

An Investigation into the Effect of Music on Immersion in Video Games

Rowan Evans
Durham University

ABSTRACT

Immersion is often a major consideration in game design, and a well-designed soundtrack should only aid it. This study investigates the relationship of music and immersion in digital games – making sure to focus on music in specific, rather than the game sound as a whole. The experiment uses the 2D platformer *Celeste*, and uses a subjective measure, in the form of an immersion questionnaire, combined with more objective measures, in the form of perceived time distortion and performance in a post-game task. The results show higher immersion scores on the questionnaire from having music present, but the objective measures do not show any meaningful outcomes. This suggests immersion is more of a personal, reported experience, rather than being easily measured by separate methods. Following on from this study, more work should be done to decipher what objective means can be used to measure immersion levels effectively.

1. INTRODUCTION

The use of sound in games has been reported as being able to increase the phenomenon of immersion when playing. It is said to “envelop players in the game space” (Collins, 2013) and can transform less immersive digital environments, into ones that are more immersive (Grimshaw, 2012). The studies into sound and immersion have been skewed towards the entire games sound, rather than distinguishing it from music. Music has been tested separately for its immersive qualities in a few previous studies, however, the results have been inconclusive, and often containing conflicting information.

What is immersion? Immersion, in games, describes a state which a person’s “thoughts, attention and goals are all focused in and around the game as opposed to attending to being concerned with anything else” (Sanders & Cairns, 2010). Common traits have been defined in a previous study as being “Lack of awareness of time. Loss of awareness of the real world. Involvement and a sense of being in the task environment” (Jennett, et al., 2008). Immersion has been broken down also into different parts, perhaps exemplifying how complex it is.

Jennett et al. (2008) defined the three main stages of immersion being engagement, engrossment and total immersion. They also described three ideas which explain engaging experiences in the vein of immersion: flow (the process of being so absorbed in activities, that irrelevancies are screened out), cognitive absorption (state of deep involvement with a software), presence (the subjective sense of being in a virtual environment) (Jennett, et al., 2008). In this sense, immersion is not just a simple definition, and while this study may not delve far into the specifics, it is worth

keeping in mind when the concept of “immersion” is mentioned generally.

While this study should illuminate some aspect of the relationship between music and immersion, the scope from the outset is severely limited due to the diverse nature of digital games. The way in which games function in immersing the player will greatly differ from game to game, depending on genre and a multitude of other factors. Immersion in games can be summed up as being, “a many faceted phenomena with different aspects that can appear and be emphasised differently in the individual cases of games and players” (Ermi & Mäyrä, 2005). In short, no one player will experience immersion in the same way, and no one game will use immersion in the same way.

Previous studies into the relationship between music and immersion. For example, a study by Grimshaw, Lindley and Nacke (2008) found music to distract from a person’s sense of flow (an aspect of immersion) during gameplay – and not having music did not affect the level of immersion in a meaningful way. In fact, it was the diegetic sound only condition which induced the highest levels of immersion, which suggests that music could be more detrimental to immersion.

Another example which throws Music’s immersive qualities into doubt is a study conducted by Sanders and Cairns (2010). In the first experiment they conducted, the music that they used for their game had a negative effect on immersion – which they found was because it was not liked by the majority of participants. Once they obtained a “likeable” piece of music to use, recorded immersion increased. In this case, we can infer that a badly designed soundtrack would struggle to immerse a player. Therefore, more clarification is needed in deciphering the relationship between music and immersion in video games.

Apart from those mentioned, most studies have focused on the role of sound as a whole – ignoring the fact that music is a separate entity to sound design. One such study by Zhang and Fu (2015), which, while in its title claims to be about the “influence of background music of video games on immersion”, only tests for the entire sound of the game. They may argue for the game being chosen because of its “stirring background music”, but this isn’t worth pointing out if it is not tested for.

So, in light of this semantic error, this following study is a semi replication of Zhang and Fu’s (2015) experiment but making sure to test specifically for the influence of music rather than the entire sound of the game. While using elements

from their experiment that I think deserve repeating (use of post-game task, measuring time distortion, use of a questionnaire), this study will come at it from a different angle by testing the music of a game on its own merits.

Diegetic and non-diegetic sound. In the case of this experiment – since it is the effect of the music being tested – the diegetic sound has been left in. Diegetic sound is the sound ‘within’ the world of the game. For example, it includes wind sounds, footsteps, character speech, etcetera. The music, on the other hand, since it exists outside the confines of the of the game world, is labelled as non-diegetic. We, the audience, may hear the soundtrack swell as the tension of the game increases – but the characters of the game world would be oblivious to this music as it exists outside their world. Diegetic and non-diegetic sound are not fixed concepts, as many games drift between the two, but in the game chosen there is an explicit divide (shown by the separate ‘music’ and ‘sound’ controls).

Hypothesis and research question. This study will aim to illuminate the positive effect music has on the phenomenon of immersion, as opposed to the effect of diegetic sound. It will aim to obtain objective data that works in tandem with subjective data from an immersion questionnaire. As such the hypotheses are as follows:

- 1) The level of immersion when playing a digital game will be greater when the musical soundtrack is playing, as opposed to when no music is playing.
- 2) The measures of immersion used will be significantly correlated with each other.

Following these hypotheses, this will paper aim to answer the research question: *To what extent does music have a positive effect on the experience of immersion when playing a digital game?*

2. METHOD AND MATERIALS

Design. A repeated measure model was used for this experiment, and to lessen order effects, counterbalancing was used. Odd numbered participants would complete condition M (with music first, followed by W (without music) – while even numbered participants would complete this in reverse. The participants would complete the task first (playing the game), then, complete each designated measurement task in a set order (perceived time, Stroop test, questionnaire).

Participants. There were nine participants in total, the majority of these enrolled in an undergraduate course (66.67%), the rest being graduates in full time employment (22.22%) and also an A-Level student (11.11%). The ages ranged from 18-50, 77.78% of these being in the age range 18-21, and 22.22% in the age range of 45-50. The mean age was 25.78. The gender was 55.56% male and 44.44% female.

Questionnaire. The immersion questionnaire used was taken from a paper by Jennet et al. (2008) which I believed to be concise and explore important aspects of immersion – ending with a useable immersion score. It did not directly mention music in any of the questions, meaning participants were less

likely to have the subject of the experiment in mind when completing it. Each question was rated on a scale from 1 to 5, of which there were 26, making 130 the maximum score achievable.

Time Distortion. Time distortion refers to the phenomena of losing track of time, or not being aware of how much time is passing. This is one of the dimensions of flow which is in turn a dimension of immersion. Higher levels of time distortion have been linked in previous studies to higher levels of reported immersion (Sanders & Cairns, 2010) (Zhang & Fu, 2015), therefore I will be measuring for time distortion in this study. This concept will be measured for by the main task lasting a set period of time, followed by the player guessing how long they estimated playing for. Guesses of shorter elapsed times should be associated with higher immersion scores, as the more immersed a person is in the game, the less time they should guess to have passed, whereas a person who is not experiencing as much immersion should guess a longer period of time.

Post-Game Task. The Stroop test was used as a post-game task to give further objective data to show increased immersion. The idea in this being, that worse performance in a task after playing the game should indicate higher amounts of immersion, (so should be associated with higher immersion scores). Different tasks have been utilised in previous studies, such as using a tangram puzzle (Jennett, et al., 2008). For this experiment, the post-game task chosen was a Stroop test as in Zhang and Fu’s study (Zhang & Fu, 2015). A Stroop test is a series of coloured words, where the participant has to answer the colour of the word, rather than the colour the word says. It’s an easy concept to grasp, but it’s visual dissonance will mean participants will have to focus on the task. Results were split into congruent and incongruent average times. Congruent answers mean the colour and word matched, whereas incongruent means the colour and word did not match. Scores were given based on a combination of culminate response time and accuracy.

Game Device. The device used to play the games on was the Nintendo Switch games console. Due to being portable, this console allowed the experiment to be conducted in a number of different locations without having to compromise the size of the screen. While the different locations used might affect the overall immersion, and the Switch screen is small, having the screen size stay the same was most important control variable. In combination with this, the same set of AKG Y50 Headphones were used at 50% volume for all participants.

Game Used. The game played by all participants was *Celeste* (Thorson, 2018) – because of the accolades soundtrack had achieved, and because it has the option to remove music while keeping diegetic sound intact. The soundtrack has had acclaim from many prominent gaming outlets, Kotaku saying, “The soundtrack complements the games colourful pixel art” (Alexandra, 2018) and Imagine Games Network saying, “Celeste has one of the best game soundtracks I have heard in years” (Marks, 2018). In this case, the soundtrack is known as being well designed, so it follows that it should be able to

increase immersion. None of the participants had previously played this game.

The Stroop Test. The Stroop test was conducted from an online website (Stroop Test Online.com, 2005-2018), able to be accessed quickly in whatever location I was in. The participants were required to do this test on the same MacBook Pro 15” model.

Procedure. The Switch game *Celeste* (Thorson, 2018) was either set to full music or no music, with diegetic sound left intact for both conditions. The level began at the same point each time the experiment was carried out, a new save file being made each time. Participants were briefed on the controls and then were signalled to begin at the same time as a timer being set (which they did not know the duration of). The participants then played uninterrupted for ten minutes until the timer was finished and were asked to immediately put down the console and to take off their headphones.

They were then asked to guess approximately how long they had been playing for (in integers of a minute) and shortly after this completed the Stroop test. The participants then had to complete the questionnaire in whatever time they saw fit and were able to ask questions of clarification during this period. The participants then repeated the experiment with the opposite condition to which they started in.

3. RESULTS

The average immersion score music was 98.62 out of a possible 130, compared to an average immersion score without music of 83.33. This shows a substantially higher score in perceived immersion for the music condition. Furthermore, it is worth pointing out that the range of immersion scores was large across the participants as a whole. Immersion scores with music ranged from 56 to 117 and without music ranged from 47 to 114. Only one of the participants (11.11%) taking part in the experiment scored higher on their immersion score on the no music condition.

Table 1. Mean scores on immersion questionnaire divided whether there was music.

Condition	Mean Score	Standard Deviation
Music	98.67	20.39
No Music	83.33	23.69

The average time elapsed guessed was lower on the music condition at 12.78 minutes compared to the 13.89 minutes for the no music condition. The average times guessed for both conditions were higher than the 10 minutes of actual play. This difference is not as substantial as the questionnaire

results but is still noticeable. Out of the participants, three guessed the same time on each play through, the majority guessing a shorter time on the music condition (55.56%) and only one guessing a shorter time on the no music condition (11.11%).

Table 2. Mean time guessed in minutes divided by whether there was music

Condition	Mean Score	Standard Deviation
Music	12.78	4.6
No Music	13.89	3.98

The results from the Stroop test show worse performance on the no music condition with the average time guessed for congruent results 19.12 seconds, and average incongruent time being 16.31 seconds. For the music condition, the average congruent time was 19.54 and the average incongruent time was 17.17. While the worse performance matching higher immersion suggests a link, the difference of a less than a second should throw that idea into doubt. However, with a larger sample size this difference would have the possibility to become more defined.

Table 3. Mean score in seconds on online Stroop test

Condition	Congruent Response	Incongruent Response
Music	19.54	17.17
No Music	19.12	16.31

4. DISCUSSION

What can be said of these results, therefore, is that they suggest that the soundtrack of *Celeste* (Thorson, 2018) has a positive effect on the level immersion experienced while playing. The large difference in the average immersion scores indicate that at least the subjective experience of immersion is increased by the presence of music. This can only be seen as a positive thing for designers of game soundtracks.

What of the objective data recorded that was meant to link with these questionnaire scores however? The difference in perceived elapsed time was small, but recognisable. Although it wasn't as substantial as it could have been, it falls in line

with previous studies (Sanders & Cairns, 2010) (Zhang & Fu, 2015) which have also reported this. The difference in post-game task performance was even smaller, while not casting doubt on the immersion scores themselves, does indicate that immersion is not as immediately able to be measured through more objective means. However, with a larger sample size, there might start to be a more noticeable difference.

This has been found previously by Zhang and Fu (2015) who, using the Stroop test like this study, were unable to garner statistically significant results. This highlights a further issue of finding suitable objective measurements that are able to measure the experience of immersion. Grimshaw, Lindley, and Nacke (2012) found this to be an issue when measuring physiological responses, that the “tonic measurements of physiological response from an accumulated game session were not significantly affected by different sound modalities” (2008). Not finding much difference in the different Stroop test performances therefore is not an unusual thing to find, and perhaps is testament to immersion is not being a concrete, measurable concept. Even further that immersion is made up of many different independent elements.

Looking at the questionnaire scores alone agrees with the first hypotheses in saying that a musical soundtrack increases the experience of immersion when playing a game. At least in the case of *Celeste* (Thorson, 2018) the soundtrack designed so that it is able to elevate the experience of immersion. This sheds further light on the inconsistencies of results found in Sanders and Cairns study (2010). According to their study, to be immersive, the soundtrack of *Celeste* must at least have been ‘likeable’. The factor of ‘likeability’ is one that I would consider measuring in a subsequent study.

However, the results garnered by Grimshaw Lindley and Nacke (2008) seem to disagree with what my study has found, as they found music on its own did not increase the amount of immersion experienced by players. This could be for a number of reasons related to both the game and the soundtrack in question. The games used themselves are totally different - their study uses a first-person shooter game, while mine uses a 2D platformer¹. A study by Kristine Jørgensen shows that music affects the way people play in different ways depending on the genre (2008). This only further indicates music and immersion is not so easily understood by one study alone, and results can vary depending on genre or effectiveness of the soundtrack.

Limitations and scope for a future study. Past studies (Sanders & Cairns, 2010) (Zhang & Fu, 2015) have had more noticeable results than this study, in regard to measuring the time distortion. I believe that this is in part down to my use of the repeated measure model, rather than the individual measure model. To avoid confusion each instance of the game task was timed to 11 minutes, however, because the players played it twice, there was an increased chance of them either getting better at guessing the second time, or in my case, an

increased chance of them guessing the same time twice. Those who guessed the same time for both music and no music, might have done so because they were assured in the initial guess, or did not want to get the question wrong (even if I assured them that the two game play periods could have different timings).

If I had more time and resources I would have attempted using music not composed specifically for the game as an added condition. This could prove illuminating in deciphering whether it is a well-crafted soundtrack that improves immersion – or the presence of music alone. Sanders and Cairns’ study (2010) would suggest the former – their immersion experiment failed to produce the expected results when the music was not liked. In this way, the music in my chosen game, *Celeste* (Thorson, 2018), can be judged as successful, but would I get similar results from a completely different soundtrack? An added complication to this is games which use both diegetic and non-diegetic musical soundtracks – how would these parameters map onto the experience of immersion in games?

There is certainly scope for more empirical research exploring the relationship between music and immersion, especially given the broad nature of what can be tested. A study incorporating multiple genres, demographics and platforms would be most ideal to further illuminate this field of study.

REFERENCES

- Alexandra, H. (2018, January 25). *Kotaku*. Retrieved from <https://kotaku.com/celeste-the-kotaku-review-1822420606>
- Collins, K. (2008). *Game Sound*. Massachusetts: MIT Press.
- Collins, K. (2013). *Playing with Sound - A Theory of Interacting with Sound and Music in Video Games*. Massachusetts: MIT Press.
- Ermí, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. *Changing Views: Worlds in Play. Selected Papers of the 2005 Digital Games Research Association's Second International Conference* (pp. 15-27). DiGRA.
- Grimshaw, M. (2012). Sound and player immersion in digital games. In T. Pinch, & K. Bijsterveld (Eds.), *The Oxford Handbook of Sound Studies* (pp. 347-365). Oxford: Oxford University Press.
- Grimshaw, M., Lindley, C. A., & Nacke, L. (2008). Sound and immersion in the first-person shooter: mixed measurement of the player's sonic experience. *Games Computing and Creative Technologies: Conference Papers* (pp. 1-7). Bolton: UBIR: University of Bolton Institutional Repository.
- Jørgensen, K. (2008). Left in the dark: Playing computer games with the sound turned off. In K. Collins (Ed.), *From Pac-Man to Pop Music. Interactive Audio in Games and New Media* (pp. 163-176). Farnham: Ashgate.
- Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *Int. J. Human-Computer Studies*, 66, 641-661.

¹ A 2D perspective game wherein a player character jumps between platforms from one side of the screen to the other.

- Marks, T. (2018, January 25). *Ign*. Retrieved from <http://uk.ign.com/articles/2018/01/25/celeste-review>
- Sanders, T., & Cairns, P. (2010). Time perception, immersion and music in videogames. *BCS HCI 2010*. Dundee: ACM Press.
- Stroop Test Online.com. (2005-2018). *Stroop Test Online*. Retrieved March, April 2018, from The Stroop Effect Test: http://www.onlinestrooptest.com/stroop_effect_test.php
- Thorson, M. (Director). (2018). *Celeste* [Motion Picture].
- Zhang, J., & Fu, X. (2015). The influence of background music of video games on immersion. *Journal of Psychology & Psychotherapy*, 5(4), 1.