

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/355845214>

Music and Time Perception in Audiovisuals: Arousing Soundtracks Lead to Time Overestimation No Matter Their Emotional Valence

Article in *Multimodal Technologies and Interaction* · October 2021

DOI: 10.3390/mti5110068

CITATIONS

0

READS

149

4 authors:



Alessandro Ansani
Sapienza University of Rome

24 PUBLICATIONS 52 CITATIONS

[SEE PROFILE](#)



Marco Marini
Italian National Research Council

8 PUBLICATIONS 33 CITATIONS

[SEE PROFILE](#)



Luca Mallia
Italian University of Sport and Movement "Foro Italico"

94 PUBLICATIONS 1,268 CITATIONS

[SEE PROFILE](#)



Isabella Poggi
Università Degli Studi Roma Tre

183 PUBLICATIONS 3,894 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Other stuffs [View project](#)



Risky behavior and aging [View project](#)



Article

Music and Time Perception in Audiovisuals: Arousing Soundtracks Lead to Time Overestimation No Matter Their Emotional Valence

Alessandro Ansani ^{1,2,*} , Marco Marini ^{1,3}, Luca Mallia ⁴ and Isabella Poggi ²

¹ Department of Psychology, Sapienza University of Rome, 00185 Rome, Italy; mar.marini@uniroma1.it

² Cosmic Lab, Department of Philosophy, Communication, and Performing Arts, Roma Tre University, 00146 Rome, Italy; isabella.poggi@uniroma3.it

³ Institute of Cognitive Sciences and Technologies (ISTC), 00185 Rome, Italy

⁴ Department of Movement, Human and Health Sciences, University of Rome, Foro Italico, 00135 Rome, Italy; luca.mallia@uniroma4.it

* Correspondence: alessandro.ansani@uniroma1.it

Abstract: One of the most tangible effects of music is its ability to alter our perception of time. Research on waiting times and time estimation of musical excerpts has attested its veritable effects. Nevertheless, there exist contrasting results regarding several musical features' influence on time perception. When considering emotional valence and arousal, there is some evidence that positive affect music fosters time underestimation, whereas negative affect music leads to overestimation. Instead, contrasting results exist with regard to arousal. Furthermore, to the best of our knowledge, a systematic investigation has not yet been conducted within the audiovisual domain, wherein music might improve the interaction between the user and the audiovisual media by shaping the recipients' time perception. Through the current between-subjects online experiment ($n = 565$), we sought to analyze the influence that four soundtracks (happy, relaxing, sad, scary), differing in valence and arousal, exerted on the time estimation of a short movie, as compared to a no-music condition. The results reveal that (1) the mere presence of music led to time overestimation as opposed to the absence of music, (2) the soundtracks that were perceived as more arousing (i.e., happy and scary) led to time overestimation. The findings are discussed in terms of psychological and phenomenological models of time perception.

Keywords: soundtrack; film music; audiovisual; time estimation; time perception



Citation: Ansani, A.; Marini, M.; Mallia, L.; Poggi, I. Music and Time Perception in Audiovisuals: Arousing Soundtracks Lead to Time Overestimation No Matter Their Emotional Valence. *Multimodal Technol. Interact.* **2021**, *5*, 68. <https://doi.org/10.3390/mti5110068>

Academic Editor: Insook Choi

Received: 10 September 2021

Accepted: 26 October 2021

Published: 29 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Music is considered to be effective in influencing a broad variety of cognitive and affective mechanisms, from work performance [1] to memory tasks [2] and learning [3], from economic decision making [4–6] to advertising [7], moral judgement [8–11] and prosocial behavioral intentions [12], and also personal enhancement [13].

In the tradition of studies conducted regarding music's influence, a promising thread of the research focused on time perception, starting from the seminal work of Rai [14]. As we will see, aside from the research traditionally focusing on music (i.e., perceived duration of musical excerpts), some work has pointed at the domain of audiovisual stimuli [15] and film-induced mood [16], with some interest in the dynamic interaction between the auditory (i.e., musical) and visual elements [17]. For example, it has been found that music is capable of modulating and altering visual perception [18–20] due to phenomena such as auditory driving [21].

In this study, we are interested in the modalities through which music affects time perception. We firmly believe that this peculiar research theme has the merit of meeting the needs of a growing number of scientists, artists, and professionals who are

interested in shaping the interaction between users and audiovisuals with some background music, whether in movies [22], educational videos [23], interactive games [24], or videogames [25,26]. For instance, there may be some utility in decreasing the self-perceived passage of time in interactive educational or tutorial contexts, especially considering that the majority of educational or tutorial videos have background music within. In such contexts, the challenge is to reach a balance between the entertaining effects of music that improve learning [3,27] and its distracting properties [28] that dampen attention. We are confident that a proficient management of the background music could act upon the recipients' time perception, thus improving the interaction between the users and the audiovisual devices at hand.

Before discussing the existing tradition of research, a clarification is needed on the assessment of time estimation. As suggested by [29], in the literature, two main paradigms exist for the study of time estimation: prospective and retrospective; the former implies that the participants are aware that, after the presentation of a given stimulus, there will be questions on the perceived duration of elapsed time, thus studying the experienced time (i.e., the subject attends to the passage of time itself), whereas the latter does not have such an implication, that is, the experimental participants are not informed about the questions that will follow concerning time perception, thus analyzing the remembered time (i.e., the subject's attention is not focused on time perception) (for a detailed discussion of the different cognitive processing of prospective and retrospective time judgement, see also [30–32]). In the current work, as we are interested in analyzing the soundtrack's influence on the time estimation of an audiovisual experience, we focus on the retrospective paradigm (i.e., remembered time). We opted for this paradigm because we wanted our experimental viewers to be completely unaware of the task at hand. In other words, we wanted to avoid conscious time counting.

To begin with, in Section 2, a summary of previous studies on the relationship between music and time perception is presented, alongside a focus on the musical parameters that have been proven to perform a role in affecting time perception in several contexts. In Section 3, more attention is devoted to two psychological models (i.e., Dynamic Attending Theory and Scalar Expectancy Theory) that attempt to explain how time perception works within the audiovisual domain. In Section 4, we present our online experiment, which we discuss in Section 5, where we also list some of the limitations of this work and a few suggestions for future research. A brief conclusion is presented in Section 6.

2. Previous Works on Music and Time Perception

A variety of studies focused on how music alters our subjectively perceived time in an indirect fashion, for instance, considering the waiting times [33] in retail settings [34], restaurants [35], queue contexts [36,37], and on-hold waiting situations [38].

North and Hargreaves [33] compared the waiting times of four groups of participants who were waiting for an experiment to begin; three of the groups were provided with background music differing in complexity, while the last group was not provided with music (control condition). They found no differences between the music conditions, but the controls showed a significantly lower waiting time. Areni and Grantham [39] reported that when waiting for an important event to begin, their participants tended to overestimate the waiting time when they disliked the background music that was playing, whereas they underestimated the waiting time in the presence of background music they liked. Fang [35] found that slow-paced background music extended the customer waiting time in a randomly selected restaurant, whereas fast-paced background music shortened the waiting time of customers, i.e., they decided to leave earlier. Guéguen and Jacob [38] shed further light on the issue by analyzing the cognitive mechanisms that come into play in an on-hold telephone scenario. Their results proved that in comparison with the no music condition, the simple presence of music led to both an underestimation of the time elapsed and an overestimation of the projected time passed before a person would hang up. Finally, the most up-to-date meta-analytic review of the effects of background music

in retail settings [34] (p. 761) concludes that, “A higher volume and tempo, and the less liked the music, the longer customers perceive time duration. Tempo has the greatest effect on arousal.”

In most of these studies, the dependent variables are indirect measures of time perception because participants do not explicitly report their time awareness, but their behavior is simply annotated. In some other cases [40,41], an actual self-report of the wait length was assessed.

Evidence that music alters the representation of time also stems from qualitative research on altered states of consciousness (ASCs) [42]. In such studies, subjects’ reports often mention feelings of timelessness, time dilation, and time-has-stopped in correspondence to music listening activities [43].

To sum up, there exists a consensus that music has a robust role in influencing how we perceive time [15,25,37,40,44–46], although fostering small-to-moderate effects, as Garlin and Owen [34] pointed out in their meta-analytic review.

On the contrary, less consensus exists on the musical parameters that are responsible for the alteration of time perception. Research with both direct and indirect measures of time estimation has focused on diverse music parameters.

Musical Parameters and Time Perception

Several studies have investigated time estimation in the dependence of several music parameters and types of music, primarily by assessing the perceived length of musical excerpts.

To begin with, the musical structure complexity has been found to increase the time estimation [47]; on the contrary, the results on tempo are not consistent; if [44] found no evidence, other works proposed that a slower tempo seems to lead to time underestimations [37,48,49]. Coherently, [37] found temporal perception (perceived minus actual wait duration) to be a positive function of musical tempo. In relation to musical modes, a study [50] proved that the Locrian mode (diminished, thus more likely to be unpleasant) led to time overestimation as opposed to the Ionian and Aeolian modes. The modes of ancient Greece can be described as a kind of musical scale coupled with a set of particular melodic behaviors. The Ionian mode is equal to modern-day major scale. The Aeolian mode is today’s natural minor scale (i.e., 1, 2, b3, 4, 5, b6, b7). The Locrian mode is a minor scale with the second and fifth scale degrees lowered a semitone (i.e., 1, b2, b3, 4, b5, b6, b7). Arguably, the Locrian mode tends to yield more negative valence than the Minor due to the different composition of the tonic triads (i.e., the principal chord that determines the tonality of the mode). When compared with the Minor mode, wherein the tonic triad is constituted by a minor 3rd and a perfect 5th (i.e., minor chord), the Locrian mode’s tonic is composed of minor 3rd and diminished 5th (i.e., diminished chord). Because of their composition, the diminished chords are considered dissonant and prove to be responsible for conveying the lowest valence and very high tension [51].

The music volume also might play a role; indeed, [45] proposed that people listening to quieter music tend to underestimate the time passed. Lastly, [52] reported that an overestimation of passed time was observed for pop music played in a major vs. minor mode, while [53] concluded that listening to familiar music leads to an underestimation of passed time.

3. Time Perception in Audiovisuals—Models and Mechanisms

When it comes to audiovisuals (i.e., complex multimodal stimuli that present at least two channels of information: auditory and visual), it can be assumed that different mechanisms come into play, especially since the viewer engages with a conscious elaboration of the overall stimulus by integrating its various interacting parts. The integration process can be easy or difficult depending on the stimulus’ internal congruency, and this can, in turn, impact the time estimation. For instance, as elaborated by [54] (p. 504), “Because of its effects on information processing, stimulus congruity may influence the retrospective estimation of event duration. Specifically, underestimation of lapsed time might be expected

when the elements comprising the event are incongruent, because incongruent information tends to be more difficult to encode and retrieve because of the absence of a preexisting cognitive schema [. . .], and because of weaker linkages between unrelated nodes in an associative network. [. . .] exposure to incongruent information, like elevated arousal states, may create a distraction that reduces attention to one’s internal “cognitive timer [55]”.

In their recent review on time perception in audiovisual perception, Wang and Wöllner [15] clarify that two Internal Clock models can account for the effects of music on time perception in the audiovisual context, which are the Dynamic Attending Theory (DAT), also known as the oscillator model [56], and the Scalar Expectancy Theory (SET), also referred to as pacemaker-counter model [57]. Contrarily to the SET, which postulates regularly emitted pulses by an independent internal clock, the DAT claims that the time estimation of the duration of past events depends on the coupling between attentional pulses and the occurrences of external events [56]. For this reason, this model is sometimes referred to as an attention-based model. Crucial to this theory is the idea that the emission of attentional pulses or oscillations is a non-linear (i.e., dynamic) process, and that the attention regulates the pulses to a greater extent than the working memory (on the relationships among time perception, attention, and working memory, see [32]). Indeed, the adjective “dynamic” stems from the fact that, contrary to the SET model, the emission of the attentional pulses is not a static process, rather, it varies depending on the salience of the external events (i.e., stimuli).

The DAT model also suggests that when the attention toward a stimulus is low, a fewer number of pulses is emitted, thus leading to an underestimation of time. Furthermore, other results agreeing with the DAT model indicate a positive correlation between musical tempo and time estimation in both auditory [58–60] and visual perception [61].

Conversely, the latter model (i.e., Scalar Expectancy Theory) proposes a linear cumulation of regularly emitted attentional pulses and a pacemaker that counts them. According to this memory-based theoretical framework, more akin to that of [62], a more important role is assigned to the working memory as it is a three-step model in which the memorization constitutes the second and central phase, following the clock phase and followed by the time judgement. In particular, the storage-size model of memory of time perception [62] claims that richer stimuli (i.e., with a high amount of information) lead to the perception that a greater number of events occur in a given interval, thus favoring time overestimation.

To sum up, among the models that account for the influence of music on time perception, two share the hegemony: the attention-based (DAT) and the memory-based (SET and storage-size model of memory of time perception). Nevertheless, various results exist within the literature because various studies claim that different features of the stimuli influence time estimates (Table 1).

Table 1. Factors influencing time perception.

Factor	Effect	Reference
Attention	overestimation	[63]
Engagement	overestimation	[31]
Arousal	overestimation	[16,59,64]
Arousal	underestimation	[65]
Negative emotions	overestimation	[66]
Positive emotions	underestimation	[66]
Music familiarity	underestimation	[53]
Liked vs. disliked music	underestimation	[39]
Fast musical tempo	overestimation	[58,60]
Slow musical tempo	underestimation	[37,48,49]
Volume	overestimation	[45]
Pitch and metrical variations	overestimation	[67]

Table 1. *Cont.*

Factor	Effect	Reference
Musical structure complexity	overestimation	[47]
Stimulus complexity	overestimation	[62]
Locrian mode (vs. Ionian and Aeolian)	overestimation	[50]
Major mode	overestimation	[52]
Minor mode	underestimation	[52]

As we are interested in the time estimation of an audiovisual, we focus our attention on two basic parameters, namely valence and arousal of the emotions conveyed by music, for two reasons:

1. It is known that certain pieces of music can, through their emotional valence, foster positive affective states, to the point that music has traditionally been considered as a valid mood inductor [68]. Therefore, in accordance with previously collected results from outside of the audiovisual domain [66], we can hypothesize that the positive affect experienced by the recipients while viewing may be negatively correlated with the estimation of the time elapsed [31], that is, the better the viewers feel as they watch the scene (i.e., positive affective state), the less they perceive the passing of time.
2. A great deal of research suggests that the arousal (i.e., the physiological and psychological state of activation) conveyed by music might lead to time overestimation [34,54], possibly due to an effect on the internal clock system speed (both in attention- and memory-based models of time perception). Nevertheless, no one, to our knowledge, has ever shown such a phenomenon in an audiovisual domain.

These two points underpin our research questions, which are introduced in the following section.

4. The Present Study

As stated above, in the literature on the influence of music on time perception, one cannot draw definitive conclusions about a variety of factors. With this study, we aim to investigate how the perceived length of a visual scene is affected by the background music (i.e., soundtrack), and, more specifically, the following two particular features conveyed by the music: emotional valence and arousal.

4.1. Research Questions

We hypothesize that both the emotional valence and arousal conveyed by music have a key role in the time estimation of an audiovisual piece, although in contrasting ways.

First, coherently the studies on waiting times [38], we expect the mere presence of music to lead to a decrease in the perception of the elapsed time (i.e., time underestimation) (Hypothesis 1).

Secondly, considering the literature on music pieces [66], we expect positively valenced music to result in time underestimation, and negatively valenced music to induce time overestimation (Hypothesis 2).

Third, in accordance with both the attention and memory-based models of time perception, we hypothesize that the arousal level should lead to time overestimation (Hypothesis 3).

Below, we describe the experimental paradigm, each construct, and its related measurement separately.

4.2. Method

We designed a between-subjects experiment wherein the participants watched a modified version (01'30") of a short movie by Calum Macdiarmid [69] (Figure 1).

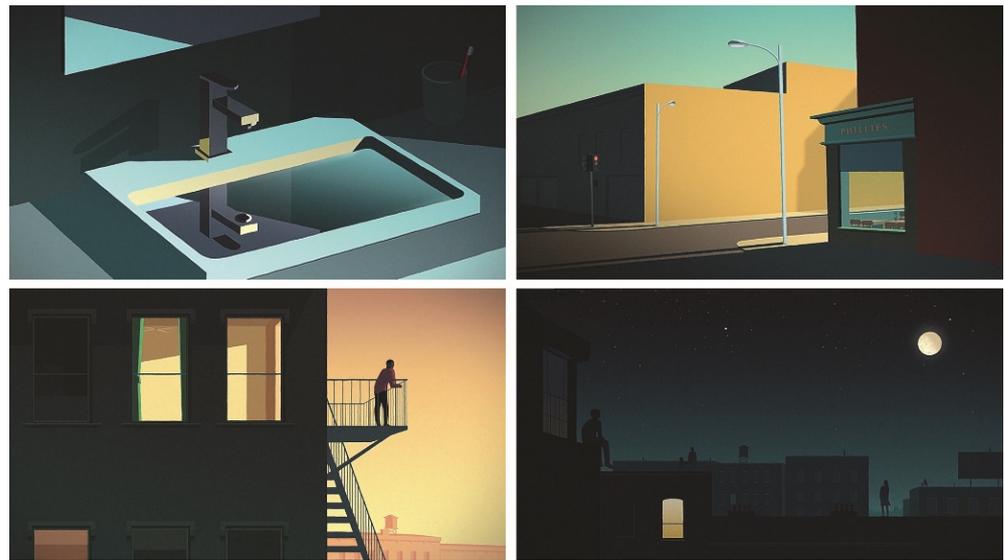


Figure 1. Illustration of four representative frames of the scene.

Using Reaper 6.29, we created five versions of the short movie—varying under the five experimental conditions—with the video accompanied respectively by a happy piece (*Appalachian spring—VII: doppio movimento*) by A. Copland), a sad cello melody accompanied by a piano, (*After Celan* by D. Darling and K. Bjørnstad), a frightening track from the Original Motion Picture Soundtrack of the film *Proxy* (*Murder* by the Newton Brothers), a relaxing piece specifically composed to control anxiety [70], or by no music at all (i.e., control condition). This method allowed us to present all the possible combinations of valence and arousal (Table 2).

Table 2. Soundtracks’ emotional valence and arousal.

Track	Emotion	Valence	Arousal
Appalachian spring (VII: doppio movimento) [A. Copland]	happiness	+	+
After Celan [D. Darling and K. Bjørnstad]	sadness	-	-
Murder [Newton Brothers]	fear	-	+
World’s most relaxing music [R. Wiseman]	relaxation	+	-

Similarly to [20], the four pieces were chosen by considering the findings of [71], and the subsequent studies enumerated by [72] concerning a plethora of psychoacoustic parameters associated with emotional expression in music. Two of the pieces evoked negative affects but differed in the arousal dimension: the *After Celan* track’s soft tone and morbid intensity fosters sadness and tenderness [73]. Conversely, the Newton Brothers’ track’s great sound level variability and the rapid changes in its sound level could be associated with the experience of fear [71], while its increasingly louder volume can evoke restlessness, agitation, tension [74] or rage, fear [75] and scariness [76].

In a similar way, the two other pieces both foster positive feelings, but with a marked difference in the arousal dimension: if Copland’s piece’s orchestration, fast tempo, and high pitch all cultivate a sense of highly exciting joy, the relaxing piece was specifically composed with the goal of controlling anxiety. To elaborate, it presents a relatively constant volume, narrow melodic range, legato articulation, and regular beat [70]. To mitigate any loudness perception effects, the perceived loudness of all the tracks was normalized via a Loudness, K-weighted, relative to Full Scale (LKFS) [77].

We also aimed at ecological validity; thus, in order to allow people to participate in a less detached situation than a lab, we built an online procedure on Qualtrics.com. The participants accessed a single-use link (An anti-ballot box stuffing was employed to avoid multiple participations from the same device) through which they could run the experiment. As a result of the online procedure, they were able to participate directly from home on their laptops, smartphones, or tablets, just as if they were watching an actual movie. An introductory screen summarily presented the task to the participants without mentioning the question about the time estimation. Immediately after this introductory screen, the informed consent statement was presented. After viewing the scene, a questionnaire was administered with three questions: the first two, which might be considered as a manipulation check, were designed to verify whether the emotional valence and arousal self-reported by participants were the same as those expected for each music condition. The last question aimed to assess the dependent variable, namely the participants' perception of elapsed time (i.e., time estimation). To avoid sequence effects (i.e., the theoretical possibility that a previous question could affect the following one in any possible way), the order of questions was completely randomized for each participant.

4.3. Measures

4.3.1. Affective States of the Recipients

To measure the affective state of the viewers, we needed to identify what we might call the emotional nuclei of the viewing session and the emotional nuances each soundtrack could add to the narration. To this aim, as we were interested in a fast and immediate answer that caught the gist of the emotional content of the vision, we decided against a Likert scale with several emotions as the items, because this would have resulted in increased fatigue for the recipients. On the contrary, we resorted to Plutchik's wheel of emotions [78]. In brief, we presented our participants with the image of the wheel (Figure 2), asking them to select with a click the region that best represented the emotion they were experiencing while viewing the video.

4.3.2. Arousal

To assess emotional arousal, we used a 100-point slider, asking our participants how active they felt while viewing the scene. The slider was initially set to 0; the recipients were required to place it at their desired point. As it would have been suboptimal to use a single adjective to refer to the concept of arousal unambiguously, in this assessment, we provided a note in the question to our participants that read: "When we say active, we also mean awake or ready".

4.3.3. Time Estimation

We asked our participants to indicate the length of the video by dragging a slider that ranged between 60 and 120 s (i.e., minimum and maximum values admitted); the slider was initially placed at the center of the bar (i.e., 90 s). Later, as was the case with [37], we created a measure of the gap between the estimated time and the actual time, according to the formula:

$$\Delta\text{Time estimation} = \text{Selfreported time} - \text{Actual time elapsed (90 s)}$$

In this way, a positive number indicates an overestimation of time, whereas a negative number represents an underestimation.

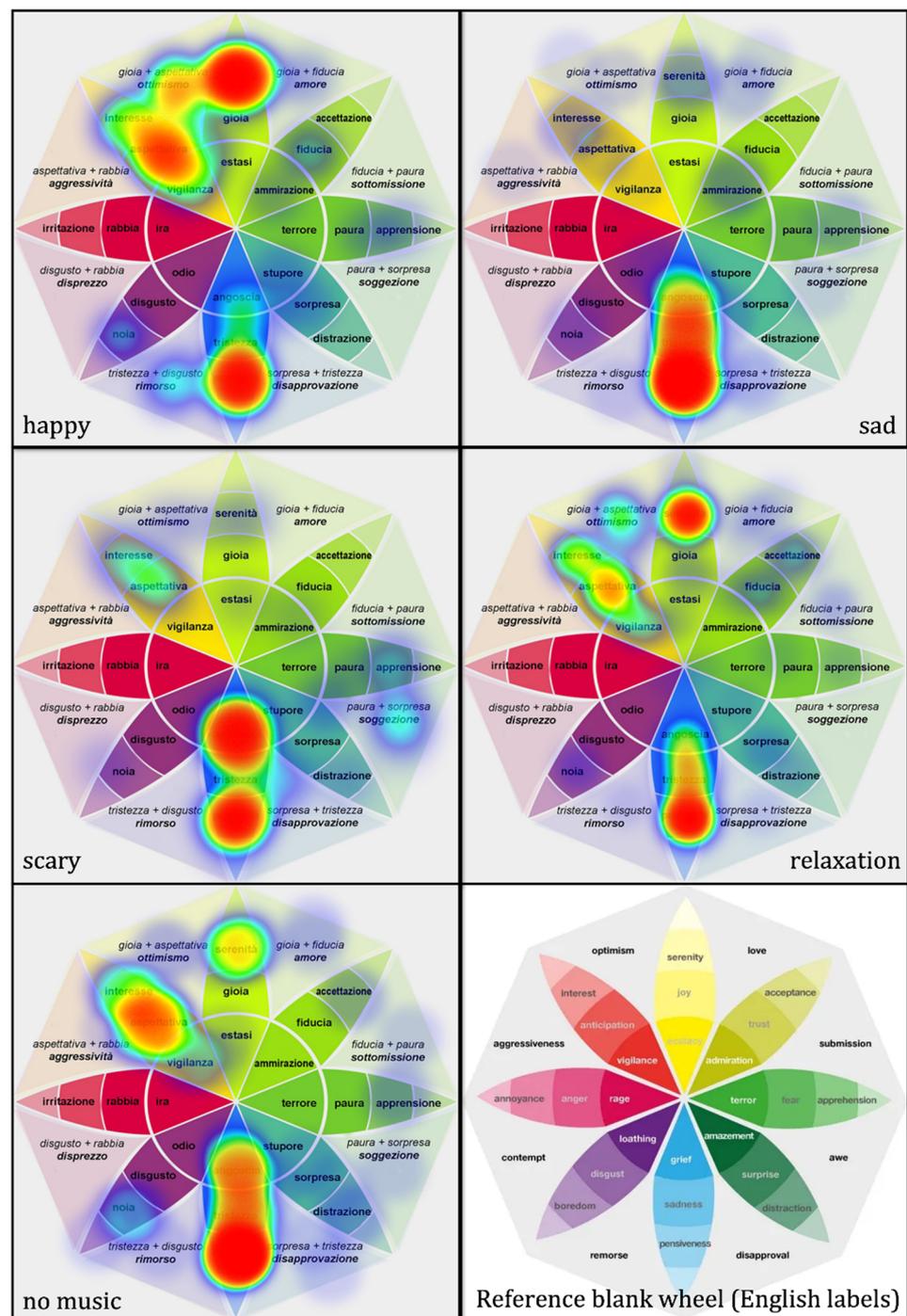


Figure 2. Heatmaps of the participants’ affective state in each condition. Heatmaps visualize, in an aggregated fashion, the most frequently clicked (hot) and unclicked (cold) emotions using colors on a scale from red to blue. At the bottom right, a blank wheel with English labels is presented to facilitate comparisons.

4.4. Participants and Preliminary Sample Data Analysis

As a first step, six hundred and three ($n = 603$) Italian participants were recruited by sharing the link of the study on social media and through university mailing lists (i.e., snowball procedure). Their participation was provided on a voluntary basis, and the participants were not incentivized with any reward.

Before our data analysis, to improve the reliability of our sample, we performed exclusions based on the following pre-established criteria:

- An attention check question in which a short Likert scale was presented with the explicit instruction that asked participants to avoid completing it; we excluded all those participants who completed such a scale.
- A time counter on the screen displaying the video was incorporated (it was visible to the experimenters only) so as to exclude all participants who had not watched the whole video (i.e., time spent on that screen < 90 s).
- All those participants who completed the task in less or more than the mean duration \pm 3SD were excluded.
- All participants who did not complete the questionnaire in all its parts were also excluded.

After the above exclusions, our sample size decreased from 603 to 565 valid participants (mean age = 26.01 SD = 10.53, 339 females, 60%). The five experimental groups were comparable in the number of participants (range 104–119) and were gender-balanced ($p = 0.41$).

4.5. Results

For the statistical analyses, IBM SPSS 26.0 was used; the path analysis was processed through Mplus 8.5 [79]. The violin plots were made by means of R (ggplot2 package). For each test, the effect size is provided by employing η (eta squared, for chi-square and ANOVA statistics). In the ANOVA tests, the post-hoc computed observed power is provided in terms of $(1-\beta)$. In the results of the model (Section 4.5.2), for each path, we provide the standardized path coefficient (β), the relative Standard Error (S.E.), the level of statistical significance (p value), and a 95% Confidence interval (95% CI). In the case of the indirect effects, 95% Bias-Corrected Confidence Intervals are indicated (BCa).

4.5.1. Affective States of the Recipients

The heatmaps of Figure 2 provide a first and intuitive point of view of the participants' affective states. A common emotional nucleus emerges in all conditions, specifically the bottom region of Plutchik's wheel of emotions, which is the axis that includes pensiveness, sadness, and grief. It is worth mentioning that the other soundtracks add or subtract diverse emotional nuances in comparison with the control condition. For instance, comparing the controls with the happy group, the region of the serenity/joy becomes more populated. When considering the scary condition, the serenity/joy axis loses relevance, while the expectancy area remains active, and apprehension and awe gain saliency. Conversely, when considering the sad condition, all the other axes aside from the pensiveness/sadness/grief axis become unnoticeable.

Upon further analyses, considering that our participants simply clicked once on the Plutchik's wheel of emotions image in correspondence with the emotion they were feeling, we created an emotional score by assigning 1 point to the participants who chose a positively valenced emotion (21.9%), 0 points to non-valenced emotions (expectation, interest, surprise, and distraction, 17.3%), and -1 point to negatively valenced emotions (60.7%). We then performed a chi-square test to evaluate the distribution of the emotion valence in dependence of the condition, finding it to be significant, $\chi^2(8565) = 101.34$, $p < 0.001$, η condition dependent = 0.08, η aff. state dependent = 0.40 (Table 3).

Table 3. Affective states across conditions.

Soundtrack	Affective State		
	Negative	Neutral	Positive
happy	35.96%	23.68%	40.35%
relaxation	41.18%	21.85%	36.97%
no music	64.42%	21.15%	14.42%
sad	84.68%	6.31%	9.01%
scary	78.63%	13.68%	7.69%

4.5.2. Time Estimation

Before proceeding with the analysis of variance, we studied the descriptive statistics. The first aspect to consider is that the majority of participants in our sample (71.3%) underestimated the actual length of the scene ($M = -14.98$ $SD = 27.01$, $\min = -62$, $\max = 60$). We then proceeded to the verification of our hypotheses (Section 4.1).

Hypothesis 1 (H1). *Does the presence of music lead to time underestimation?*

To verify Hypothesis 1, that is, whether the mere presence of music negatively influenced time estimation, a one-way ANOVA was performed, which revealed the main effect of the music [$F(1, 563) = 6.46$, $p = 0.011$ $\eta^2 = 0.011$ ($1 - \beta) = 0.72$]. Contrarily to the hypothesis, the control group reported the video to be shorter ($M = -21.03$ $SD = 26.10$) as opposed to the music group ($M = -13.62$ $SD = 27.01$) (Figure 3). We can therefore state that Hypothesis 1 was not verified.

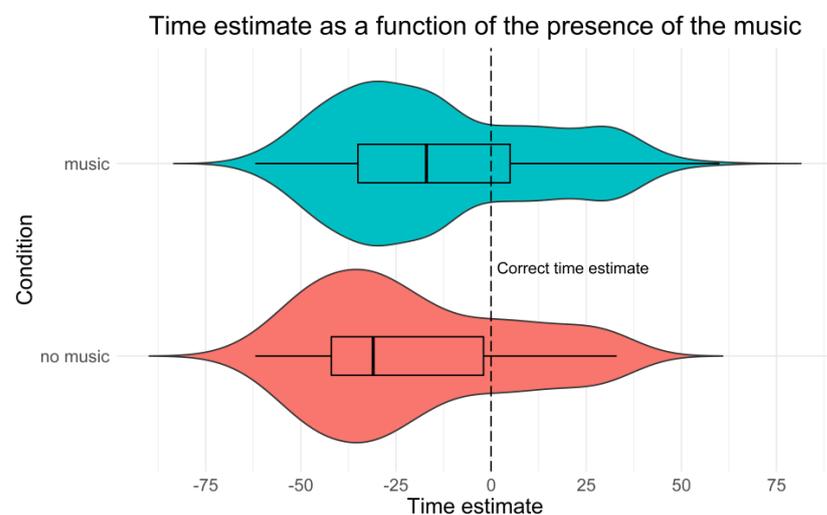


Figure 3. Time estimate (delta) as a function of the presence of the music (violin plot). The form of the violin indicates the distribution curve. The boxplots within each violin represent interquartile ranges (IQRs). Black vertical lines within the boxplots indicate median values. Values above zero represent overestimation; negative values indicate underestimation. Participants in the music conditions presented significantly higher time estimation.

After analyzing all the groups in greater detail, we still found an effect of the music [$F(4, 560) = 4.93$, $p = 0.001$ $\eta^2 = 0.034$, ($1 - \beta) = 0.96$]. Subsequent custom hypothesis contrasts revealed the significant differences against the control condition to be those of the happy ($M = -14.00$ $SD = 24.74$, $p = 0.050$), scary ($M = -7.37$ $SD = 29.08$, $p < 0.001$), and relaxation conditions ($M = -13.28$ $SD = 27.88$, $p = 0.031$) (Table 4 and Figure 4). As concerns the specific roles of valence and arousal, we resorted to a path analysis that we describe in the following paragraph.

Table 4. Time estimate (delta) as a function of the condition.

Soundtrack	Mean (s)	SD	N
happy	-14.00	24.74	114
relaxation	-13.28	28.88	119
no music	-21.03	26.10	104
sad	-20.16	23.65	111
scary	-7.37	29.08	117
Total	-14.98	27.01	565

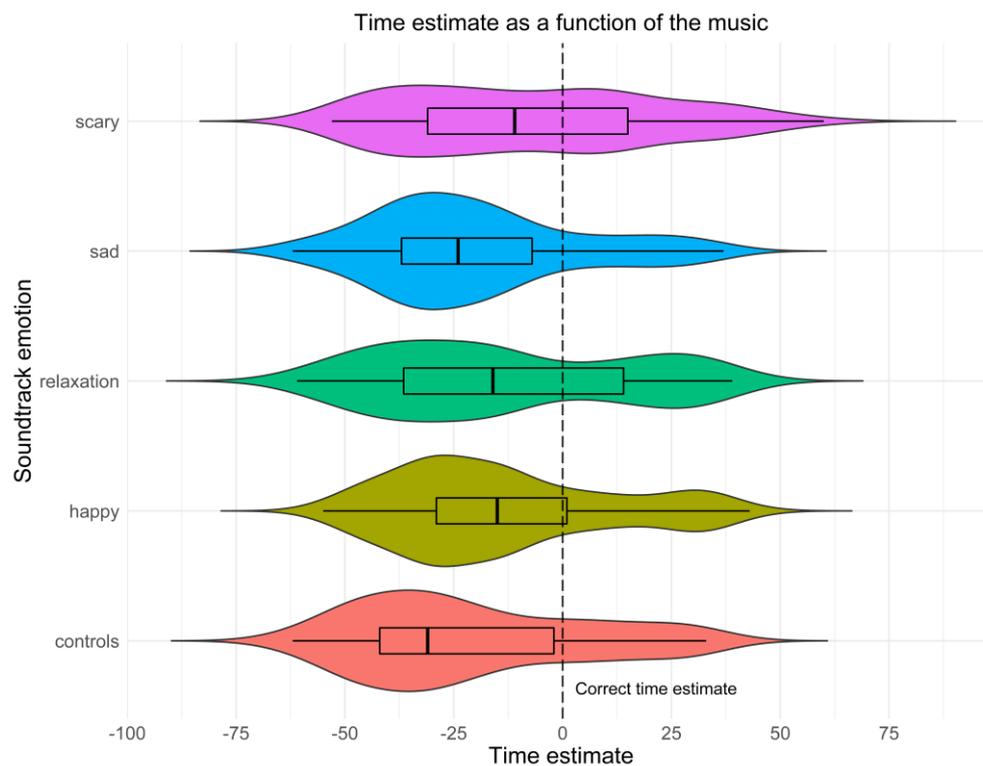


Figure 4. Time estimate (delta) as a function of the soundtrack. The form of the violin indicates the distribution curve. The boxplots within each violin represent interquartile ranges (IQRs). Black vertical lines within the boxplots indicate median values. Values above zero represent overestimation; negative values indicate underestimation. Participants in the happy, relaxation, and scary conditions presented significantly higher time estimations than in the no-music condition (custom hypothesis contrasts).

Hypothesis 2 (H2). *The role of emotional valence.*

Hypothesis 3 (H3). *Arousal in time estimation.*

As for the verification of Hypotheses 2 and 3, a path analysis was performed to analyze the role of the valence and arousal as conveyed by the music and self-reported by our participants with regard to time estimation. The model presents two exogenous variables, namely the valence and the arousal conveyed by music (i.e., the experimental conditions). Both the variables were operationalized on three levels; the valence denoted as -1 (negative valence: sad and scary), 0 (neutral valence/no music), and 1 (positive valence: happy and relaxation); and the arousal denoted as -1 (low arousal: relaxation and sad), 0 (neutral arousal/no music), and 1 (positive arousal: happy and scary). For the next step (i.e., order of the model), the endogenous variables were the self-reported affective state and arousal. The first part of our model can be considered as a manipulation check that is conducted to ensure that our participants' affective state and arousal were effectively and coherently affected by the pieces of music that we selected. Finally, the last endogenous variable was the time estimate.

To avoid normality issues, Robust Maximum Likelihood (MLR) was used as the estimator.

All the fit indices presented a good fit for the tested model with the empirical data [80]: $\chi^2(2)$ test of model fit = 3.77 ($p = 0.151$) CFI = 0.986; RMSEA = 0.040 (90% CI = 0.001–0.101); SRMR = 0.018 (see also Figure 5 caption).

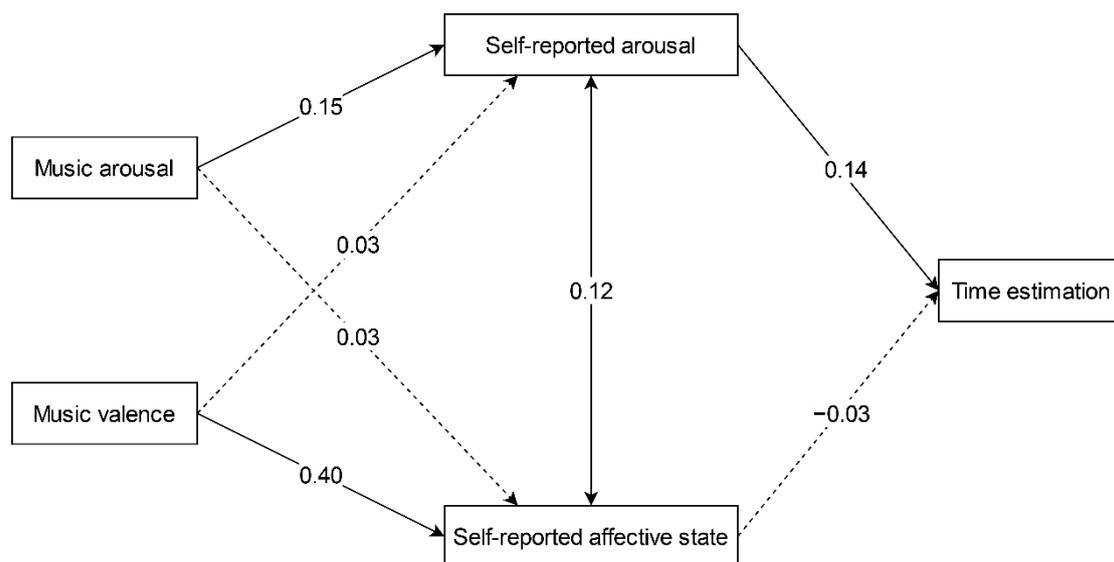


Figure 5. Time estimate (delta) as a function of the soundtrack. *Note.* Results of path analysis: χ^2 test of model fit = 3.77 ($p = 0.151$) Comparative Fit Index (CFI) = 0.986. Absolute fit indexes: Root Mean Square Error of Approximation (RMSEA) = 0.040 (90% CI = 0.001–0.101); Standardized Root Mean square Residual (SRMR) = 0.018. Parameters estimates are standardized. Dotted lines represent the insignificant relationships. Continuous lines represent paths with $p < 0.003$.

As expected, the emotional valence conveyed by the music significantly impacted the participants' affective state ($\beta = 0.403$, S.E. = 0.036, $p < 0.001$, 95% CI = 0.332–0.475) and not the self-reported arousal ($\beta = 0.035$, S.E. = 0.040, $p = 0.380$, 95% CI = -0.039 –0.112). Conversely, the arousal conveyed by the music influenced the self-reported arousal ($\beta = 0.147$, S.E. = 0.040, $p < 0.001$, 95% CI = 0.068–0.225) and not the participants' affective state ($\beta = 0.037$, S.E. = 0.039, $p = 0.341$, 95% CI = -0.039 –0.112). The two self-reported measures were scarcely but significantly, correlated ($r = 0.121$, $p = 0.002$, 95% CI = 0.043–0.199).

As hypothesized in Hypothesis 3, the self-reported arousal positively predicted the time estimation ($\beta = 0.146$, S.E. = 0.042, $p < 0.001$, 95% CI = 0.064–0.229), whereas, contrarily to Hypothesis 2, the self-reported affective state did not reach the statistical significance ($\beta = -0.030$, S.E. = 0.041, $p = 0.458$, 95% CI = -0.110 –0.049).

Lastly, the indirect effects were measured using bootstrapped bias-corrected confidence interval estimates (95% Confidence Interval with 10,000 bootstrap resamples), unsurprisingly, the total indirect effect of the emotional valence of our musical pieces on the time estimate was not significant, $\alpha\beta = -0.007$, S.E. = 0.017, $p = 0.684$, 95% BcCI = -0.041 –0.027. Rather, as further support for Hypothesis 3, the indirect effect of the musical arousal was significant, $\alpha\beta = 0.020$, S.E. = 0.009, $p = 0.018$, 95% BcCI = 0.003–0.037 (Figure 5).

5. Discussion

Firstly, our results suggest that the mere presence of music causes an increase in time estimation in an audiovisual context. This finding seems to contradict that of other studies that found that music presence, as opposed to music absence, led to longer waiting times, therefore suggesting a decrease in the perception of elapsed time (i.e., time underestimation) [33,38,81] due to the fact that music leads to perceive the time passing by as being slower. We can account for such an apparent contradiction by considering that music has a twofold nature: on the one hand, when it is reproduced in the background (as in most of the studies on waiting times mentioned above), it may be conceived as a distractor that draws the focus of attention away from the conscious time perception. On the other hand, when music is paired with a visual stimulus (as in the film music domain), it becomes a key part of the meaning of that scene, the integration processing of which requires added attentional and memory resources.

Indeed, concerning the above-mentioned models of time perception (Section 3), this effect of the presence of music might go in the direction of some memory-based phenomenon related to an added complexity. In further detail, an audiovisual stimulus requires more information to be processed than a visual stimulus alone. Not only do audiovisuals require several parallel levels of processing, such as visual, music, kinesthetic, and, possibly, speech, sound FX (i.e., sounds recorded and presented to make a specific storytelling or creative point without the use of words or soundtrack, ex.: sounds of real weapons or fire), and text [82], but they also require their coherent integration aimed at building a working narrative, namely the subjective interpretation of a scene. Such integration involves both bottom-up (sensory-perceptual) and top-down (expectative) processes: on the one hand, a recipient perceives information using their senses; on the other, one integrates this information using previous knowledge and cognitive schemas stored in the long-term memory [82]. Other studies have already revealed that, under the influence of differently valenced soundtracks for the same video, not only do the viewers generate diverse plot expectations [20,83] and alter their recall of the scene [84,85] (i.e., high-level processing), but they can also be driven and even deceived in a way that impacts their visual perception (i.e., low-level processing) [18–21]. Therefore, it can be assumed that such a to-be-processed integration, only present in the music conditions, could be the cause of time overestimation in accordance with the memory-based model of time perception.

Elaborating on the differences in the soundtracks more in detail, when considering the four music conditions separately, both the positively valenced soundtracks (i.e., happy and relaxing) and highly arousing ones (i.e., happy and scary) seemingly result in time overestimation (Figure 4).

Nevertheless, to better clarify the roles of valence and arousal, we implemented a more sophisticated path analysis that considered not just the experimental conditions but the subjectively perceived affective state and arousal level as plausible predictors of time estimation. The results of the model (Figure 5) clarified that the soundtracks' impact on the study participants was coherent with our hypotheses and, more importantly, that only the subjectively perceived level of arousal positively predicted the time estimation (in contradiction with [54,65]).

It appears that our results contrast the well-known traditional adage that “time flies when you're having fun”, or at least they correct this adage in quite a counter-intuitive fashion, that is: “Time flies when you're not activated”.

This outcome is coherent with those studies that furthered several music parameters, showing that fast musical tempi [37,58,60] and high musical structure complexity [47], all features present only within the highly arousing soundtracks, led to overestimations. Similarly, our findings also overlap with those studies that found music in major mode (widely associated with positive affect) and music in minor mode (largely associated with negative affect) do not differ in their influence on time perception [64]. Were this the case, then the music valence should have behaved as a negative predictor, given that the two positively valenced pieces were both in the major mode, whereas the two negatively valenced ones were both in the minor mode.

It is also worth noting that these findings are in contradiction with those of [66], where systematic overestimation in the judgment of the duration of joyful musical excerpts was found, and the opposite was noticed for the sad tracks. We may account for such a difference by bringing attention to two significant differences between their procedure and ours: first, although both the studies employ a retrospective paradigm, Bisson and colleagues [66] inserted a cognitive task between the two musical excerpts; thus fostering a relevant change in the participants' foci of attention that could have created a bias in the internal clock mechanism. Secondly, and most importantly, it should be considered that their results (i.e., positive valence music in major key fosters time overestimation as opposed to negative valence in minor key) might also be explained in terms of arousal. In fact, the positive valence musical piece that was used by Bisson et al. [66] (i.e., the 1st movement of Johann Sebastian Bach's *Brandenburg Concerto No. 2 in F major*, BWV 1047O) can undoubtedly be

considered to be an arousing composition, incomparably more arousing when contrasted with the negative valence musical piece that was used (Samuel Barber's *Adagio for strings in B \flat minor* from the 2nd movement of *String Quartet, Op. 11*).

As for the diatribe between the attention and memory-based models, it is worth mentioning that no safe conclusion can be drawn from the current study. The attention-based model posits that time overestimations are due to a higher number of attentional pulses emitted in highly arousing situations, whereas the memory-based model posits time overestimation to be a phenomenon caused by the stimulus complexity. The more complex the stimulus, the more processing is required, and a greater number of traces remain in memory, thus leading to overestimations. To put this in terms of the Scalar Expectancy Theory, the pacemaker regularly emits attentional pulses, but the counter device, in the presence of richer stimuli, counts an increased number of pulses. The issue here is that the two arousing soundtracks both present, apart from the faster tempi, greater perceived complexity compared to both the relaxing and sad tunes.

To disambiguate between these two differently oriented models, in future works, it could be profitable to compare, for instance, two arousing soundtracks differing in the degree of harmonic and melodic complexity (for example, a very fast bebop jazz tune with a techno track), and the same might be done with two scarcely arousing pieces.

Rather than through the attention and memory-based models, a phenomenological approach appears to be promising in explaining this and other aforementioned results. Such a phenomenological approach, promoted by Flaherty [86], has philosophical roots in the thought of Heidegger, Husserl [87], and Merleau-Ponty [88,89]. It proposes that time consciousness cannot be fully analyzed through perception because of the intrinsic nature of time, which is considered as a construction more than an objectively perceived entity. As such, no fixedly emitted pulses of sort can exist; on the contrary, our experience of the *now* rises from the integration of diverse perceivable stimuli into a single unit of content within consciousness. Yet, the number (i.e., how many of these stimuli we process) and the saliency of these stimuli vary depending on several factors, including memory, personality, affect, and physiological conditions. Two of Flaherty's forms of temporal experience (i.e., "temporal compression" and "protracted duration") deal with retrospective time judgement. Temporal compression happens when the listening activity is not so engaging (e.g., in our sad and relaxing conditions); in these cases, the listener's brain works almost automatically, so that "time will be experienced and retroactively constructed as having flowed quickly" [90] (p. 256) [31]. Conversely, the protracted duration phenomenon arises in cases of intense, novel, or extraordinary experiences (e.g., the highly arousing soundtracks of our study, as opposed to the less arousing, might belong to this category to a greater extent), and, similarly to the memory-based models, it is mainly due to a more complex structure of information that needs to be processed. Such a difference is eminently important as it accounts for our results in a coherent fashion.

Limitations

Lastly, the four main limitations of this study must be highlighted. Firstly, all of the measures employed are self-reported. On the one hand, self-reported measures in psychological studies on music have been consistently applied to the study of musically elicited emotions over the last 35 years and presented good reliability as long as they stem from validated theories or models of emotion [91]. Moreover, there is some evidence suggesting a consistent overlap between self-reported and psychophysiological measures such as skin conductance levels [92–95], heart rate [94,96], finger temperature, and zygomatic facial muscle activity [95]. On the other hand, it must be acknowledged that these two sets of measures cannot always be considered as equally valid in all contexts, and other studies found more complex relationships between them [97], even in the audiovisual domain [98]. For these reasons, it would be good practice to replicate these findings in a laboratory setting by employing one or several psychophysiological measures [99].

Secondly, some studies insist on the role of music preference [25] and familiarity [100] in time perception. In our study, to construct a more condensed online task (and to avoid further losses in participation), we did not ask our participants to express their musical preferences, nor did we measure the extent to which they were previously exposed to the genre of the soundtrack they were listening to. Similarly, we did not ask for their movie preferences. All these personal characteristics could have slightly biased our findings.

Thirdly, as regards the stimulus complexity in audiovisuals, we need to mention that an assessment of the subjectively perceived musical fit [101], that is, the degree to which, according to a viewer, musical and visual information overlap each other with no semantical frictions, could have been profitable. Nevertheless, so far, such an assessment has been validated for audiovisual advertising only [101,102].

Lastly, the short movie we used as the stimulus was not completely neutral from an affective standpoint; indeed, we found that 64.42% of the viewers in the no-music condition reported a negative affect during their viewing. Although we are confident that such an “affect negativity” of the visual stimulus could not have jeopardized the validity of the results *per se*, we are less certain that it did not impact the perceived congruity of the audiovisual; namely, the fact that a negative visual stimulus was in some conditions paired with a positive soundtrack could have led to a decrease in the stimulus congruity, thus eliciting a slightly different (and perhaps more complex) processing. Indeed, in this design, we did not include a measure of the musical experience *per se*. In other words, we did not assess the self-reported valence and arousal of the musical pieces separately from the video. As a consequence, what we have referred to as a self-reported measure of the affective state of the participants is a measure of the overall audiovisual stimulus, subsequent to the aforementioned cognitive process of integration of the visual and auditory channels of information. On the one hand, the results of our model (Section 4.5.2) support the claim that the musical pieces were representative of the desired valence; on the other, there is also evidence that visual information influences the perception and memory of music [103]. In future studies, we plan to include the investigation of the bi-directional influence between music and video stimuli, especially with reference to musical fit [101,102] interactions with time perception.

6. Conclusions

To conclude, two main results have been found in this study. The first is that the mere presence of music, regardless of its valence and arousal, leads to time overestimation in an audiovisual context, possibly due to the cognitive process of integrating the visual and auditory information. Secondly, and most importantly, the primary result is that the subjectively perceived level of arousal, which is in turn increased by faster musical tempi and greater stimuli complexity (i.e., happy and scary soundtracks), positively predicts the time estimation of an audiovisual (i.e., arousal leads to time overestimation). In the light of the studies mentioned in Section 3, the supposedly causal role of the arousal in time overestimation appears to be solid. Further studies need to identify the cause by distinguishing between attention and memory-based models of time perception.

It is our intention to underline the potential that these findings and this research niche might present in the audiovisual domain. The notion that the interaction between the soundtrack and the moving image can affect the viewers’ time perception should receive further attention from media psychologists, video content creators, filmmakers, and, in general, any scholars or professionals interested in shaping and improving the interaction between viewers and an audiovisual. As the development of new technologies continues, their interactive uses become more and more explored and exploited. It is not negligible to claim that an ameliorated management of the background music within the audiovisuals could improve the interaction between the user and the audiovisual devices by shaping the recipients’ time perception.

Our results confirm previously collected evidence [16,59,64] revealing that the musically conveyed arousal, and not specific emotions, fosters time overestimation within a narrative audiovisual scene.

We are aware that, from a naïve point of view, the fact that arousing music steers the listeners towards time overestimations might appear paradoxical. Instead, this is far from being unknown among music composers. For instance, it is told that Maurice Ravel was very disappointed by Wilhelm Furtwängler's rendition of his Boléro, which was so fast that he thought it would have lasted forever [104].

Author Contributions: Conceptualization, A.A.; methodology, A.A. and M.M.; validation, L.M.; formal analysis, A.A. and M.M.; investigation, A.A.; data curation, A.A.; writing—original draft preparation, A.A.; writing—review and editing, A.A., M.M., and I.P.; supervision, L.M. and I.P. All authors have read and agreed to the published version of the manuscript.

Funding: No funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of Roma Tre University.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are openly available here: https://www.dropbox.com/s/pjz63dtc5bdyy1i/TimePerceptionDATA_1.sav?dl=0, accessed on 27 October 2021.

Acknowledgments: AA is deeply grateful to Alice Cappella for her assistance in creating the audiovisual stimuli.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lesiuk, T. The Effect of Music Listening on Work Performance. *Psychol. Music* **2005**, *33*, 173–191. [CrossRef]
2. Nguyen, T.; Grahn, J.A. Mind Your Music: The Effects of Music-Induced Mood and Arousal across Different Memory Tasks. *Psychomusicol. Music Mind Brain* **2017**, *27*, 81–94. [CrossRef]
3. Lehmann, J.A.; Seufert, T. The Influence of Background Music on Learning in the Light of Different Theoretical Perspectives and the Role of Working Memory Capacity. *Front. Psychol.* **2017**, *8*, 1902. [CrossRef] [PubMed]
4. Palazzi, A.; Wagner Fritzen, B.; Gauer, G. Music-Induced Emotion Effects on Decision-Making. *Psychol. Music* **2018**, *47*, 621–643. [CrossRef]
5. Israel, A.; Lahav, E.; Ziv, N. Stop the Music? The Effect of Music on Risky Financial Decisions: An Experimental Study. *J. Behav. Exp. Financ.* **2019**, *24*, 100231. [CrossRef]
6. Mentzoni, R.A.; Laberg, J.C.; Brunborg, G.S.; Molde, H.; Pallesen, S. Type of Musical Soundtrack Affects Behavior in Gambling. *J. Behav. Addict.* **2014**, *3*, 102–106. [CrossRef] [PubMed]
7. Cuesta, U.; Martínez-Martínez, L.; Niño, J.I. A Case Study in Neuromarketing: Analysis of the Influence of Music on Advertising Effectiveness through Eye-Tracking, Facial Emotion and GSR. *Eur. J. Soc. Sci. Educ. Res.* **2018**, *5*, 73–82. [CrossRef]
8. Ansani, A.; D'Errico, F.; Poggi, I. "It Sounds Wrong . . ." Does Music Affect Moral Judgement? In *Computational Science and Its Applications—ICCSA 2017*; Gervasi, O., Murgante, B., Misra, S., Borruso, G., Torre, C.M., Rocha, A.M.A.C., Taniar, D., Apduhan, B.O., Stankova, E., Cuzzocrea, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; Volume 10409, pp. 753–760. ISBN 978-3-319-62406-8.
9. Ansani, A.; D'Errico, F.; Poggi, I. 'You Will Be Judged by the Music I Hear': A Study on the Influence of Music on Moral Judgement. *Web Intell.* **2019**, *17*, 53–62. [CrossRef]
10. Ziv, N. Music and Compliance: Can Good Music Make Us Do Bad Things? *Psychol. Music* **2015**, *44*, 953–966. [CrossRef]
11. Ziv, N.; Hoftman, M.; Geyer, M. Music and Moral Judgment: The Effect of Background Music on the Evaluation of Ads Promoting Unethical Behavior. *Psychol. Music* **2012**, *40*, 738–760. [CrossRef]
12. Strick, M.; de Bruin, H.L.; de Ruiter, L.C.; Jonkers, W. Striking the Right Chord: Moving Music Increases Psychological Transportation and Behavioral Intentions. *J. Exp. Psychol. Appl.* **2015**, *21*, 57–72. [CrossRef] [PubMed]
13. Brown, S.; Theorell, T. The Social Uses of Background Music for Personal Enhancement. In *Music and Manipulation: On the Social Uses and Social Control of Music*; Berghahn Books: New York, NY, USA; Oxford, UK, 2006; pp. 126–162.
14. Rai, S. Comparison of Time-Estimation of Music, Noise, Light-Filled and Unfilled Intervals. *Indian J. Psychol.* **1973**, *48*, 37–43.
15. Wang, X.; Wöllner, C. Time as the Ink That Music Is Written with: A Review of Internal Clock Models and Their Explanatory Power in Audiovisual Perception. *Jahrb. Musik.* **2020**, *29*, e67. [CrossRef]

16. Droit-Volet, S.; Fayolle, S.L.; Gil, S. Emotion and Time Perception: Effects of Film-Induced Mood. *Front. Integr. Neurosci.* **2011**, *5*, 33. [[CrossRef](#)] [[PubMed](#)]
17. Chen, L.; Zhou, X.; Müller, H.J.; Shi, Z. What You See Depends on What You Hear: Temporal Averaging and Crossmodal Integration. *J. Exp. Psychol. Gen.* **2018**, *147*, 1851–1864. [[CrossRef](#)] [[PubMed](#)]
18. Wallmark, Z.; Nghiem, L.; Marks, L.E. Does Timbre Modulate Visual Perception? Exploring Crossmodal Interactions. *Music Percept.* **2021**, *39*, 1–20. [[CrossRef](#)]
19. Jolij, J.; Meurs, M. Music Alters Visual Perception. *PLoS ONE* **2011**, *6*, e18861. [[CrossRef](#)]
20. Ansani, A.; Marini, M.; D'Errico, F.; Poggi, I. How Soundtracks Shape What We See: Analyzing the Influence of Music on Visual Scenes Through Self-Assessment, Eye Tracking, and Pupillometry. *Front. Psychol.* **2020**, *11*, 2242. [[CrossRef](#)]
21. Boltz, M.G. Auditory Driving in Cinematic Art. *Music Percept.* **2017**, *35*, 77–93. [[CrossRef](#)]
22. Herget, A.-K. On Music's Potential to Convey Meaning in Film: A Systematic Review of Empirical Evidence. *Psychol. Music* **2019**, *49*, 21–49. [[CrossRef](#)]
23. Richards, D.; Fassbender, E.; Bilgin, A.; Thompson, W.F. An Investigation of the Role of Background Music in IVWs for Learning. *Res. Learn. Technol.* **2008**, *16*, 231–244. [[CrossRef](#)]
24. Berndt, A.; Hartmann, K. The Functions of Music in Interactive Media. In *Interactive Storytelling*; Spierling, U., Szilas, N., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5334, pp. 126–131. ISBN 978-3-540-89424-7.
25. Cassidy, G.; MacDonald, R.A. The Effects of Music on Time Perception and Performance of a Driving Game. *Scand. J. Psychol.* **2010**, *51*, 455–464. [[CrossRef](#)] [[PubMed](#)]
26. Sanders, T.A.; Cairns, P. Time Perception, Immersion and Music in Videogames; In Proceedings of the Human Computer Interaction (HCI) 2010, Dundee, UK, 6–10 September 2010.
27. Savan, A. The Effect of Background Music on Learning. *Psychol. Music* **1999**, *27*, 138–146. [[CrossRef](#)]
28. Furnham, A.; Strbac, L. Music Is as Distracting as Noise: The Differential Distraction of Background Music and Noise on the Cognitive Test Performance of Introverts and Extraverts. *Ergonomics* **2002**, *45*, 203–217. [[CrossRef](#)]
29. Ziv, N.; Omer, E. Music and Time: The Effect of Experimental Paradigm, Musical Structure and Subjective Evaluations on Time Estimation. *Psychol. Music* **2011**, *39*, 182–195. [[CrossRef](#)]
30. Block, R.A. Chapter 9 Experiencing and Remembering Time: Affordances, Context, and Cognition. In *Advances in Psychology*; Elsevier: Amsterdam, The Netherlands, 1989; Volume 59, pp. 333–363. ISBN 978-0-444-87379-8.
31. Schäfer, T.; Fachner, J.; Smukalla, M. Changes in the Representation of Space and Time While Listening to Music. *Front. Psychol.* **2013**, *4*, 508. [[CrossRef](#)]
32. Block, R.A.; Gruber, R.P. Time Perception, Attention, and Memory: A Selective Review. *Acta Psychol.* **2014**, *149*, 129–133. [[CrossRef](#)]
33. North, A.C.; Hargreaves, D.J. Can Music Move People?: The Effects of Musical Complexity and Silence on Waiting Time. *Environ. Behav.* **1999**, *31*, 136–149. [[CrossRef](#)]
34. Garlin, F.V.; Owen, K. Setting the Tone with the Tune: A Meta-Analytic Review of the Effects of Background Music in Retail Settings. *J. Bus. Res.* **2006**, *59*, 755–764. [[CrossRef](#)]
35. Fang, Z. The Study on the Effect of Background Music on Customer Waiting Time in Restaurant. *Open Cybern. Syst. J.* **2015**, *9*, 2163–2167. [[CrossRef](#)]
36. McDonnell, J. Music, Scent and Time Preferences for Waiting Lines. *Int. J. Bank Mark.* **2007**, *25*, 223–237. [[CrossRef](#)]
37. Oakes, S. Musical Tempo and Waiting Perceptions. *Psychol. Mark.* **2003**, *20*, 685–705. [[CrossRef](#)]
38. Guéguen, N.; Jacob, C. The Influence of Music on Temporal Perceptions in an On-Hold Waiting Situation. *Psychol. Music* **2002**, *30*, 210–214. [[CrossRef](#)]
39. Areni, C.; Grantham, N. (Waiting) Time Flies When the Tune Flows: Music Influences Affective Responses to Waiting by Changing the Subjective Experience of Passing Time. *ACR N. Am. Adv.* **2009**, *36*, 449–455.
40. Park, J.-S.; Stoel, L.D. How Background Music Affects Consumer Perception of Waiting Time?—A Mediating Role of Emotions. *J. Fash. Bus.* **2018**, *22*, 16–29. [[CrossRef](#)]
41. Cameron, M.A.; Baker, J.; Peterson, M.; Braunsberger, K. The Effects of Music, Wait-Length Evaluation, and Mood on a Low-Cost Wait Experience. *J. Bus. Res.* **2003**, *56*, 421–430. [[CrossRef](#)]
42. Vaitl, D.; Birbaumer, N.; Gruzelier, J.; Jamieson, G.A.; Kotchoubey, B.; Kübler, A.; Lehmann, D.; Miltner, W.H.R.; Ott, U.; Pütz, P.; et al. Psychobiology of Altered States of Consciousness. *Psychol. Bull.* **2005**, *131*, 98–127. [[CrossRef](#)] [[PubMed](#)]
43. Gabrielsson, A.; Wik, S.L. Strong Experiences Related to Music: A Descriptive System. *Musicae Sci.* **2003**, *7*, 157–217. [[CrossRef](#)]
44. North, A.C.; Hargreaves, D.J.; Heath, S.J. Musical Tempo and Time Perception in a Gymnasium. *Psychol. Music* **1998**, *26*, 78–88. [[CrossRef](#)]
45. Kellaris, J.J.; Altsech, M.B. The Experience of Time as a Function of Musical Loudness and Gender of Listener. *ACR N. Am. Adv.* **1992**, *19*, 725–729.
46. Droit-Volet, S.; Ramos, D.; Bueno, J.L.O.; Bigand, E. Music, Emotion, and Time Perception: The Influence of Subjective Emotional Valence and Arousal? *Front. Psychol.* **2013**, *4*, 417. [[CrossRef](#)] [[PubMed](#)]
47. Bueno, J.L.O.; Firmino, E.A.; Engelman, A. Influence of Generalized Complexity of a Musical Event on Subjective Time Estimation. *Percept. Mot. Ski.* **2002**, *94*, 541–547. [[CrossRef](#)] [[PubMed](#)]

48. Caldwell, C.; Hibbert, S.A. Play That One Again: The Effect of Music Tempo on Consumer Behaviour in a Restaurant. *ACR Eur. Adv.* **1999**, *4*, 58–62.
49. Kellaris, J.J.; Kent, R.J. Exploring Tempo and Modality Effects, on Consumer Responses to Music. *ACR N. Am. Adv.* **1991**, *18*, 243–248.
50. Bueno, J.L.O.; Ramos, D. Musical Mode and Estimation of Time. *Percept. Mot. Ski.* **2007**, *105*, 1087–1092. [[CrossRef](#)]
51. Lahdelma, I.; Eerola, T. Single Chords Convey Distinct Emotional Qualities to Both Naïve and Expert Listeners. *Psychol. Music* **2016**, *44*, 37–54. [[CrossRef](#)]
52. Kellaris, J.J.; Kent, R.J. The Influence of Music on Consumers' Temporal Perceptions: Does Time Fly When You're Having Fun? *J. Consum. Psychol.* **1992**, *1*, 365–376. [[CrossRef](#)]
53. Yalch, R.F.; Spangenberg, E.R. The Effects of Music in a Retail Setting on Real and Perceived Shopping Times. *J. Bus. Res.* **2000**, *49*, 139–147. [[CrossRef](#)]
54. Kellaris, J.J.; Mantel, S.P. Shaping Time Perceptions with Background Music: The Effect of Congruity and Arousal on Estimates of Ad Durations. *Psychol. Mark.* **1996**, *13*, 501–515. [[CrossRef](#)]
55. Zakay, D. Chapter 10 Subjective Time and Attentional Resource Allocation: An Integrated Model of Time Estimation. In *Advances in Psychology*; Elsevier: Amsterdam, The Netherlands, 1989; Volume 59, pp. 365–397. ISBN 978-0-444-87379-8.
56. Jones, M.R.; Boltz, M. Dynamic Attending and Responses to Time. *Psychol. Rev.* **1989**, *96*, 459–491. [[CrossRef](#)]
57. Gibbon, J. Scalar Expectancy Theory and Weber's Law in Animal Timing. *Psychol. Rev.* **1977**, *84*, 279. [[CrossRef](#)]
58. Hammerschmidt, D.; Wöllner, C. Sensorimotor Synchronization with Higher Metrical Levels in Music Shortens Perceived Time. *Music Percept.* **2020**, *37*, 263–277. [[CrossRef](#)]
59. Droit-Volet, S.; Fayolle, S.; Lamotte, M.; Gil, S. Time, Emotion and the Embodiment of Timing. *Timing Time Percept.* **2013**, *1*, 99–126. [[CrossRef](#)]
60. Wang, S.; Shi, Z. Temporal Entrainment Effect: Can Music Enhance Our Attention Resolution in Time? [poster presentation]. In Proceedings of the 12th International Conference of Students of Systematic Musicology, SysMus, Berlin, Germany, 10–12 September 2019.
61. Ortega, L.; López, F. Effects of Visual Flicker on Subjective Time in a Temporal Bisection Task. *Behav. Process.* **2008**, *78*, 380–386. [[CrossRef](#)] [[PubMed](#)]
62. Ornstein, R.E. *On the Experience of Time*; Penguin: Harmondsworth, UK, 1975.
63. Polti, I.; Martin, B.; van Wassenhove, V. The Effect of Attention and Working Memory on the Estimation of Elapsed Time. *Sci Rep.* **2018**, *8*, 6690. [[CrossRef](#)]
64. Droit-Volet, S.; Bigand, E.; Ramos, D.; Bueno, J.L.O. Time Flies with Music Whatever Its Emotional Valence. *Acta Psychol.* **2010**, *135*, 226–232. [[CrossRef](#)]
65. Herbert, R. *Everyday Music Listening: Absorption, Dissociation and Trancing*; Routledge: London, UK, 2016; ISBN 978-1-4724-8060-6.
66. Bisson, N.; Tobin, S.; Grondin, S. Remembering the Duration of Joyful and Sad Musical Excerpts: Assessment with Three Estimation Methods. *NeuroQuantology* **2009**, *7*, 46–57. [[CrossRef](#)]
67. Boltz, M.G. Tempo Discrimination of Musical Patterns: Effects Due to Pitch and Rhythmic Structure. *Percept. Psychophys.* **1998**, *60*, 1357–1373. [[CrossRef](#)]
68. Västfjäll, D. Emotion Induction through Music: A Review of the Musical Mood Induction Procedure. *Musicae Sci.* **2001**, *5*, 173–211. [[CrossRef](#)]
69. Macdiarmid, C. On Lockdown [Short Movie]. Available online: <https://vimeo.com/435128203> (accessed on 27 October 2021).
70. Elliott, D.; Polman, R.; McGregor, R. Relaxing Music for Anxiety Control. *J. Music Ther.* **2011**, *48*, 264–288. [[CrossRef](#)]
71. Juslin, P.N.; Laukka, P. Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. *J. New Music Res.* **2004**, *33*, 217–238. [[CrossRef](#)]
72. Cespedes-Guevara, J.; Eerola, T. Music Communicates Affects, Not Basic Emotions—A Constructionist Account of Attribution of Emotional Meanings to Music. *Front. Psychol.* **2018**, *9*, 215. [[CrossRef](#)] [[PubMed](#)]
73. Quinto, L.; Thompson, W.F.; Taylor, A. The Contributions of Compositional Structure and Performance Expression to the Communication of Emotion in Music. *Psychol. Music* **2014**, *42*, 503–524. [[CrossRef](#)]
74. Fabian, D.; Schubert, E. Expressive Devices and Perceived Musical Character in 34 Performances of Variation 7 from Bach's Goldbergvariations. *Musicae Sci.* **2003**, *7*, 49–71. [[CrossRef](#)]
75. Scherer, K.R.; Sundberg, J.; Tamarit, L.; Salomão, G.L. Comparing the Acoustic Expression of Emotion in the Speaking and the Singing Voice. *Comput. Speech Lang.* **2015**, *29*, 218–235. [[CrossRef](#)]
76. Eerola, T.; Friberg, A.; Bresin, R. Emotional Expression in Music: Contribution, Linearity, and Additivity of Primary Musical Cues. *Front. Psychol.* **2013**, *4*, 487. [[CrossRef](#)]
77. Grimm, E.M.; Van Everdingen, R.; Schöpping, M. Toward a Recommendation for a European Standard of Peak and LKFS Loudness Levels. *SMPTE Motion Imaging J.* **2010**, *119*, 28–34. [[CrossRef](#)]
78. Plutchik, R. A general psychoevolutionary theory of emotion. In *Theories of Emotion*; Elsevier: Amsterdam, The Netherlands, 1980; pp. 3–33. ISBN 978-0-12-558701-3.
79. Muthén, L.K.; Muthén, B. *Mplus User's Guide: Statistical Analysis with Latent Variables, User's Guide*; Muthén & Muthén: Los Angeles, CA, USA, 2017; ISBN 0-9829983-2-5.

80. Hu, L.; Bentler, P.M. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Struct. Equ. Modeling A Multidiscip. J.* **1999**, *6*, 1–55. [[CrossRef](#)]
81. Stratton, V.N. Influence of Music and Socializing on Perceived Stress While Waiting. *Percept. Mot. Ski.* **1992**, *75*, 334. [[CrossRef](#)]
82. Cohen, A.J. Congruence-Association Model of Music and Multimedia: Origin and Evolution. *Psychol. Music Multimed.* **2013**, 17–47.
83. Vitouch, O. When Your Ear Sets the Stage: Musical Context Effects in Film Perception. *Psychol. Music* **2001**, *29*, 70–83. [[CrossRef](#)]
84. Boltz, M.G. The Cognitive Processing of Film and Musical Soundtracks. *Mem. Cogn.* **2004**, *32*, 1194–1205. [[CrossRef](#)] [[PubMed](#)]
85. Boltz, M.G. Musical Soundtracks as a Schematic Influence on the Cognitive Processing of Filmed Events. *Music Percept.* **2001**, *18*, 427–454. [[CrossRef](#)]
86. Flaherty, M.G. *A Watched Pot: How We Experience Time*; NYU Press: New York, NY, USA, 1999; ISBN 0-8147-2687-9.
87. Husserl, E. *The Phenomenology of Internal Time-Consciousness*; Heidegger, M., Ed.; Indiana University Press: Bloomington, Indiana, 2019; ISBN 978-0-253-04199-9.
88. Ciavatta, D. 8. Merleau-Ponty and the Phenomenology of Natural Time. In *Perception and its Development in Merleau-Ponty's Phenomenology*; Jacobson, K., Russon, J., Eds.; University of Toronto Press: Toronto, ON, Canada, 2017; pp. 159–190. ISBN 978-1-4875-1285-9.
89. Shen, Y. The Trio of Time: On Merleau-Ponty's Phenomenology of Time. *Hum. Stud.* **2021**. [[CrossRef](#)]
90. Holmer Nadesan, M.; Flaherty, M.G. A Watched Pot: How We Experience Time [book review]. *Hum. Stud.* **2002**, *25*, 257–265. [[CrossRef](#)]
91. Zentner, M.; Eerola, T. Self-report measures and models. In *Handbook of Music and Emotion: Theory, Research, Applications*; Oxford University Press: Oxford, UK, 2010; pp. 187–221.
92. Ribeiro, F.S.; Santos, F.H.; Albuquerque, P.B.; Oliveira-Silva, P. Emotional Induction Through Music: Measuring Cardiac and Electrodermal Responses of Emotional States and Their Persistence. *Front. Psychol.* **2019**, *10*, 451. [[CrossRef](#)] [[PubMed](#)]
93. White, E.L.; Rickard, N.S. Emotion Response and Regulation to “Happy” and “Sad” Music Stimuli: Partial Synchronization of Subjective and Physiological Responses. *Musicae Sci.* **2016**, *20*, 11–25. [[CrossRef](#)]
94. Krumhansl, C.L. An Exploratory Study of Musical Emotions and Psychophysiology. *Can. J. Exp. Psychol./Rev. Can. Psychol. Expérimentale* **1997**, *51*, 336–353. [[CrossRef](#)]
95. Lundqvist, L.-O.; Carlsson, F.; Hilmersson, P.; Juslin, P.N. Emotional Responses to Music: Experience, Expression, and Physiology. *Psychol. Music* **2009**, *37*, 61–90. [[CrossRef](#)]
96. van der Zwaag, M.D.; Westerink, J.H.D.M.; van den Broek, E.L. Emotional and Psychophysiological Responses to Tempo, Mode, and Percussiveness. *Musicae Sci.* **2011**, *15*, 250–269. [[CrossRef](#)]
97. Lynar, E.; Cvejic, E.; Schubert, E.; Vollmer-Conna, U. The Joy of Heartfelt Music: An Examination of Emotional and Physiological Responses. *Int. J. Psychophysiol.* **2017**, *120*, 118–125. [[CrossRef](#)] [[PubMed](#)]
98. Ellis, R.J.; Simons, R.F. The Impact of Music on Subjective and Physiological Indices of Emotion While Viewing Films. *Psychomusicol. J. Res. Music Cogn.* **2005**, *19*, 15–40. [[CrossRef](#)]
99. Hodges, D.A. Psychophysiological measures. In *Handbook of Music and Emotion: Theory, Research, Applications*; Oxford University Press: Oxford, UK, 2010; pp. 279–311.
100. Bailey, N.; Areni, C.S. When a Few Minutes Sound like a Lifetime: Does Atmospheric Music Expand or Contract Perceived Time? *J. Retail.* **2006**, *82*, 189–202. [[CrossRef](#)]
101. Herget, A.-K.; Schramm, H.; Breves, P. Development and Testing of an Instrument to Determine Musical Fit in Audio-Visual Advertising. *Musicae Sci.* **2018**, *22*, 362–376. [[CrossRef](#)]
102. Herget, A.-K.; Breves, P.; Schramm, H. The Influence of Different Levels of Musical Fit on the Efficiency of Audio-Visual Advertising. *Musicae Sci.* **2020**, 1029864920904095. [[CrossRef](#)]
103. Boltz, M.G.; Eberndorf, B.; Field, B. Audiovisual Interactions: The Impact of Visual Information on Music Perception and Memory. *Music Percept.* **2009**, *27*, 43–59. [[CrossRef](#)]
104. Nichols, R. *Ravel*; Yale University Press: New Haven, CT, USA, 2011; ISBN 0-300-10882-6.