gasseleder's [6]. The difference between their studies is that Jorgensen [10] turned off the sound completely and compared music to silence, while Gasseleder [6] changed the music and compared the effects of the different soundtracks with each other.

Yet another perspective is delivered by Sanders and Cairns [16], who investigated how the sound affects the level of time perception in the players. *Time perception* is described by the authors as the perception of how much time have passed. While playing games this perception is changed as the players immerse themselves in the game world. By measuring the participants' perception of time they investigated how music affect the immersion compared to silence.

Sanders and Cairns [16] conducted two experiments in total. In the first experiment the music had negative impact on the players' immersion because it was perceived to clash with the game. In the second experiment, where the music was changed after feedback from the participants, immersion was higher. An interesting aspect of this is that the researchers first intended to only make one experiment, but since the music they chose in the first experiment clashed with the game they had to change the music and conduct a second experiment.

Their results indicated that the type of music used in a game can have a substantial impact on how players perceive that game. If the music clashes with the theme, feel and mechanics of the game, the immersion becomes lower than if the music fits the game. Music is therefore according to Sanders and Cairns [16], an important aspect of how a game feels and plays.

RESEARCH QUESTION

In order to investigate how changes in the sound and music in the video game *Transistor* [19], affects perceived immersion I've formulated the following research question:

"*How does changing the soundscape in a video based on gameplay from the video game* Transistor [19] *affect the watcher's level of immersion?*"

Hypothesis

The study investigates the difference between the complete soundscape and manipulated soundscapes where sound elements have been edited out. Since previous research has shown that music is an important aspects of immersion in video games, the first hypothesis is that there will be significant differences in immersion between the *Close to original* soundscape and the manipulated soundscapes. The second hypothesis is that the result will be as described in the estimate picture (Figure 1) which predicts that the impact on presence will be biggest for *Close to original*, followed by *Music only*, *No voice acting*, *No music* and *Silent*, in that specific order.

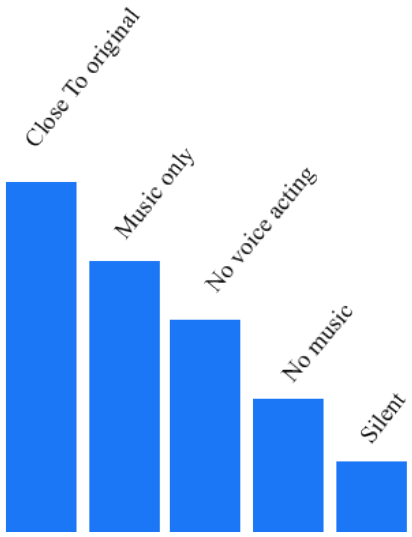

Figure1: After hypothesis formulated preliminary image concerning how the level of immersion for each of the cases are predicted to relate to each other.

In relation to the hypotheses I made a preliminary image placing the cases in order after which level of immersion I believe they would cause the viewer to perceive.

As shown in Figure 1, the estimation is that the *Music only* will have the next highest level of immersion after the *Close*

to original case, followed by *No voice acting*, *No music* and *silent.*

METHOD

Choice of study material

Transistor [19] (the game chosen for the prototypes) is a top down action RPG (Role play game) with a focus on combat and strategy. A striking feature of the games overall aesthetics is the distinct and strongly emotional music. The narrative of the game is mostly told through the players actions alongside commentary from a talking sword the player character brings with her.

The part of the game featured in the video prototypes (see figure 2 above) is the first major encounter with a member of the in-game villain organization the Camerata (the boss character). It consists of the event that sets off the battle, the battle itself, and the aftermath where the boss character dies and is stored inside the talking sword of the player character.

The battle music is rather distinct and features a lyrical song, which two thirds into the battle changes into a slightly metallic corrupt version of itself. This, alongside other sound elements, such as the boss shouting replies and the sword's commentary, creates a memorable soundscape during the fight.

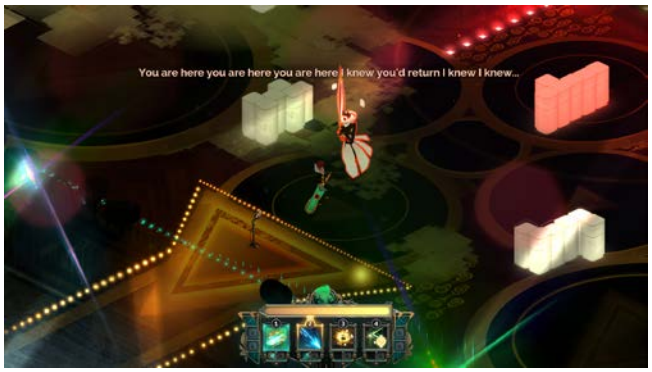

Figure 2: Footage from the battle scene used in the video prototypes featuring the rather distinct lyrical song "*In circles*".

Design of videos

For the experiment of this study, five video prototypes featuring gameplay from *Transistor* [19] were made. All the video prototypes were created based on a video clip made in Adobe Premiere Pro. The clip was composed of the sequence of recorded gameplay described in the material choice section above. It was cut and phased to be roughly one minute long. The clip contained the sound effects, music and voice acting in three different layers and could be turned on and off depending on which aspects of the soundscape that were desired. In the sound effect layer free generic crash sounds were added in order to replace the attack sounds from the game, as these were not possible to retrieve from the game files. Five video prototypes with different soundscapes were then created by turning on and off different layers.

Close to original [21] (the first video prototype) is as close to the original soundscape as possible, which means that it features most of the original sounds of the game, but has been compiled in Adobe Premiere Pro so that the sequence is not a direct copy of the game but has been slightly modified.

No music [22] (the second video prototype) contains the voice acting and sound effects only. The music is omitted.

No voice acting [23] (the third video prototype) contains sound effects and music. The voice acting is omitted.

Music only [24] (the fourth video prototype) contains only the music. All other sounds are omitted.

Silent [25] (the fifth prototype) is completely silent.

The way the videos were designed means that the visual stimuli is exactly the same for all prototypes, whilst the auditory stimuli differs. This is similar to how Moore [14] set up his experiment, although he used the same music for different visuals and tested for reaction to clashing visuals and images. The current study also shares some similarities with the experiment by Gasseleder [6], who investigated perceived emotions in relation to different characteristics in music. However, the participants in the current study were not playing the game but watched the audiovisual prototypes.

It was important to achieve high and equal sound quality of the video prototypes. Many small adjustments of the soundscapes were made. Repeated testing and several iterations led up to the prototypes currently used in the study. The quality of the *Close to original* case was increased by a slight change of the volume of the music layer. For *Music only* and *No voice acting* the music layers were adjusted in order to achieve even volume, in order for these to work as separate video prototypes. No adjustments were made to *No music* and *Silent*.

Details turned out to be important when the design was tested by students who beta tested the videos. For instance the prototypes were perceived as more immersive when the introduction to the battle was included in the clip, as that added a setup for the events that followed. It also paid off to do fine adjustments to the sound effects layer, in order to make the sound effects feel softer and more incorporated into the soundscape.

Survey design

A web survey including the five video prototypes was created in google forms, based on the PENS questionnaire [15], specifically the part that works with presence and immersion. The PENS questionnaire has been validated and found to perform well in comparison with the GEQ

questionnaire [3]. It showed an explained variance of 80% for the items investigated by Brühlmann [3]. The GEQ questionnaire on the other hand had an explained variance of 68%, which is high, but not as high as the explained variance found with the PENS questionnaire [3]. The PENS questionnaire was also used in the study by Dennie [4] who conducted a study with some similarities to the current study.

In the survey the participants were tasked to watch the video prototypes and then answer eight questions per video. In order to answer the questions the participants rated each of them on a 7 point Likert scale, based on how much the participants agreed with the statement. Agreement ranged between 1=totally disagree and 7=totally agree. Examples of questions were *"When watching the video, I feel transported to another time and place"* and*" When watching the characters moving through the game world I feel as if I am actually there"*

Since the participants were watching videos instead of playing the game, the questions were slightly adjusted to be about watching instead of playing. These adjustments were kept close to the original phrasings and did not fundamentally change the question, in order to minimizing the risk of losing statistical reliability.

As the order of exposure to material may influence the outcome of the results, this study made use of counterbalance, specifically the Latin square model as described by Shuttleworth [17]. According to Shuttleworth [17] counterbalancing can be done in several ways. If there are four or fewer items to counterbalance it is possible to provide all the possible orders the items could come in.

However when the items become more than four the combinations are so many, that it is more effective to select a few of the possible orders, and randomly provide them to the participants. This is called the *Latin square* [17], which is a model used for incomplete counterbalance in situations where more than four items need counterbalancing.

In this study with five videos, the *Latin square* [17] was used, as it reduced the number of alternating questionnaires that had to be created. This means that there were five versions of the questionnaire where the prototypes were presented in different order. The five versions of the questionnaire were randomly assigned to the participants, in order to counter the risk ordering effect. [17]

Data collection

The questionnaire was provided online via *Facebook* and different game oriented forums such as *Steam* and *Game Spot*, and gathered 28 participants in total. Out of these was 15 men, 12 women and one of other gender identification. Ages ranged from 19 to 57 (mean= 28.03; SD = 0.75.)

Five participants had played *Transistor* [19] before and a majority of the participants responded that they played games often. Out of the 28 participants 11 played games every day, 8 played games some times per week while 7 seldom played games and 2 never played games.

Due to that the study was carried out online it is not possible to know how many who visited the questionnaire but did not send a response nor how many were exposed to it but chose to not answer.

Data analysis

Data was analyzed with the statistics software *R*. Five presence variables, based on each of the five cases, *Close to original, No music, No voice acting, Music only* and *Silent*, were calculated as the mean values of their respective item scores, using the *lsmean* function in *R*. Means and standard deviations were calculated for each of the variables.

Of eight questions one was reversed. This question was particularly similar to one of the other questions that were included in the questionnaire. The reversed question was therefore left out. This means that the presence scores calculated in the analysis are based on seven of the questions.

The study used a mixed linear model with a within subjects design. The type of analysis conducted with the model is a linear regression. Winter [20] describes how to conduct mixed model analysis with R. Using the instructions he provides for mixed linear regression a regression was carried out. All the variables were entered as one block. As part of the regression a Tukey test was performed by R, where the cases were compared pairwise with each other in order to see if there were any statistically significant differences between them.

RESULTS

The mean values and standard deviations for the variables can be found in Table 1 and the confidence intervals are presented in Figure 3. The result of the regression can be found Table 2.

Not all the *differences* in the impacts between the cases were significant. Close *to original* had significantly higher impact on the presence than all other cases. *Silent* had significantly lower impact on presence than all other cases.

The case with the highest difference in perceived presence from *Close to original* was *Silent* with a difference of -1.9 followed by *No music* with a difference of -1.1, *Music only*, with a difference of -0.8 and *No voice acting* with a difference of -0.6.

The impacts on presence of the remaining three cases (*No music*, *No voice acting* and *Music only)* did not differ significantly from each other.

Case	Mean	Standard deviation
Close to original	3.30	1.38
No music	2.11	0.86
No voice acting	2.61	1.24
Music only	2.42	1.08
Silent	1.39	0.52

Table 1: The table shows mean values and standard deviations for the study's variables

The Tukey test with pairwise comparisons points towards a significant difference between the *Close to original* case and all other cases as well as a significant difference between the *Silent* case and the others.

The comparisons between *Close to original* and *No Music* and *Silent* both had a statistical significance of $p < 0.001$. *Close to original- Music only* also had statistical significance of $p < 0.001$ and *Close to original-No voice acting* statistical significance of $p = 0.013$

When *Silent* was compared to the other cases statistical significance was generally high. *Silent- No voice acting* had a statistical significance of $p<0.001$. *Silent- Music only* also had a statistical significance of $p < 0.001$, while *Silent-No music* had statistical significance of $p = 0.007$

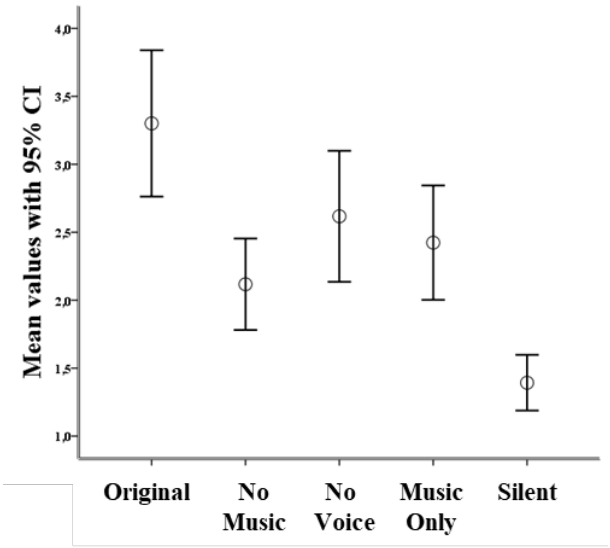

Figure 3: The graph shows the mean values and confidence intervals for the cases. *Close to original* perform the best of the cases and *Silent* perform the worst. In the middle we find a segment containing *No voice music, No voice acting* and *Music only*.

Model	Presence
Intercept	3.30 *** (0.20)
No music	-1.18 *** (0.22)
No Voice acting	-0.68 ** (0.22)
Music only	-0.88 *** (0.22)
Silent	-1.91 *** (0.22)
N	140
N groups: id	28
Var: id (Intercept)	0.48
Var: Residual	0.66

***p < 0.001, ** p < 0.01, * p < 0.05

Table 2: The table shows the results of the regression analysis wih presence as dependent variable.

Thus the statistically significant differences that are possible to spot are those between *Silent* and all other cases and *Close to original* and all other cases.

The remaining three cases *No music*, *No voice acting* and *Music only* did not show statistically significant differences when compared to each other, meaning that they cannot be placed in any special order.

DISCUSSION

In relation to hypotheses

The study investigated the differences between the original soundscape and several altered soundscapes in the video game *Transistor* [19].

The first hypothesis was that there would be a significant difference between the *Close to original* case and the other cases. This hypothesis was supported by the results.

The second hypothesis stated that the result would be as described in the preliminary picture (Figure 1) presented in relation to the hypothesis, which predicted that the impact on presence would be biggest for *Close to original*, followed by *Music only*, *No voice acting*, *No music* and *Silent*, in that specific order. This was not supported by the results.

Main findings

The main findings of this study is that the more complete the soundscape is the more presence is perceived, meaning that more elements in the soundscape creates a stronger sense of immersion. The study also showed that the perceived presence decreased when the soundscape was altered by removing elements, and that a completely silent soundscape gave very low scores of presence. This indicates that when

you remove something from a game soundscape the level of perceived immersion will decrease.

The highest presence score was achieved by the *Close to original* case and the lowest score was claimed by the *Silent* case. It is quite logical that a soundscape generates a higher sense of presence when there is sound effects than when there is not, and the immersion drops substantially when a game with a strong soundscape goes silent.

The significant differences were between the *Close to original* and all the other cases, as well as between the *Silent* and all the other cases. Which of the other cases that differed the most from *Close to original* beside the *Silent* case, is uncertain and may be different in another investigation.

If the differences between the cases *No music, Music only* and *No voice acting* had been statistically significant, these could have indicated unique contributions to immersion by sound elements like music or the voice acting, as well as providing information concerning how the different soundscapes affect the player immersion. The impacts of the differences between those cases and how they perform in relation to each other could be highly interesting to investigate further in future study.

Another interesting aspect of the results of this study was that *Silent* contributed significantly less to immersion than all the other possible soundscapes. This means that even the cases that only had a fraction of the soundscape (For instance *No music*) performed significantly better than the *Silent* case. This indicates the importance the mere existence of any kind of sound has for a game to feel immersive.

In relation to theories and previous research

Compared to theories and previous research, the results of this study show some interesting things. First I want to bring up Gasseleder [6] again. The experiment he conducted with *Batman Arkham city* is quite similar to the current study, except that he focused on dynamics and emotional arousal and manipulated the sound while participants were playing the game. Despite the differences there is still a possibility to compare the finding of this study with the types of results Gasseleder [6] got. Gasseleder's [6] results that indicated the importance of dynamics in the music, while this study got results that indicated that the more complete the soundscape is, the more presence the participants perceive. The two results concern different aspects of a game's sound design, and both may contribute to the experience of the sound in a game, but on different levels. The music may contribute to the immersion by capturing the environment and feel of the game, while the other sound effects may add that little bit of extra vividness that immerses the player even more in the game world. In the current study one of the strongest evidences is for the importance of those extra sound effects, as *Silent* performed worse than all other cases. The *Transistor* [19] sounds from the speaking sword and from the boss, the attacks, and the music all contributes to making the soundscape be perceived as more alive, and therefore they generate a higher level of presence.

Gasseleder's reasoning [7] about expressive fidelity (the flow of the sound effects and music) in games is also interesting to compare the results of the current study with, especially as the results of this study indicate that the expressive fidelity may increase with each element that is added. This may mean that the sound effects and voice acting in the case of *Transistor* [19] may have increased the expressive fidelity of the game, at least when participants viewed the game as video prototypes.

The experiment in Gasseleder's study [6] also comes close to the experiment by Moore [14], who exposed his test participants to samples of audio and visuals and swapped the music for key scenes with each other, in order to investigate how music that clashed with the visuals would affect the viewers. However, Moore's experiment did not involve interaction. There is also a difference in which media is investigated. Moore [14] investigated film, while Gasseleder's [6] study focuses on video games. The current study had the goal to investigate the soundscape as a whole, while Moore [14], focused solely on the music and its impact. Similarities both studies were the way in which sound was studied. A similarity with Moore's [14] study was the lack of interaction. Gasseleder [6] and Moore [14] both focused on music only while the current study focused on the full soundscape, but there are similarities in the findings for instance that sound impacts immersion

A highly interesting comparison can be made between the current study and the study by Jorgensen [10]. Both those studies focus on how player experience is affected by changes in the soundscapes. While Jorgensen [10] investigated how the players were affected by turning the sound completely off when they were playing, the current study compares the effect on presence for different edited soundscapes when the participants watched videos. The similarity of the approaches is that the effect on the player experience when sound was removed was investigated. The differences are that the current study removed parts of the soundscapes while Jorgensen [10] removed the sound as a whole element, and of course that the experiment in the current study was non-interactive. Regarding the results, Jorgensen's [10] findings showed a negative effect on player immersion and emotional engagement with the game when sound was removed, which is similar to the result of the current study, which found that immersion decreased when participants watched videos where elements of the soundscapes were omitted. Jorgensen [10] also found that

player performance was affected negatively due to loss of crucial gameplay information. This was something that the current study did not investigate.

Regarding Jorgensen's[11] reasoning about different functions of sound in video games, the *Atmospheric functions* are the ones investigated in the current study, as the main focus was to investigate the change in perceived immersion when different aspects of the soundscape was omitted. Since she identifies *Atmospheric functions* as linked to the sense of presence, the results of the current study are interesting, as they show that the more sounds a soundscape contains the more immersive it is percieved, while any type of sound is more immersive than silence. This may indicate that the *Atmospheric functions* of sounds are affected by how complete the soundscape feels and not just by music and voice overs such as in Jorgensen's[11] examples.

The perceived presence can be related to perceived emotions as these may contribute to the amount of immersion perceived. In the PENS questionnaire [15] some of the questions concern the emotions percieved by the players while playing the game. For instance the participants were asked to rate a statement reading: *"I experience feelings as deeply in the video as I have in real life"*. This as well as other questions were about emotional engagement with the game. The relationship between emotion and presence can also be associated with the study by Jennet, Cox and Cairns [9]. The players they interviewed especially mentioned perceived emotions, when they described how they had felt like they stepped out of themselves and into a fictional character. Based on this the hope of the current study was to find differences between the cases that pointed towards the importance of the music, as previous research have linked music to perceived emotion [5,8,12,13,18].

The influence of the music alone is hard to estimate. However the fact that the *Silent* case had a significantly lower presence score than all the other cases, may indicate some sort of emotional impact added to the other dimensions of the soundscape. This can be related to what Lankoski [13] writes about sound as a major component in how a player perceives a game they play, in particular by adding a layer of emotion to the equation. Since the presence score increased when the other cases were compared to *Silent,* it is likely that the added sounds also added a layer of emotion and therefore increased the feeling of being there. The *Close to original* case had a significantly higher score than the other cases, which also may hint towards an increased emotional impact when more sound is added.

According to Ekman and Lankoski [5], the sound had a unique own effect on the perception of the environments in the games they investigated. The games they investigated (*Silent hill 2* and *Fatal frame*) had a strong focus on ambient noise, while *Transistor* [19] examined in this study has a stronger focus on music in its soundscape. Therefore the impact of the sound effects may be different in *Transistor* [19].

Taking into account theories and previous experiments, such as those by Hevner [8] or Juslin and Lindstrom [9], both working from the perspective that music evokes emotions in the listener, there was an expectation in the current study that the music would have a significant impact on the soundscape. While the previous studies did not directly test emotion, the models presented by them hinted towards a strong connection between music and perceived emotions. In turn this would lead to that music alone could have a certain role in the soundscape of a game. The sound effects were considered an interesting aspect of the soundscape when the prototypes were created. However, there was no special sound effects prototype made as with the Music only case. This was because previous research indicated that the music alone should have a higher importance than all the other aspects of the soundscape in the game.

Findings such as those by Sloboda and Juslin [18], also point towards a stronger importance of the music compared to other sounds. They found that emotions are strengthened by music and in addition also aroused by music. People can feel their emotions change while they are listening to music. These findings made it logical to expect higher impact on immersion by the music in the current study. However there was no significant difference between the *No music, No voice acting* and *Music only* cases, which should have been needed in order to point out music as being a specific contributor.

In relation to method

All the cases were videos and not different modifications of the game, which means that the participants did not play the game, but watched it in videos. Regarding the interaction aspect it can be mentioned that music in video games differ from music in films [ref 3] as music in games has to respond with the player's interaction and change dependent on the situation the player is in. The sound of games may be perceived differently when people are watching compared when they are playing. A factor that can have affected the result is thus that the participants did not play the game, but instead watched a number of video prototypes. Due to this, the presence scores may not represent what one may feel when playing the game. This can be one reason for the lack of statistically significant differences between *No music, No voice acting* and *Music only*.

Another reason why the more edited cases (*No voice acting, No music* and *Music only*) ended up as not significant could be issues with the prototypes. The videos in the current study used some custom sound effects for player attacks. This could be one issue and the result may be different if the

original attack sounds from the game would be used instead. A solution to this could be to modify the game's audio files and for instance replace the desired files with new edited ones. For instance it could possible to replace certain audio files with empty audio files, in order to quiet specific parts of the soundscape. That was the way Gasseleder [6] handled sound manipulation in his experiment. He used an external third party program to access the sound files of *Batman Arkham city*, and replaced sound files while participants were playing the game.

Questions for future research

There are a number of possibilities for future research.

A specialized study focused on the music in video games and similar to Gasseleder [6] or Sanders and Cairns [16] (they made alternations while players are interacting with the game) could be a way to further study the changes in immersion in video games.

A study that could serve as an interesting piece of inspiration is that by Allouche et al. [1] who focused on user centered customization of the music in a video game. It would be possible to develop a game which (similarly to the game in the study by Allouche et al.[1] allow the users to choose from different soundtracks to accompany the gameplay, and keep track of which of the soundtracks the players choose, then interview them about their reasons. Such a study could possibly give more information regarding the influence the music has on the gaming experience.

To more specifically investigate the effect music has on presence, it could be an idea to change the original music in an existing game, preferably while players are playing the game. That would take the dimension of interaction into account. Another idea could be to measure more than one factor. Presence is one way to measure immersion, however one might use *Time perception* as an additional instrument of measurement too, like Sanders and Cairns [16] did.

Yet another possibility is to keep this study's concept with a number of videos with gameplay from one game, (*Transistor* [19] or another game that has an interesting soundscape) and improve the quality of the videos, by polishing the presented soundscapes more and only use sound found in the games audio files. Simultaneously the aim could be to both interview participants and have them fill in a questionnaire. This could maybe lead to deeper understanding and more significant results and help to explain why the differences between the cases in the middle segment was not statistically significant. Information that is not possible to read from the quantitative data alone could be added from the interviews. For instance it could explain why the presence scores are coming out a certain way or give hints towards other factors than the soundscapes alone that may affect the immersion.

A variation of the above described suggestion is to investigate music as an element in sound design of games, by replacing the *Silent* prototype with music from outside the game. This would be interesting since the investigation by Sanders and Cairns [16] unintentionally showed that players could react negatively to music they thought was clashing with the rest of the game. A prototype with clashing music could add an interesting twist to a study similar to this one.

Another suggestion is to create several prototypes that study singular aspects of the soundscape in a game. This could for instance be a comparison between, *Music only*, *Voice acting only* and *Ambient noise/sound effects only*, in order to investigate specific parts of the sound in the game.

In a future study it would also be interesting to compare the differences of perceived presence between people who have previously played the game and people that have not played it before. This would require equal amounts of respondents who have played *Transistor* [19] before and participants who have not. A similar setup could be done comparing gamers and non-gamers. Both these possible study setups would require more specific samples of participants.

Finally I think that it would be interesting with a study that directly tries to measure the emotional impact of music on immersion, similar to Gasseleder [6], but also measure the level of presence and make it possible to draw parallels between perceived emotions and perceived immersion. This could give an increased insight into how music as a specific element affects the way players perceive games and their degree of immersion in these games.

CONCLUSION

In essence this study has shown something interesting about sound in games, namely the importance of the *completeness of the soundscape*. The result is unique in that it indicates that every aspect that make up the soundscape in a game is important for the overall feel of the game and the immersion perceived by the player.

The first main result of this study is, that the more complete the soundscape in a game is the higher is the player's sense of immersion. This is shown quite clearly in the results as there are significant differences in perceived presence between the *Close to original* case and all the other cases, showing that the *Close to original* immersed participants more than the other cases.

In order to differentiate the effect the music as an isolated factor had on the overall immersion, the setup of the study could have been done differently. It would have been interesting to also focus more on the effect of the music alone as well as voice acting and sound effects. With a more

detailed focus on these individual parts of the game sound it would be possible to investigate more.

However there are some interesting facts emerging from this study in the form of the *completeness of the soundscape*. It is interesting that if one removes something from the sound in a game the outcome is less immersive which in turn proves the need of all the elements in the sound design to be there, the effect when the player hit an enemy, a line said by an opposing character as well as a specific piece of music may all contribute to the experience of the game and its soundscape overall.

The second main result is that a silent game is much less immersive than a game that incorporates sound and thus indicates the importance of sound as an element in game design. It stood quite clear that despite stunning graphics and visual effects the soundscape of *Transistor* [19] became better whenever sound was added even if the added sound was just a fragment of the original soundscape.

In essence the conclusion of this study is that sound in game matters and contribute to the level of immersion perceived by the player or in this case viewer. And the more complete the soundscape is or feels, the more the audience will immerse in the game. This could indicate that every single element of the sound is important and contributes to the experience and thus should be given attention in game design, in order to maximize the effect of the sounds.

ACKNOWLEDGMENTS

I would like to thank my supervisor Petri Lankoski at Sodertorn University for providing help and feedback during the different stages of the project, as well as for giving feedback on written material produced during the course of the project. I also want to thank all of my alpha and beta testers who provided me with feedback on the video prototypes and questionnaires during the design process.

REFERENCES

1. Allouche K., Albeiriss H., Zarghoune R., Arrasvuori J., Eronen A. & Holm J.2007. Implementation and evaluation of a background music reactive game. *IE '07 Proceedings of the 4th Australasian conference on Interactive entertainment.* Melbourne, Australia. Article No. 1. http://dl.acm.org/citation.cfm?id=1367957
2. Bridget, R., 2008. Dynamic Range: Subtlety and silence in video game sound. In Collins, K(ed). 2008. *From Pac-Man to pop music*, Burlington. Ashgate Publishing Company.
3. Brühlmann, F. & Schmid, G.M., 2015. How to Measure the Game Experience?: Analysis of the Factor Structure of Two Questionnaires CHI EA '15 *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* pp 1181-1186. DOI: 10.1145/2702613.2732831
4. Dennie, T.M. 2012. *Perception of autonomy and its effect on intrinsic motivation, immersion, and performance.* Master's Thesis at Graduate School of Western Carolina University. https://libres.uncg.edu/ir/wcu/f/Dennie2012.pdf
5. Ekman, I. & Lankoski, P. 2009. Hair-Raising Entertainment: Emotions, Sound, and Structure in Silent Hill 2 and Fatal Frame. In Perron, B. (ed.) *Horror Video Games: Essays on the Fusion of Fear and Play*. McFarland.
6. Gasseleder, H.P.2014. Dynamic Music and Immersion in the Action-Adventure an Empirical Investigation. *In proceedings of the 9th Audio Mostly*: *A Conference on Interaction With Sound*. New York, NY, USA, 2014. DOI: 10.1145/2636879.2636908
7. Gasseleder H.P. 2015. The Role of Agency in Ludoacoustic Immersion: Experiencing Recorded Sound and Music in Situational Context. In *proceedings of the Audio Mostly 2015 on Interaction With Sound*. New York, NY, USA, 2015. DOI: 10.1145/2814895.2814899
8. Hevner, K. 1936. Experimental studies of the elements of expression in music. *American Journal of Psychology*, 48, pp 246-268. http://www.jstor.org.till.biblextern.sh.se/stable/1415746?sid=primo&origin=crossref&seq=23#page_scan_tab_contents
9. Jennet, C.I., Cox, A.L. & Cairns, P.2008. Being "In the game". In *proceedings of the 2008 Philosophy of computer games conference*, Potsdam, available Through https://www-users.cs.york.ac.uk/~pcairns/papers/JennettPHIGames08.pdf
10. Jorgensen, K.2006. On the functional aspects of computer game audio. In *Proceedings of the audio mostly conference, 2006.* Piteå, Sweden. Available through: http://hdl.handle.net/1956/6734
11. Jorgensen, K, 2008. Left in the dark: Playing computer games with the sound turned off. In Collins, K (ed.) 2008. *From Pac-Man to pop music*, Burlington. Ashgate Publishing company
12. Juslin, P. N. & Lindström, E. 2011. Musical Expression of Emotions: Modelling Listeners' Judgements of Composed and Performed Features. *Music Analysis, Musical Expression of Emotions*, 29, 1-3, pp. 334-364. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-2249.2011.00323.x/full
13. Lankoski, P. 2012. Computer Games and Emotions. In. Sageng, Fossheim & Larsen (eds.). *The Philosophy of*

Computer Games. Springer. DOI: 10.1007/978-94-007-4249-9_4.

14. Moore, J. D. 2010. *The Impact of Visual-Music Interaction on Music Perception*. The Influence of Agreement and Disagreement. Master Thesis at Baylor School of Music, Baylor, TX. https://beardocs.baylor.edu/xmlui/handle/ 2104/8043 (retrieved 24 July 2014). https://baylor-ir.tdl.org/baylor-ir/bitstream/handle/2104/8043/jordan_moore_masters.pdf?sequence=1
15. Rigby, S. & Ryan, R. 2007. *The Player Experience of Need Satisfaction (PENS)* Immersyve Team available through http://immersyve.com/?s=Player+Experience+of+Need+Satisfaction
16. Sanders, T. & Cairns, P. 2010. Time perception, immersion and music in videogames. In *Proceedings of the 24th BCS Interaction Specialist Group Conference* (Dundee, Scotland), pp. 160-167 https://www-users.cs.york.ac.uk/~pcairns/papers/Sanders_HCI2010.pdf
17. Shuttleworth, M. 2014. Counterbalanced Measures Design. *Experimental research.* 2014. Retrieved 2017-04-06 from Accessible research guide Explorable:https://explorable.com/counterbalanced-measures-design
18. Sloboda, J. A. & Juslin, P. 1993. At the interface between the inner and outer world. In Sloboda J.A. & Juslin, P. (eds.) *Handbook of Music and Emotion: Theory, Research and Applications*. Oxford Scholarship online. DOI: 10.1093/acprof:oso/9780199230143.003.0004
19. Supergiant games. 2014 *Transistor*(Video Game, Pc.) .Supergiant Games
20. Winter, B. 2014. *A very basic tutorial for performing linear mixed model analyses(Tutorial 2)*University of California: Merced, Cognitive and informational sciences. Retrieved 2017-05-11 from: http://www.bodowinter.com/tutorial/bw_LME_tutorial2.pdf
21. Video prototype: *Close to original.* Available through: *https://youtu.be/Ybsa1v4-KeI*
22. Video prototype: *No Music.* Available through: https://youtu.be/wwD8UPhsk4M
23. Video prototype: *No Voice acting.* Available through: https://youtu.be/vmcDo5fgsKc
24. Video prototype: *Music Only.* Available through: https://youtu.be/d3ungcf4OSA
25. Video prototype: *Silent.* Available through: https://youtu.be/upLz54qoNo4