

When Music Drives Vision: Influences of Film Music on Viewers' Eye Movements

Karin Auer,^{*1} Oliver Vitouch,^{*2} Sabrina Koreimann,^{*3} Gerald Pesjak,^{#4} Gerhard Leitner,^{#5} Martin Hitz^{#6}

^{*}*Dept. of Psychology, University of Klagenfurt, Austria*

[#]*Interactive Systems Group, University of Klagenfurt, Austria*

¹karin.auer@gmx.at, ²oliver.vitouch@aau.at, ³sabrina.koreimann@aau.at,

⁴gerald.pesjak@gmail.com, ⁵gerhard.leitner@aau.at, ⁶martin.hitz@aau.at

ABSTRACT

Various studies have shown the co-determining strength that film music has on the viewers' perception. We here try to show that the cognitive processes of watching a film, observed through viewers' scanpaths and eye-movement parameters such as number and duration of fixations, are different when the accompanying film music is changed. If this holds, film music does not just add to a holistic impression, but the visual input itself is actually different depending on features of the soundtrack. Two film clips, 10 seconds each, were presented with three different musical conditions (horror music, documentary music, no music) in a between-subjects design. Clip 2 additionally contained a cue mark (red X in the bottom left corner, shown for 1 s). Participants' scanpaths were recorded using a *ASL H6* head-mounted eye-tracking system based on corneal reflection of infrared light. The resulting scanpaths of $N = 30$ participants showed distinct patterns dependent on the music condition. Specific trajectory categories were found for both film clips (five for clip 1, nine for clip 2). Systematic differences ($p < .05$) could be shown in most of these categories and variables. The additional cue mark was consciously perceived significantly more often in both music conditions than in the silent condition. Our results suggest that the slogan "What you see is what you hear" can be true on a very fundamental, first-layer level: Visual input varies with different scores, resulting in viewers not seeing the same film anymore in a straight sense.

I. INTRODUCTION

From a cognitive perspective, the use of music within film provides an interesting context in which to investigate effects of the auditory modality upon the cognitive processings of visual information.

Boltz (2001), p. 428

Eye fixations are intimately involved with our ability to visually encode spatially distributed information.

Just & Carpenter (1976), p. 444

...when looking at a visual display and completing a task, the location of one's gaze point corresponds to the thought that is "on top of the stack" of mental operations.

Goldberg & Wichansky (2003), p. 507

Film music is an important part of the total work of art of a movie. Even if film music may not be fully consciously perceived most of the time, it still affects the viewer. Research results in this field demonstrate the influence of film music regarding the meaning of a scene (Lipscomp & Kendall, 1994), the perceived activity of a scene (Bolivar, Cohen, & Fentress, 1994), the emotional meaning (Thayer & Levenson, 1983; Ellies & Simons, 2005), its communicational value (Infante &

Berg, 1976), the scene continuation (Vitouch, 2001), and the cross-modally perceived tempo (Vitouch, Sovdat, & Höller, 2006), but also regarding viewers' memory (Boltz, 2001) and the focusing of their attention (Marshall & Cohan, 1988). Several authors additionally included physiological parameters, such as skin conductance levels, into their measurements (Thayer & Levenson, 1983; Ellies and Simons, 2005). However, to the best of our knowledge, there is not a single study so far assessing the effects of film music on viewers' eye movements.

The design of our experiment relates to well-established psychological concepts. Theories on the effects of film music assume that listening to a particular music activates a specific scheme. This scheme influences the cognitive processing of visual input (Boltz, 2001). New information gets tied to the available material and the recipient's expectations. The old information gathered via the automatically triggered scheme influences the interpretation and judgment of new information.

Studies in music psychology found indicative characteristics of musical pieces, which serve as information carriers for listeners (Gabrielsson & Juslin, 2003; Koelsch, Kasper, Sammler, Schulze, Gunter, & Friederici, 2004). Film music operates on the same characteristics. Additionally, film music varies its attributes depending on the film genre (Schmidt, 1976; Libscomp & Kendall, 1994).

Representatives of the meanwhile classic *eye-mind hypothesis* assume that inner cognitive processes are reflected in the idiosyncratic scanpaths of eye movements. A given task is responsible for directing attention to the information that is required to solve this very task. Different tasks (Yarbus, 1976), different cognitive processes (Just & Carpenter, 1976; Loftus & Mackworth, 1978) and different daily acts (Pelz & Canosa, 2001) can be correctly identified based on the respective scanpaths' attributes. Also, eye movement onto a particular area is necessary for conscious attention in this sector (Deubel & Schneider, 1996).

Consequently, if film music activates a specific scheme, the cognitive processes that are typical for this scheme should be demonstrable in scanpaths. Different scanpaths, as a function of this particular scheme, should indicate that film music is not just responsible for influencing and "tuning" higher cognitive processes, and for adding to a holistic impression, but also for the basic process of gathering visual input. Since information with no representation in the scanpath cannot be processed, this information cannot be perceived by the viewer.

If film music activates a scheme which influences the eye movements of the viewer, then it co-determines which stimuli are perceived and therefore influence the perception of a scene. To investigate potential effects of film music along these lines, we developed the following experimental design.

II. METHOD

A. Design

The design of this study closely follows the basic design of Thayer and Levenson (1983) and of Ellies and Simons (2005). Two short film clips, each 10 seconds long, were coupled with three musical conditions – horror score, documentary score, and no music at all. Each participant was presented with both clips, and the musical condition rotated across subjects. During clip 2, for the duration of one second, a red X was visible in the bottom left corner. This hidden information was used to assess the location of participant's attention with the clip. Theoretically, the X should have a higher likelihood of conscious perception in the horror score condition – the scheme "horror" should direct participants' attention to the dark corners of a scene. We assume that media experience causes people to look there for visual affirmation.

A head-mounted eye-tracking device was used to record participants' scanpaths. After watching the film clips, a questionnaire with a three-point self measurement scale concerning the subjective activity of participants' gaze was given. Three questions concerning the hidden information (red X) in clip 2 were included as well. All participants were successively asked (a) if they had noticed anything special in the last scene, (b) if they had noticed an X, and, if applicable, (c) which color the X had.

B. Materials

Scenes from a documentary (*The Incredible Human Journey*, 2009, referred to as clip 1) and a movie (*The Painted Veil*, 2006, referred to as clip 2) were used as visual stimuli. Scenes were chosen because they permit equivocal interpretations. There was no sudden cut or change of scenery during these scenes. The clips differ from each other in two structural respects: (1) Clip 1 uses a pan shot whereas clip 2 shows a steady scene, and (2) the light in clip 1 is brighter. The visibility of the red X in clip 2 was tested in a pilot study ($N = 10$).

Two different musical pieces were used as film music. For the horror condition, *Big Fight* by Steve Jablonsky was used; for the documentary condition, *Amazonia I* from Alexandre Mabeix was chosen. The music was chosen as-it-is, without fine-tuned compositional coordination with the succession of visual patterns in the scenes.

An ASL H6 eye tracking device was used to record participants' gaze. This apparatus is based on pupil/corneal reflections of infrared light.

C. Sample

A total of 39 subjects participated in this study. Due to technical difficulties, the scanpaths of nine participants were excluded post-hoc from the study. Therefore, the scanpath data of $N = 30$ subjects (age: 14-51 years, 53% female) were analyzed. Since the subjects had no way of knowing if the eye tracking device malfunctioned or not, the questionnaires of all $N = 39$ subjects (age range same as above, 62% female) were analyzed.

III. RESULTS

A. Hypothesis 1: Film music influences participants' fixation rate

We assumed that the fixation rate, like other physiological measurements, is influenced by the musical condition. Similar to the physiological measures applied by Thayer and Levenson (1983), the horror condition should lead to a higher rate of fixations as compared to the no-music condition, while the documentary music condition should lead to the lowest rate of fixations. Due to technical difficulties, however, there was no objective way of counting participants' fixations. Therefore, just the three-point self-measurement scale was used to determine participants' gaze activity. In addition, a seven-point scale assessment was taken by the test supervisor.

1) *Self Measurement of Participants.* Each participant was asked to make an assumption concerning his or her gaze activity for each clip. Neither clip 1 nor clip 2 showed significant differences between the musical conditions in these subjective data (clip 1: $H = 0.94$, $p = .625$; clip 2: $H = 2.23$, $p = .328$).

2) *Measurements by the Test Supervisor.* The scanpaths of $N = 30$ subjects were analyzed separately by the test supervisor. Gaze activity was rated for each participant from one (minimal movement) to seven (extreme movement). Both clips were judged by the supervisor twice, without any essential differences between rating trials. The measurements of the test supervisor did not reveal significant between-groups differences (clip 1: $H = 1.6$, $p = .455$; clip 2: $H = 4.0$, $p = .134$; see Figure 1).

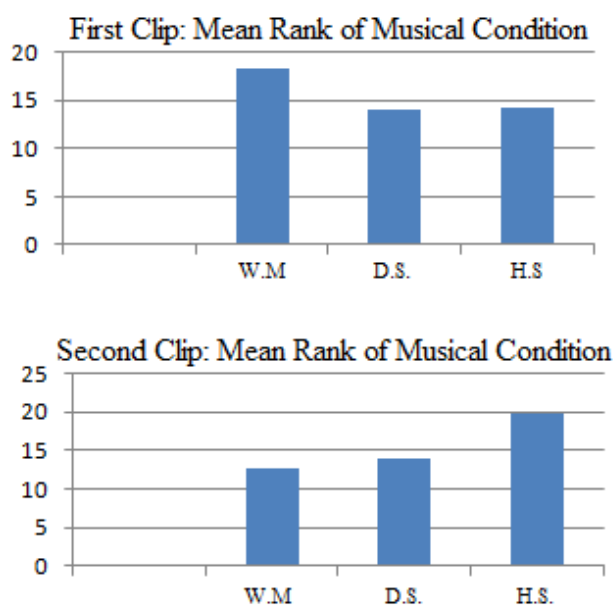


Figure 1. Results of the test supervisor seven-point measurement scale (mean ranks); W.M. = without music, D.S. = documentary score, H.S. = horror score

B. Hypothesis 2: Film music influences subjects' scanpaths

If film music activates a scheme, cognitive processes in accordance to this scheme should be visible in subjects' scanpaths. Data from $N = 30$ participants were analyzed here. Like before, clips were judged separately. Categories tailored to the film material and the gaze characteristics were derived inductively.

1) *Clip 1.* Analysis of clip 1 lead to five trajectory categories, four of them showing significant differences between musical conditions (see Table 1). The trajectory category "mountain" was found significantly more often in scanpaths from the documentary music and the no-music condition, whereas the category "gaze between persons" was not found in the documentary condition at all. The variable "calm" occurred significantly more often in the documentary and the no-music condition, the variable "hectic" in the horror condition.

Table 1. Trajectory categories for clip 1 and their statistical evaluations. (See text for details.)

categories for the first clip	statistical evaluation
pan shot orientation of gaze	χ^2 (df = 2) = 1.491, $p = .575$
mountain	χ^2 (df = 2) = 9.600, $p = .008$
gaze between persons	χ^2 (df = 2) = 6.477, $p = .039$
calm	χ^2 (df = 2) = 11.518, $p = .003$
hectic	χ^2 (df = 2) = 8.532, $p = .014$

2) *Clip 2.* This clip led to nine trajectory categories, seven of them showing significant between-groups differences (see Table 2). We assume that clip 2 caused a broader range of categories because there is no pan shot here. (In clip 2, subjects' gaze followed the pan shot most of the time – the so-called "pan shot orientation" of gaze.)

Table 2. Trajectory categories for clip 2 and their statistical evaluations. (See text for details.)

categories for the second clip	statistical evaluation
dark area	χ^2 (df = 2) = 18.373, $p < .000$
left hand region	χ^2 (df = 2) = .282, $p = .866$
center	χ^2 (df = 1) = 4.593, $p = .033$
boat	χ^2 (df = 2) = 1.491, $p = .475$
length of fixations	χ^2 (df = 1) = 4.593, $p = .032$
calm	χ^2 (df = 2) = 7.602, $p = .022$
hectic	χ^2 (df = 2) = 6.477, $p = .039$
slow alternations	χ^2 (df = 2) = 11.518, $p = .003$
fast alternations	χ^2 (df = 2) = 19.234, $p < .000$

Two out of nine categories (center and length of fixations) did not show significant differences between the three music conditions in the statistical standard evaluation. It was obvious from the data, however, that in these two categories, two of the three musical conditions (documentary music and no music) show parallel distributions of frequency. The distribution of those two musical conditions differed significantly from the third musical condition. The category "dark area" was significantly more often found in the scanpaths from the horror condition, whereas the category "center" did not feature in the

horror condition at all. The variable "length of fixation" was significantly larger in the documentary and the no-music condition, as were the variables "calm" and "slow alternations". Contrary to this, the categories "hectic" and "fast alternations" were significantly more frequent in the horror score condition.

C. Hypothesis 3: Film music increases and guides attention

To examine this hypothesis, additional information, namely a red X, was positioned in the bottom left corner of clip 2. Data from all $N = 39$ subjects were analyzed. Results show significant differences between the three music conditions (see Figure 2): χ^2 (df = 2) = 6.1, $p = .047$, $C = .37$.

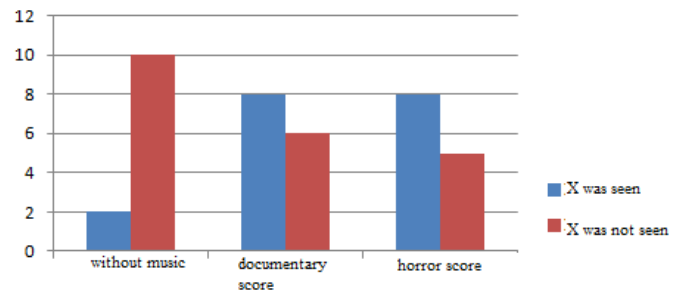


Figure 2. Frequency of consciously perceived hidden information between musical conditions

As shown in Figure 2, the frequency of consciously perceiving the red X is very similar in the documentary condition and the horror condition. These two musical conditions were therefore combined in a further analytical step. This difference between music and no-music conditions is significant with χ^2 (df = 1) = 6.1, $p = .014$, $C = .37$.

IV. CONCLUSION

The choice of film music did not influence the sheer number of fixations in this study. However, film music influences gaze patterns, and therefore, in a very direct and basic form, the perceived visual input. Music co-determines which visual facets of a scene will be processed. An unexpected event was perceived significantly more often in the music conditions.

Taken together, this study presents clear empirical evidence that film music systematically affects viewers' eye movements, and therefore the effectively perceived visual input. Once more, music makes the difference, even when it comes to the core visuals.

REFERENCES

- Bolivar, V. J., Cohen, A. J., & Fentress, J. C. (1994). Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action. *Psychomusicology*, 13, 28-59.
- Boltz, M. G. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception*, 18, 427-454.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, 36, 1827-1837.
- Ellis, R. J., & Simons, R. F. (2005). The impact of music on subjective and physiological indices of emotion while viewing films. *Psychomusicology*, 19, 15-40.
- Gabrielsson, A., & Juslin, P. N. (2003). Emotional expression in music. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 505-534). New York: Oxford University Press.

- Goldberg, J. H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 493-516). Amsterdam, The Netherlands: Elsevier.
- Infante, D. A., & Berg, C. B. (1979). The impact of music modality on the perception of communication situations in video sequences. *Communication Monographs*, 46, 135-141.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441-480.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, 7, 302-307.
- Lipscomb, S. C., & Kendall, R. A. (1994). Perceptual judgement of the relationship between musical and visual components in film. *Psychomusicology*, 13, 60-98.
- Loftus, R. G., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 565-572.
- Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, 41, 3587-3596.
- Schmidt, H. (1976). Musik als Einflussgröße bei der filmischen Wahrnehmung. In H. Schmidt (Ed.), *Musik in den Massenmedien Rundfunk und Fernsehen* (pp. 126-169). Mainz, Germany: Schott.
- Thayer, J. F., & Levenson, R. W. (1983). Effects of music on psychological responses to a stressful film. *Psychomusicology*, 3, 44-52.
- Vitouch, O. (2001). When your ear sets the stage: Musical context effects in film perception. *Psychology of Music*, 29, 70-83.
- Vitouch, O., Sovdat, S., & Höller, N. (2006). Audio-vision: Visual input drives perceived music tempo. In M. Baroni, A. R. Adessi, R. Caterina & M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception & Cognition* (pp. 86-89; CD-ROM). Bologna, Italy: Bononia University Press.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.