



Short Papers

Empfohlene Struktur-Gleichung? Einfluss von strukturell passender Filmmusik auf Emotionen, immersives Filmerleben und Filmbewertung

Influence of Structurally Fitting Film Music on Emotions, Immersive Film Experience, and Film Evaluation

Miguel Machulla¹ , Friederike Mertens¹, Alexander Frank², Ann-Kristin Herget^{*1} 

[1] Institut für Musik und Musikwissenschaft, Technische Universität Dortmund, Dortmund, Deutschland. [2] Medien- und Wirtschaftskommunikation, Institut Mensch-Computer-Medien, Julius-Maximilians-Universität Würzburg, Würzburg, Deutschland.

Zusammenfassung

Obwohl Expert*innen aus der Medienwelt immer wieder die Bedeutung struktureller Passung zwischen Musik- und Filmebene audiovisueller Medien betonen, gibt es bisher nur wenige Studien, die die möglichen Effekte struktureller Passung in den Fokus nehmen. Die vorliegende Studie untersucht daher die Auswirkungen struktureller Passung auf die Emotionalität, das räumliche Präsenzerleben und die allgemeine Filmbewertung. Basierend auf einem Arktisreisefilm (3 min) wurden drei Stimulusversionen mit verschiedenen Passungsgraden (konsonant, out-of-phase, dissonant) der Musik- und Filmebene für ein einfaktorielles between-subjects Online-Experiment erstellt. Musik, die strukturell besonders gut zum Film passt, kann die Gefühle der Rezipierenden beeinflussen—ANCOVAs (Kovariate: musikalische Vorbildung) zeigen signifikante Einflüsse für die Emotionen Transzendenz und Beruhigung. Strukturelle Passung kann ebenfalls bedingen, zu welchem Grad Rezipierende den Eindruck haben, körperlich in die Welt des Films einzutauchen. Dass eine strukturelle Passung von Musik und Film beeinflusst, wie Rezipierende den Film bewerten, konnte mit den vorliegenden Daten jedoch nicht gezeigt werden. Musikalische Vorbildung, als Kovariate, hatte keinen Einfluss. Die Ergebnisse des Experiments mit realistischer Rezeptionssituation lassen sich in den bisherigen Forschungsstand einordnen und deuten auf zumindest kleine Effekte der strukturellen Passung hin. Als Gründe für die zum Teil geringen gemessenen Unterschiede werden die Online-Durchführung und die komplexe Bearbeitung des Stimulusmaterials diskutiert. Für zukünftige Forschung in diesem bislang wenig beachteten Feld wird empfohlen, die Validität des Stimulusmaterials zu verbessern und weitere Kovariaten (wie Musikgeschmack, visuelle Expertise) zu berücksichtigen. Abschließend werden als mögliche Erklärungsansätze der Effekte struktureller Passung die Konzepte Fluency und Entrainment diskutiert.

Schlüsselwörter: Hintergrundmusik, audiovisuell, Musical Fit, Strukturelle Passung, Emotionalität, Filmbewertung, Räumliches Präsenzerleben

Abstract

Experts in the media industry frequently stress the significance of structural congruence between music and film in audiovisual media. Despite this emphasis, few studies have been conducted to investigate the effects of this type of musical fit. This study addresses this gap by examining how structural fit impacts emotionality, spatial presence, and overall film evaluation. Three different stimuli versions were created for a one-way between-subjects online experiment, each with varying degrees of fit (consonant, out-of-phase, and dissonant) between the music and film in an Arctic travel film (3 min). Music that fits the film's structural attributes can influence viewers' emotions. ANCOVAs (covariate: music expertise) reveal that music significantly affects feelings of transcendence and peacefulness. In addition, structural fit can determine how much viewers get immersed in the film's world. However, there is insufficient evidence to demonstrate that a structural fit between music and film affects viewers' evaluation of the movie. Music expertise, as a covariate, has no impact whatsoever. The results of the study can be classified according to the current state of research and indicate minor structural fit effects. The small effect sizes may be attributed to the implementation as online-experiment and the challenges of audio-visual stimulus material. For future studies, improving the validity of the stimulus material and considering additional covariates, such as music preferences and visual expertise, is recommended. The potential connections between structural fit, fluency, and entrainment are discussed as possible explanations.

Keywords: background music, audio-visual, musical fit, structural fit, emotionality, film evaluation, spatial presence

Jahrbuch Musikpsychologie, 2024, Vol. 32, Article e185, <https://doi.org/10.5964/jbdgm.185>

Received: 2023-11-16. Accepted: 2024-01-26. Published (VoR): 2024-05-15.

Reviewed by: Nicolas Ruth; Tim Loepphen.

*Corresponding author at: Institut für Musik und Musikwissenschaft, Technische Universität Dortmund, Emil-Figge-Str. 50, 44227 Dortmund, Deutschland. E-mail: Ann-Kristin.Herget@tu-dortmund.de



This is an open access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), CC BY 4.0, which permits unrestricted use, distribution, and reproduction, provided the original work is properly cited.

Was verbindet Tom, der Jerry jagt, einen mit einer Tarantel kämpfenden James Bond und Peter Parker, als er als Spiderman zum ersten Mal an einem Gebäude hochklettert? In den zugehörigen Filmszenen haben Rezipierende das Gefühl, dass Musik und Film in ihrer Struktur besonders gut aufeinander abgestimmt sind. Der Filmschnitt passt zu Tempo und Takt der Musik, Bewegungen in der Musik werden visuell eingefangen, Akzente in Bild und Ton erfolgen gleichzeitig—eine sogenannte “strukturelle Passung” (z. B. Cohen, 2005; Kloppenburg, 2017, S. 434) liegt vor. Dass eine Passung von Musik und Medienformat ganz allgemein eine Rolle dabei spielt, wie vorhersagbar und intensiv die Wirkung von Film- und Hintergrundmusik ausfällt, wird seit über 30 Jahren empirisch untersucht (z. B. Macinnis & Park, 1991; Marshall & Cohen, 1988). Bisherige Studien fokussieren die Übereinstimmung von Musik und medialem Kontext dabei vor allem auf emotionaler und auch assoziativer Ebene (Herget, 2021b): Greift die Musik die Emotionalität der Filmszene auf, ergänzt sie diese oder kontrastiert sie diese sogar gezielt? Werden durch bestimmte musikalische Stereotype, wie Instrumentalklischees oder spezifische Musikgenres, Assoziationen ausgelöst, die in ihrer Bedeutung zum Film passen? Vereinzelt empirische Studien (z. B. Costabile & Terman, 2013; Tan et al., 2007), vor allem jedoch Stimmen aus der Praxis betonen immer wieder auch die Relevanz einer Kongruenz auf struktureller Ebene, in die im finalen Filmschnitt oft besonders viel Zeit und Feingefühl investiert wird (Saltzman, 2022). Es stellt sich die Frage, welche Arten von struktureller Kongruenz im Filmkontext bedacht werden müssen und inwiefern diese die Wirkung eines Films beeinflussen können.

Musikalische Kongruenz und das Congruence Association Model (CAM)

Das *Congruence Association Model* (CAM) von Cohen erklärt, wie Musik die Wahrnehmung von Rezipierenden steuern und dabei Bedeutungen vermitteln kann (Cohen, 2010, 2015). Nicht nur im Kontext von Filmmusik, sondern in multimodalen Medien mit Musik jeglicher Art (Lipscomb, 2013; Tan, 2017), sind Rezipierende dazu in der Lage strukturell ähnliche Elemente verschiedener Ebenen des Films als zueinander passend wahrzunehmen. Wird beispielsweise Musik als strukturell passend zur Bildebene identifiziert, können durch die Musik hervorgerufene Emotionen und geweckte Assoziationen Elemente der Bildebene mit neuer Bedeutung aufladen. Zudem ist es gerade strukturelle Passung von Musik und Bild, die zu einer multisensorischen Integration der verschiedenen Ebenen des Films (Bild, Text, Sprache, Musik, Soundeffekte/Geräusche, kinästhetische Informationen) beitragen kann (Cohen, 2010, 2015; für die Anwendung des CAM auf ein konkretes Filmbeispiel, siehe Cohen, 2015).

Grundsätzlich können Musik und Bild nicht nur auf struktureller, sondern auch auf emotionaler und assoziativer Ebene zueinander passen. Am häufigsten erforscht worden ist die emotionale Kongruenz (Cohen, 2010; Unz et al., 2008). Dabei können affektiv-kongruente Musik- und Bildebenen sich hinsichtlich ihrer kommunizierten (Rosenfeld & Steffens, 2019) und induzierten Emotionen (Ellis & Simons, 2005) gegenseitig verstärken. Eine affektiv-inkongruente Kombination

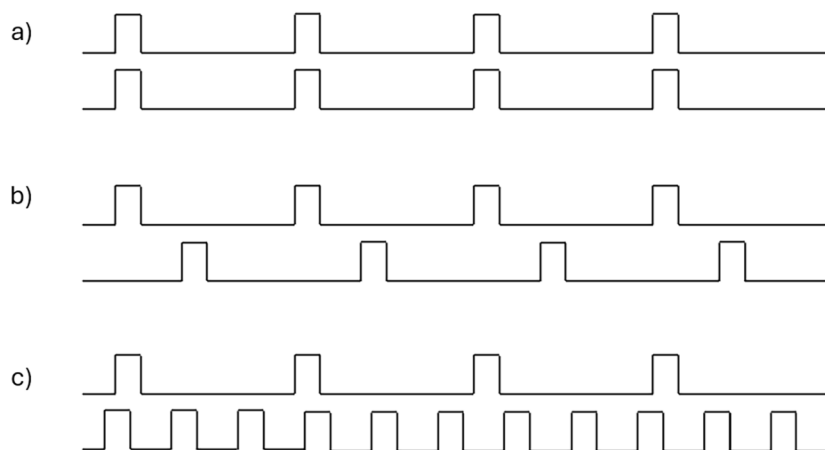
aus Bild- und Musikebene verringert hingegen die emotionalisierende Wirksamkeit der Hintergrundmusik (Bolivar et al., 1994; Rosenfeld & Steffens, 2019). Musik kann jedoch nicht nur emotionalisieren, sondern auch überzufällig Assoziationen wecken, welche durch aktivierte Schemata—organisierte Netzwerke von Erfahrungen, Erinnerungen und Wissen—hervorgerufen werden (Boltz, 2001; Brosius & Kepplinger, 1991; Shevy, 2008). Ähnlich wie bei der emotionalen Kongruenz können sich auch bei der assoziativen Kongruenz Bild- und Musikebenen eines Mediums in ihrer jeweiligen bedeutungsvermittelnden Wirkung verstärken (Lalwani et al., 2009; Hoekner et al., 2011; van den Stock et al., 2009). Lassen sich musikalische und visuelle Struktur als kongruent wahrnehmen, liegt eine strukturelle Passung von Bild- und Musikebene vor, wie sie auch im CAM beschrieben wird (Cohen, 2015). Von den drei Passungsarten wurde die strukturelle Passung bislang am seltensten empirisch untersucht (Herget, 2021a). Welche Wirkungen bisher angenommen werden können, wird im folgenden Kapitel erläutert.

Arten und mögliche Wirkungen struktureller Passung

Bullerjahn (2001) definiert strukturelle Passung als “Abstimmung von filmischen und musikalischen Parametern” durch eine “Bezugnahme von Rhythmus und Tempo der Filmmusik auf die Bildbewegungen und die Schnittgeschwindigkeit” (S. 140). Lipscomb konkretisiert in seinen Überlegungen zu “accent structure alignment” (2013, S. 193) dazu passend drei Grade von struktureller Passung. Bei der konsonanten Passung (a) erfolgen Akzente in Musik und im Film im selben zeitlichen Intervall und sind (nahezu) perfekt angeglichen. Ist die Passung hingegen out-of-phase (b), erfolgen die Akzente zwar im selben zeitlichen Abstand, sind aber verschoben. Von einer dissonanten Passung (c) wird gesprochen, wenn visuelle und musikalische Akzente in voneinander unabhängigen Intervallen auftauchen (siehe [Abbildung 1](#)).

Abbildung 1

Arten struktureller Passung nach Lipscomb



Anmerkung. Darstellung verschiedener „accent structure alignments“ nach Lipscomb (2005, S. 44).

Weitere musikalische Parameter können von Relevanz sein, wenn physische Veränderungen im Bild mit Veränderungen in der Musik korrespondieren sollen: Beispielsweise können die vertikale Position einer Person oder eines Gegenstandes im Bild mit Tonhöhe und Melodieverlauf und die Größe mit Lautstärke unterstrichen werden (für einen Überblick, siehe [Iwamiya, 2013](#)). Diese Form der plakativen strukturellen Passung findet sich etwa in der bekannten Filmmusik-Technik des “Mickey Mousing”, welches auch die Filmmusikbeispiele der Einleitung kennzeichnet.

Gezielte Bedeutungsvermittlung durch strukturelle Passung

Wie bereits beschrieben, kann dem CAM zufolge strukturelle Passung die Aufmerksamkeit Rezipierender auf bestimmte Elemente im Film lenken, die dann gegebenenfalls mit durch Musik transportierter Bedeutung (Emotionen und Assoziationen) aufgeladen werden. In einem animierten Kurzfilm mit geometrischen Figuren, der mit verschiedenen Musiken unterlegt wird, wird der Charakter eines bestimmten Dreiecks, das sich im Takt der Musik bewegt, je nach emotionalem Ausdruck der Musik unterschiedlich eingeschätzt. Die Wahrnehmung der anderen geometrischen Figuren des Films, die sich nicht im Takt bewegen, wird durch die Musiken weniger stark beeinflusst—ein klares Indiz für strukturelle Passung von Musik und Dreieck (Marshall & Cohen, 1988). Mera und Stumpf (2014) beschreiben für eine Eye-Tracking-Studie zudem, dass Rezipierende einen durch Musik gezielt strukturell hervorgehobenen Filmprotagonisten länger fokussieren als in der Bedingung mit strukturell weniger passender Musikuntermalung. Dies unterstützt die These, dass strukturelle Passung die Aufmerksamkeit Rezipierender auf bestimmte Filmelemente lenken kann, die dann durch die von Musik transportierten Emotionen und Assoziationen mit Bedeutung aufgeladen werden (siehe auch Millet et al., 2021).

Wahrgenommene Filmkohärenz und immersives Filmerleben durch strukturelle Passung

Das Vorliegen von struktureller Passung von Bild und Musik wird von Rezipierenden intuitiv wahrgenommen (Bolivar et al., 1994). Ob auch feinere Unterschiede identifiziert werden, testete Lipscomb (1995). Rezipierende schreiben seinem Stimulusmaterial mit konsonanter, out-of-phase und dissonanter Passung den Graden entsprechend mehr oder weniger Synchronität zu. Dass durch strukturelle Passung ein Film stärker als in sich geschlossene Einheit wahrgenommen wird und dadurch einzelne Wirkparameter verstärkt werden, deutet sich bei Costabile und Terman (2013) an: Rezipierende bewerten die Filmprotagonist*innen positiver und äußern anschließend häufiger filmkongruente Einstellungsänderungen bei einem Stimulus mit strukturell und assoziativ passender Musik (im Vergleich zum Stimulus ohne Musik—strukturelle Passung als Auslöser ist also nicht sichergestellt). Beispielsweise wird im Filmstimulus das Kümmern um ältere Angehörige dargestellt, das—je nach Musikkonnotation—von den Proband*innen anschließend eher als positive oder negative Verantwortlichkeit klassifiziert wird. Auch bei Lin et al. (2022) werden die Stimuli mit struktureller Passung insgesamt positiver bewertet.

Zum Medium im Allgemeinen passende Musik erleichtert grundsätzlich das Eintauchen in ein Filmgeschehen (Kock & Louven, 2018; siehe auch Cohen, 2014)—hierbei scheint strukturelle Passung jedoch von besonderer Bedeutung zu sein. Bei Costabile und Terman (2013) berichten Rezipierende bei strukturell passender Musik von einer höheren narrativen Transportation, für die auch Wallengren und Strukelj (2018) in einer Eyeblink-Studie Hinweise sehen. Ebenso zeigen sich in den offenen Angaben der Studie von Mera und Stumpf (2014) latente Hinweise auf ein stärkeres Eintauchen der Rezipierenden in die fiktive Erzählung bei strukturell passender Musik. Die Ergebnisse sind bislang aber noch lückenhaft. Voraussetzung für ein immersives Filmerlebnis ist zumeist eine Emotionalisierung der Proband*innen (Wirth et al., 2012), die zum Erleben von z. B. räumlicher Präsenz beitragen kann. Das Gefühl, körperlich in die Welt des Mediums einzutauchen und darin sogar aktiv handeln zu können, wird häufig im Kontext von Computerspielen untersucht (im Hinblick auf Musik im Computerspiel, siehe Klimmt et al., 2019), kann aber auch beim Ansehen von Filmen auftreten (z. B. Bracken, 2005). Inwiefern strukturelle Passung von Musik und Film zum Erleben räumlicher Präsenz beiträgt, wurde bislang noch nicht untersucht.

Hypothesen

Studien belegen, dass Filmmusik nicht nur eine bedeutende Rolle bei der Emotionalisierung spielt, sondern auch vorhersagbar bestimmte Assoziationen hervorrufen und die Aufmerksamkeit der Rezipierenden steuern kann. Dieser

Effekt ist am stärksten, wenn die Musik kongruent zum Bild ist. Die Kongruenz kann dabei auf verschiedene Arten vorhanden sein, wobei die strukturelle Passung in empirischer Forschung bisher am wenigsten berücksichtigt wurde (z. B. Cohen, 2010; Herget, 2021a).

Da eine intuitiv wahrgenommene allgemeine Passung von Musik und Film an sich induzierte Emotionen verstärkt und auch die Filmbewertung positiv beeinflusst (Ansani et al., 2020; Herget, 2021c, Herget & Albrecht, 2022; Steffens, 2020), ist anzunehmen, dass auch die strukturelle Passung, als ein Teil der allgemeinen Passung, Emotionalisierung und positive Filmbewertung begünstigt. Es gibt außerdem Hinweise darauf, dass gerade strukturelle Passung zu einer stärkeren Emotionalisierung der Rezipierenden und einer positiveren Filmbewertung beitragen kann (Costabile & Terman, 2013; Lin et al., 2022). Sie könnte zudem das Erleben räumlicher Präsenz vereinfachen (Costabile & Terman, 2013; Mera & Stumpf, 2014; Wallengren & Strukelj, 2018). Lipscomb (1995) zufolge kann dieser Effekt unterschiedlich stark ausfallen—je nach Grad der strukturellen Passung von Film und Musik. Auf Grundlage dieser Studienlage stellen wir folgende Hypothesen auf:

Eine perfekte, "konsonante" Passung von Musik und Film intensiviert bei Rezipierenden die Induzierung von Emotionen (**H1**) und das räumliche Präsenzerleben (**H2**). Auch die Filmbewertung (**H3**) wird von der konsonanten Bedingung am positivsten beeinflusst. Die Effekte (**H1–H3**) schwächen sich bei der "out-of-phase"-Passung gefolgt von der "dissonanten" Passung jeweils ab.

Methode

Methode, Design und Stichprobe

Eine a priori Stichprobenkalkulation mit G*Power (Faul et al., 2007) mit einer mittleren Effektgröße von $f = 0,295$ (Costabile & Terman, 2013, Transportation) ergab eine erforderliche Stichprobe von $N = 129$ (ANCOVA: fixed effects, main effects and interactions, $\alpha = ,05$, $1 - \beta = ,85$, Numerator $df = 2$, Anzahl der Gruppen = 3, Anzahl der Kovariaten = 1). Nach der Datenbereinigung (Ausschlusskriterien: ruckelnde Videowiedergabe, schlechte Bildqualität, kein Ton, Bekanntheit des Videos, zu kurze Fragebogen- und Stimulusbearbeitungszeit) flossen die Datensätze von 123 Proband*innen (61% weiblich, Alter: $M = 33,1$, $SD = 13,11$) in die Auswertung eines einfaktoriellen between-subject Online-Experiments (3×1 -Design; Befragungssoftware UNIPARK) ein. Nach der Abfrage üblicher Informationen zur Soziodemografie entsprechend der High Hurdle Technique von Reips (2002, S. 249) wurde den Proband*innen als Stimulus zu Beginn des Experiments zufällig eine Version eines Arktisreisefilms gezeigt, der mit Musik in drei strukturellen Passungsgraden vertont wurde. Nachfolgend wurden die vom Film induzierten Emotionen, sowie Technik- und Bekanntheitschecks erhoben. Um den Zweck der Studie so lange wie möglich verdeckt zu halten, folgte nach einer Frage zu den von Musik transportierten Emotionen die Messung des räumlichen Präsenzerlebens als weiteres Konstrukt ohne Musikbezug. Abschließend wurden die verschiedenen Arten der musikalischen Passung und die musikalische Vorbildung der Proband*innen erhoben.

Stimulusbeschreibung

Als Stimulus wurde ein im Original 7-minütiges YouTube-Video (<https://tinyurl.com/37pn6ez2>) verwendet, das für das Online-Experiment auf 3:14 min gekürzt wurde. Lipscomb (1995) folgend wurden Stimulus-Versionen mit unterschiedlichen Passungsgraden erstellt. In der ersten Version war das Verhältnis der Akzente von Musik und Film perfekt angeglichen ("konsonant", 4/4-Takt der Musik im Takt mit dem Videoschnitt, 90 BPM, wie im Originalvideo mit lediglich

geändertem Schluss). Für die zweite Version wurde die eigentlich perfekt passende Hintergrundmusik (Two Feet: Quick Musical Doodles and Sex) um 0,6 s nach vorne verschoben ("out-of-phase"), um die Synchronisation der audiovisuellen Akzente aufzubrechen. Um größtmögliche strukturelle Asynchronie herzustellen, wurde in der dritten Bedingung eine andere Musik (Jaki Rose: Watercolors, 3/4 Takt mit 120 BPM) gewählt, sodass die Akzente in der Musik nicht mehr mit denen des Films übereinstimmen ("dissonant"). Es wurde versucht, die emotionale und assoziative Passung von Film und Musik über die drei Versionen möglichst gleich zu halten, um eine Konfundierung zu vermeiden (Videobearbeitungssoftware: Adobe Premiere Pro). Um sicherzustellen, dass den Proband*innen die Manipulation nicht auffällt, wurde auf eine professionellen Produktionsstandards genügende Qualität des Stimulusmaterials geachtet (Herget, 2021a). Das Stimulusmaterial kann bei der korrespondierenden Autorin angefragt werden.

Messinstrumente

Die verschiedenen Arten von musikalischer Passung wurden nach Kellaris et al. (1993) erhoben (emotionale Passung, z. B. "Die Musik und das Video haben beide die gleiche generelle Stimmung erzeugt", $\alpha = ,88$; assoziative Passung, z. B. "Die Musik war in Bezug auf das Thema des Videos angemessen", $\alpha = ,81$). Sowohl die emotionale als auch assoziative Passung dienen als Konfundierungsschecks. Ein Item zur strukturellen Passung wurde durch zwei weitere Items ergänzt (angelehnt an Lipscomb, 1995, und Bolivar et al., 1994; z. B. "Musikschnitt und Videoschnitt haben perfekt zueinander gepasst", $\alpha = ,89$). Für jede der drei Passungen wurden drei Items verwendet.

Für die Erhebung der transportierten Emotionen, als weiteren Konfundierungsscheck, wurden die Dimensionen *Transzendenz* ($\alpha = ,84$) und *Beruhigung* ($\alpha = ,92$) der Geneva Emotional Music Scale (GEMS, Zentner et al., 2008; deutsche Version nach Lykartsis et al., 2013) mit je drei Items gemessen. Aus den neun Emotionsdimensionen, die die GEMS umfasst, wurden diejenigen ausgewählt, bei denen eine Beeinflussung durch das Stimulusmaterial am wahrscheinlichsten erschien. Die induzierten Emotionen wurden auf den gleichen Emotionsdimensionen ermittelt (Transzendenz: $\alpha = ,82$; Beruhigung: $\alpha = ,87$).

Wirth et al. (2012) führten Untersuchungen zur Rolle von unter anderem emotionalem Involvement bei der Bildung von Spatial Presence durch, weshalb ihre Arbeit als Grundlage für dieses Messinstrument diente. Es wurden vier Items in die Auswertung einbezogen, $\alpha = ,86$. Die verwendete Skala zur Bewertung des Films stützt sich auf Spears und Singh (2004). Die drei übernommenen Items ($\alpha = ,80$, z. B. "Das Video hat mir gefallen.") wurden inhaltlich für den Filmkontext leicht angepasst.

Alle bisherigen Konstrukte wurden zur einfacheren Bearbeitung im Online-Fragebogen einheitlich auf einer fünfstufigen Likert-Skala erhoben.

Die musikalische Vorbildung der Proband*innen wurde durch drei Items zur musikalischen Ausbildung des Gold-MSI (Müllensiefen et al., 2014; deutsche Version nach Schaal et al., 2014, $\alpha = ,86$, z. B. "Wie viele Jahre insgesamt haben Sie regelmäßig und täglich ein Instrument (einschließlich Gesang) geübt?") gemessen (für eine Übersicht der Messinstrumente siehe Machulla et al., 2023). In Studien im Kontext von Musik in audiovisuellen Medien hatte sie sich zuvor als relevante intervenierende Variable gezeigt (z. B. Bullerjahn et al., 1993; Rossmann & Rossmann, 2018).

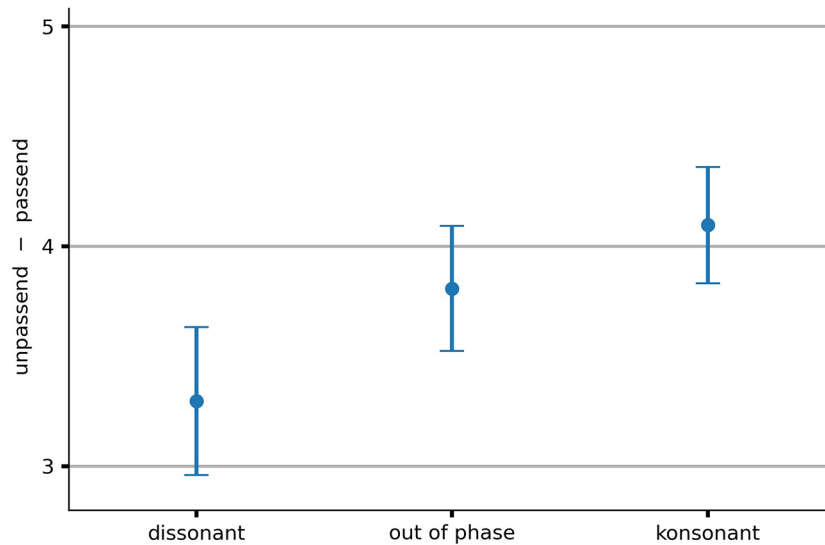
Ergebnisse

Manipulation- und Konfundierungcheck

In Abhängigkeit der strukturellen Passung der Hintergrundmusik zum Film haben die Proband*innen die konsonante ($M = 4,10$, $SD = 0,81$) Bedingung als passender wahrgenommen, als die out-of-phase ($M = 3,81$, $SD = 0,89$) oder dissonante ($M = 3,30$, $SD = 1,12$) Bedingung, $F(2, 120) = 7,49$, $p < ,001$, $\eta^2 = ,11$ (siehe *Abbildung 2*).

Abbildung 2

Wahrgenommene strukturelle Passung der drei Stimulusbedingungen



Anmerkung. Wahrgenommene strukturelle Passung (gemessen mittels 5-stufiger Likert-Skala, Abbildung zeigt Skalenausschnitt). Die Fehlerbalken bilden das 95%-Konfidenz-Intervall ab; $N = 123$.

In einer geplanten Kontrastanalyse wird der Mittelwertunterschied zwischen der konsonanten und out-of-phase Bedingung nicht signifikant, $t(120) = -1,33$, $p = ,19$, Cohen's $d = 0,30$, der zwischen der konsonanten und dissonanten Bedingung hingegen wird signifikant, $t(120) = -3,79$, $p < ,001$, Cohen's $d = 0,84$.

Um möglichst sicherzustellen, dass sich in der Hypothesenprüfung abzeichnende Unterschiede zwischen den experimentellen Bedingungen auf eine Manipulation der strukturellen Passung zurückführen lassen, wurden Konfundierungchecks durchgeführt. Das Stimulusmaterial sollte sich trotz der verschiedenen verwendeten Musiken bezüglich seiner emotionalen und assoziativen Passung sowie den von den Musiken transportierten Emotionen nicht signifikant unterscheiden. Es wurde für die Datenauswertung keine bedeutsame Konfundierung identifiziert (für Details zu diesen Auswertungen siehe Tabellen S1 und S2 in Machulla et al., 2023).

Hypothesenauswertung

Ersten Studien zur Wirkung von struktureller Passung von Musik und Film folgend sollte sich eine höhere strukturelle Passung von Musik und Film positiv auf die Emotionalisierung der Proband*innen, ihr räumliches Präsenzerleben und die Filmbewertung auswirken.

Hinsichtlich der durch die Musik in den Proband*innen ausgelöste Emotion Transzendenz unterscheiden sich die Mittelwerte der Stimulusbedingungen signifikant, $F(2, 119) = 3,72$, $p = ,03$, $\eta^2 = ,06$ (siehe Tabelle 1). Mit zunehmender struktureller Passung wird die Emotion wie erwartet stärker ausgelöst, jedoch zeigen sich in einer geplanten Kontrastanalyse (konsonant vs. out-of-phase, konsonant vs. dissonant) keine signifikanten Unterschiede (siehe Tabelle 5, erste Zeile). Zudem empfinden die Proband*innen in der out-of-phase- statt in der konsonanten Bedingung die stärkste Transzendenz (jedoch ohne signifikanten Unterschied zur konsonanten Bedingung).

Tabelle 1

Wirkung struktureller Passung auf die induzierte Emotion Transzendenz

Abhängige Variable	Exp. Bedingung	<i>n</i>	<i>EM</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Induzierte Emotion: Transzendenz	dissonant	45	3,17	3,16	0,83	3,72	2, 119	,03	,06
	out-of-phase	40	3,69	3,69	1,04				
	konsonant	38	3,56	3,57	0,89				
Musikalische Vorbildung						1,44	1, 119	,23	,01

Anmerkung. ANCOVA zur Wirkung unterschiedlicher Grade an struktureller Passung auf die induzierte Emotion Transzendenz, unter Berücksichtigung des Einflusses der musikalischen Vorbildung der Proband*innen als intervenierende Variable. $N = 123$, *EM* = geschätzte Randmittel, *M* = Mittelwerte, *SD* = Standardabweichung. Fett gedruckte Werte sind signifikant bei $p < ,05$.

Ein ähnliches Bild zeigt sich für die vom Arktisfilm ausgelöste Emotion der Beruhigung. Proband*innen in der Bedingung mit dissonanter Musik empfinden weniger Beruhigung als Proband*innen mit strukturell passenderer Musik, $F(2, 119) = 5,13$, $p = ,01$, $\eta^2 = ,08$ (siehe Tabelle 2). Hier unterscheidet sich jedoch wie erwartet in der geplanten Kontrastanalyse (konsonant vs. out-of-phase, konsonant vs. dissonant) die konsonante von der dissonanten Bedingung signifikant mit mittlerer Effektstärke. Wie bei der Transzendenz auch erleben Proband*innen in der out-of-phase-Bedingung die stärkste Emotionalisierung—die konsonante und die out-of-phase-Bedingung zeigen sich dabei jedoch erneut nicht signifikant unterschiedlich (siehe Tabelle 5, zweite Zeile). Die erste Hypothese muss verworfen werden—wobei sich eine Tendenz für eine höhere Emotionalisierung bei konsonanter als bei dissonanter Passung andeutet.

Tabelle 2

Wirkung struktureller Passung auf die induzierte Emotion Beruhigung

Abhängige Variable	Exp. Bedingung	<i>n</i>	<i>EM</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Induzierte Emotion: Beruhigung	dissonant	45	3,14	3,15	1,10	5,13	2, 119	,01	,08
	out-of-phase	40	3,73	3,73	0,74				
	konsonant	38	3,64	3,62	0,91				
Musikalische Vorbildung						3,19	1, 119	,08	,02

Anmerkung. ANCOVA zur Wirkung unterschiedlicher Grade an struktureller Passung auf die induzierte Emotion Beruhigung, unter Berücksichtigung des Einflusses der musikalischen Vorbildung der Proband*innen als intervenierende Variable. $N = 123$, *EM* = geschätzte Randmittel, *M* = Mittelwerte, *SD* = Standardabweichung. Fett gedruckte Werte sind signifikant bei $p < ,05$.

Werden die Ergebnisse der ANCOVA zum räumlichen Präsenzerleben der Proband*innen betrachtet, lassen sich über die Bedingungen hinweg signifikante Mittelwertunterschiede feststellen, $F(2, 119) = 4,60$, $p = ,01$, $\eta^2 = ,07$ (siehe Tabelle 3). Ein signifikanter Mittelwertunterschied zeigt sich zwar zwischen der konsonanten und dissonanten Bedingung, $t(120) = -2,56$, $p = ,01$, allerdings nicht zwischen der konsonanten und der out-of-phase-Bedingung, $t(120) = 0,02$, $p = ,99$ (siehe Tabelle 5, dritte Zeile). Somit muss die zweite Hypothese verworfen werden, auch wenn sich die Wahrnehmung

der Proband*innen bezüglich der räumlichen Präsenz zwischen der konsonanten und dissonanten Bedingung signifikant voneinander unterscheidet.

Tabelle 3

Wirkung struktureller Passung auf das räumliche Präsenzerleben

Abhängige Variable	Exp. Bedingung	<i>n</i>	<i>EM</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Räumliche Präsenz	dissonant	45	1,89	1,88	0,85	4,60	2, 119	,01	,07
	out-of-phase	40	2,40	2,40	0,93				
	konsonant	38	2,40	2,40	0,93				
Musikalische Vorbildung						0,19	1, 119	,66	< ,01

Anmerkung. ANCOVA zur Wirkung unterschiedlicher Grade an struktureller Passung auf das räumliche Präsenzerleben, unter Berücksichtigung des Einflusses der musikalischen Vorbildung der Proband*innen als intervenierende Variable. *N* = 123, *EM* = geschätzte Randmittel, *M* = Mittelwerte, *SD* = Standardabweichung. Fett gedruckte Werte sind signifikant bei $p < ,05$.

Die allgemeine Filmbewertung betreffend unterscheiden sich die Mittelwerte der drei Stimulusbedingungen nicht signifikant voneinander, $F(2, 119) = 1,00$, $p = ,37$, $\eta^2 = ,02$ (siehe Tabelle 4).

Tabelle 4

Wirkung struktureller Passung auf die Filmbewertung

Abhängige Variable	Exp. Bedingung	<i>n</i>	<i>EM</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Filmbewertung	dissonant	45	3,59	3,58	0,91	1,00	2, 119	,37	,02
	out-of-phase	40	3,79	3,78	0,81				
	konsonant	38	3,84	3,85	0,90				
Musikalische Vorbildung						1,47	1, 119	,23	,01

Anmerkung. ANCOVA zur Wirkung unterschiedlicher Grade an struktureller Passung auf die Filmbewertung, unter Berücksichtigung des Einflusses der musikalischen Vorbildung der Proband*innen als intervenierende Variable. *N* = 123, *EM* = geschätzte Randmittel, *M* = Mittelwerte, *SD* = Standardabweichung.

Die dritte Hypothese, nach welcher sich die Filmbewertung mit steigender struktureller Passung verbessert, muss verworfen werden, wobei sich in den Mittelwerten auch hier die vermutete Tendenz abzeichnet.

Tabelle 5

Einfache Kontrastanalyse zu induzierten Emotionen, Präsenzerleben und Filmbewertung

Abhängige Variable	Kontraste	<i>df</i>	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Induzierte Emotionen: Transzendenz	konsonant vs. out-of-phase	120	0,65	,52	0,12
	konsonant vs. dissonant	120	-1,90	,06	0,48
Induzierte Emotionen: Beruhigung	konsonant vs. out-of-phase	120	0,43	,67	0,13
	konsonant vs. dissonant	120	-2,47	,02	0,54
Räumliche Präsenz	konsonant vs. out-of-phase	120	0,02	,99	0,01
	konsonant vs. dissonant	120	-2,56	,01	0,58
Filmbewertung	konsonant vs. out-of-phase	120	-0,27	,79	0,08
	konsonant vs. dissonant	120	-1,32	,19	0,30

Anmerkung. Fett gedruckte Werte sind signifikant bei $p < ,05$.

Die in alle Analysen einbezogene Kovariate der musikalischen Vorbildung der Proband*innen hat bei keiner der abhängigen Variablen einen signifikanten Effekt (siehe Tabellen 1–4).

Diskussion

Positive Effekte struktureller Passung von Musik sind bereits jenseits des Films bekannt. So zeigen zum Beispiel Loo et al. (2019) im Kontext von Musik und Sport oder Meals et al. (2019) bezüglich der Dirigierqualität positive Effekte bei struktureller Passung. Hingegen ist Forschung zu struktureller Kongruenz im Filmkontext bislang rar (Lipscomb, 2013). Die Ergebnisse dieses Experiments legen nahe, dass zu einem Film möglichst strukturell passend eingesetzte Musik die von Rezipierenden empfundenen Emotionen (hier: Beruhigung) und den Grad der wahrgenommenen räumlichen Präsenz signifikant verbessert. Die gefundenen Effekte sind in ihrer Größe zudem vergleichbar mit Effektstärken ähnlicher Parameter vorheriger Studien. So deckt sich die gefundene mittlere Effektstärke der räumlichen Präsenz mit der gemessenen Transportation bei Costabile und Terman (2013, $\eta^2 = .07$). In der vorliegenden Studie kann bezüglich der Filmbewertung kein signifikanter Unterschied festgestellt werden. Die Mittelwerte stellen sich zwar wie erwartet dar, allerdings nur mit einer kleinen Effektstärke. Hinsichtlich der Filmbewertung konnten auch Lin et al. (2022) einen kleinen bis mittleren Effekt zu Gunsten konsonanter vs. dissonanter struktureller Passung bei audiovisuellen Medien feststellen (Cohen's $d_z = 0,40$). Mit einer größeren Stichprobe hätten sich womöglich auch in der vorliegenden Studie signifikante Effekte zur Filmbewertung darstellen lassen, was in zukünftiger Forschung noch zur Überprüfung aussteht.

In den geplanten Kontrastanalysen ließen sich zwischen der konsonanten (d. h. perfekt passenden) und der dissonanten (d. h. unpassenden) Bedingung signifikante Unterschiede finden. Obwohl bisherige Forschungsliteratur auch feinere Unterschiede der strukturellen Passung nahelegen (bspw. Lipscomb, 1995), konnten die vorliegenden Ergebnisse wider Erwarten auf keine signifikanten Unterschiede zwischen der konsonanten und der out-of-phase-Bedingung hinweisen. Könnte die meist implizite, d. h. unbewusste Verarbeitung von Hintergrundmusik (Strobin et al., 2015; Thompson et al., 1994) hier zu einem Problem geführt haben? Ein Blick in bisherige Forschungen (Bolivar et al., 1994; Lin et al., 2022; Lipscomb, 1995) und Cronbach's $\alpha = ,89$ für das vorliegende Instrument zur Erfassung der wahrgenommenen strukturellen Passung weisen darauf hin, dass Proband*innen durchaus dazu in der Lage sind, von dieser Form der Wahrnehmung explizit zu berichten.

Auswahl und Bearbeitung des Stimulusmaterials könnten die nur geringen Unterschiede zwischen den beiden passenden Stimulusbedingungen (konsonant und out-of-phase) erklären. Wie u. a. von Lipscomb (2005) gefordert, wird im Experiment mit komplexem, realistischem Stimulusmaterial gearbeitet. Das im Experiment als konsonante Bedingung verwendete Originalvideo weist aber eine Besonderheit auf: Offensichtlich wurde hier das Bild auf den 4/4-Takt der Musik geschnitten und nicht die Musik aufs Bild—ein für audiovisuelle Medienformate unübliches Vorgehen (z. B. Camp, 2013, zit. n. Strobin et al., 2015, S. 251). Die dadurch entstandene, tatsächlich nahezu perfekte strukturelle Passung des Taktes von Musik und Bild—ein idealer konsonanter Stimulus—stellte bei der Erstellung der out-of-phase-Bedingung eine Herausforderung dar. Eine Verschiebung der verwendeten Musik um nahezu jedes Zeitintervall ergab ungewollt wieder eine Passung zum Bild. Ohnehin neigen Rezipierende dazu, komplexe Stimuli bis zu einem gewissen Grad meist als strukturell passend wahrzunehmen, da das Gehirn bei der Rezeption bewusst nach Kongruenzen sucht und sie in einem gewissen Umfang auch dort herstellt, wo keine existieren (Lipscomb, 2013). Die hier gewählte Verschiebung der Musik um 0,6 s wirkte am unpassendsten—die Unterschiede in der Passung zwischen konsonant und out-of-phase sind dennoch fein. Trifft in der Realität Hintergrundmusik auf einen nicht in einem bestimmten Takt geschnittenen Filmausschnitt, ist das Potenzial einer feineren Abstufung von Nicht-Passung größer.

Da durch eine zeitliche Verschiebung der Musik, die in Takt und Tempo perfekt zum Video passt, bereits eine out-of-phase-Passung des Stimulusmaterials nur schwer herzustellen war, musste für die unpassende Bedingung eine andere Musik verwendet werden. Sind die im Experiment gefundenen Unterschiede zwischen den beiden passenden und der unpassenden Bedingung womöglich eher durch sich unterscheidende musikalische Parameter als durch Unterschiede in der strukturellen Passung zu erklären? Um eine ähnliche Wirkung der verwendeten Hintergrundmusiken möglichst sicherzustellen, wurden in umfassenden Konfundierungschecks die Passungsarten der Musiken (emotional und assoziativ) sowie die von ihnen transportierten Emotionen (Transzendenz und Beruhigung) gemessen. Mit Ausnahme der transportierten Transzendenz zwischen der konsonanten und dissonanten Bedingung transportieren die Stimulusvarianten die gleichen Emotionen und wurden in ihrer emotionalen und assoziativen Passung nicht signifikant unterschiedlich wahrgenommen. Es handelt sich zwar dementsprechend um verschiedene Musiken, die aber nach professionellen Qualitätsstandards hinsichtlich ihrer musikalischen Parameter so ausgewählt wurden, dass sie in ihrer Wirkung so ähnlich sind wie möglich. In künftiger Forschung könnte die interne Validität des Stimulusmaterials durch eigens für die Forschung komponierte Musiken, die sich tatsächlich möglichst nur im Parameter der strukturellen Passung unterscheiden, erhöht werden und die externe Validität durch einen Filmstimulus, der nicht auf die Musik geschnitten wurde.

Nicht zuletzt könnten die geringer als erwartet ausgefallenen Unterschiede zwischen den passenden Bedingungen durch die Durchführung als Online-Experiment zu erklären sein. Die nur eingeschränkte Kontrolle der Teilnahmebedingungen (z. B. unterschiedliche Sound-Wiedergabequellen) und dadurch eventuell auftretende Störfaktoren (Reips, 2002) könnten verhindert haben, dass die nur feinen Unterschiede der strukturellen Passung wahrgenommen werden. Um dieser Grundproblematik zu begegnen, werden Musikwirkungsexperimente oft im Labor durchgeführt (Eerola & Vuoskoski, 2013; Herget, 2021b). Eine höhere externe Validität der Forschung zu Musikwirkungen ist jedoch wünschenswert und kann nur durch die Untersuchung realistischerer Rezeptionssituationen erreicht werden. Dass in dieser Studie gezeigt wurde, dass Rezipierende bei der alltäglichen Mediennutzung grundsätzlich verschieden auf strukturell deutlich unterschiedlich passende Hintergrundmusiken reagieren, ist umso beachtlicher. In zukünftiger Forschung könnten jedoch etwa die Sound-Wiedergabequellen der Proband*innen erhoben werden, um einen möglichen Einfluss dieses Faktors zu kontrollieren (Wycisk et al., 2023, 2024). Ähnlich verhält es sich mit der Bildschirmgröße bzw. dem Medium, auf dem die Studie durchgeführt wird. Iwamiya merkte bereits 1994 (Iwamiya, 1994, S. 152) an, dass diese einen Einfluss auf die Wahrnehmung visueller Details haben, wodurch eine Einflussnahme auf die Wahrnehmung der strukturellen Passung zustande kommen kann.

Grundsätzlich ist die fehlende Repräsentativität der Stichprobe anzumerken. Das jüngere Durchschnittsalter der Proband*innen, die mit User Generated Content sozialisiert sind (Turner, 2015), könnte dazu führen, dass sie audiovisuelle Medien und damit auch strukturelle Passung von Musik und Medienformat grundsätzlich anders verarbeiten als ältere Generationen.

Weitere intervenierende Variablen, die die Wahrnehmung struktureller Passung beeinflussen könnten, in dieser Studie aber nicht beachtet wurden, sind das Musikgefallen (z. B. Herget et al., 2022) und die Kombination aus musikalischer Expertise und "visual literacy" (Lipscomb, 1995, S. 166), da Personen, die eine künstlerisch-visuelle und musikalische Ausbildung haben, dazu neigen, entsprechende Elemente in ihrer Umgebung verstärkt wahrzunehmen. Dadurch könnte eine Verzerrung der Ergebnisse zustande gekommen sein.

Während die vorliegende Studie erste Hinweise zu den Effekten struktureller Passung gibt, bleibt bislang offen, worauf diese Wirkungen basieren. Lin et al. (2022) vermuten, dass eine verbesserte cross-modale Integration, also die Interakti-

on zweier sensorischer Modalitäten (hier Hören und Sehen), das Verstehen von Sinneseindrücken vereinfacht. Diese „perceptual fluency“ führt zu einem besseren audiovisuellen Wahrnehmungsfluss, was wiederum eine Steigerung der Wahrnehmungsfreude nach sich ziehen könnte. Eine Synchronisation der Rhythmen von Musik und Bild könnte zudem zu Entrainment führen, was ebenfalls die Rezeption angenehmer machen würde. Eine zu perfekte Integration—der These der optimalen Aktivierung von Berlyne (1970) umgekehrt u-förmiger Kurve folgend—könnte bei Rezipierenden jedoch auch zu Langeweile führen. Inwiefern Fluency oder Entrainment als zugrundeliegende psychologische Konzepte tatsächlich eine Rolle bei der Erklärung der Effekte spielen, sollte in zukünftiger Forschung Beachtung finden.

Finanzierung

Die Autoren/die Autorinnen haben keine Finanzierung für das Forschungsprojekt erhalten.

Danksagung

Die Autoren/die Autorinnen haben keine weitere (d. h. nicht-finanzielle) Unterstützung erhalten.

Interessenkonflikte

Die Autoren/Autorinnen haben erklärt, dass keinerlei konkurrierende Interessen bestehen.

Ethikerklärung

Die vorliegende Arbeit wurde in Übereinstimmung mit ethischen Prinzipien und Standards gemäß den Richtlinien der Deutschen Gesellschaft für Psychologie (DGPs) durchgeführt.

Datenverfügbarkeit

Der SPSS-Datensatz zu diesem Beitrag wird im Open Science Framework zur Verfügung gestellt (siehe Machulla et al., 2023).

Ergänzende Materialien

Eine detaillierte Übersicht des Messinstrumentes, weiterführende Auswertungen (S1 und S2) sowie der SPSS-Datensatz zu diesem Beitrag werden im Open Science Framework zur Verfügung gestellt (siehe Machulla et al., 2023). Das Stimulusmaterial kann bei der korrespondierenden Autorin angefragt werden.

Quellenverzeichnis der ergänzenden Materialien

Machulla, M., Mertens, F., Frank, A. & Herget, A.-K. (2023). *Empfohlene Struktur-Gleichung? Einfluss von strukturell passender Filmmusik auf Emotionen, immersives Filmerleben und Filmbewertung* [Messinstrumente, Auswertungen, Datensatz]. OSF. <https://osf.io/zkhw7>

Literatur

Ansani, A., Marini, M., D’Errico, F., & Poggi, I. (2020). How soundtracks shape what we see: Analyzing the influence of music on visual scenes through self-assessment, eye tracking, and pupillometry. *Frontiers in Psychology*, 11, Article 2242. <https://doi.org/10.3389/fpsyg.2020.02242>

- Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, 8(5), 279–286. <https://doi.org/10.3758/BF03212593>
- 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: Music, Mind, and Brain*, 13(1–2), 28–59. <https://doi.org/10.1037/h0094102>
- Boltz, M. G. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception*, 18(4), 427–454. <https://doi.org/10.1525/mp.2001.18.4.427>
- Bracken, C. C. (2005). Presence and image quality: The case of high-definition television. *Media Psychology*, 7(2), 191–205. https://doi.org/10.1207/S1532785XMEP0702_4
- Brosius, H.-B., & Kepplinger, H. M. (1991). Der Einfluß von Musik auf die Wahrnehmung und Interpretation einer symbolisierten Filmhandlung. *Rundfunk und Fernsehen*, 39(4), 487–505.
- Bullerjahn, C. (2001). *Grundlagen der Wirkung von Filmmusik*. Wißner.
- Bullerjahn, C., Braun, U., & Güldenring, M. (1993). Wie haben Sie den Film gehört? *Jahrbuch der Deutschen Gesellschaft für Musikpsychologie*, 10, 140–158. <https://doi.org/10.23668/psycharchives.3293>
- Cohen, A. J. (2005). How music influences the interpretation of film and video: Approaches from experimental psychology. In R. A. Kendall & W. H. Savage (Hrsg.), *Perspectives in systematic musicology* (S. 15–36). Department of Ethnomusicology, UCLA.
- Cohen, A. J. (2010). Music as a source of emotion in film. In P. N. Juslin & J. A. Sloboda (Hrsg.), *Handbook of music and emotion: Theory, research, applications* (S. 879–908). Oxford University Press.
- Cohen, A. J. (2014). Resolving the paradox of film music through a cognitive narrative approach to film comprehension. In J. C. Kaufman & D. K. Simonton (Hrsg.), *The social science of cinema* (S. 57–83). Oxford University Press.
- Cohen, A. J. (2015). Congruence-association model and experiments in film music: Toward interdisciplinary collaboration. *Music and the Moving Image*, 8(2), 5–24. <https://doi.org/10.5406/musimoviimag.8.2.0005>
- Costabile, K. A., & Terman, A. W. (2013). Effects of film music on psychological transportation and narrative persuasion. *Basic and Applied Social Psychology*, 35(3), 316–324. <https://doi.org/10.1080/01973533.2013.785398>
- Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, stimuli. *Music Perception*, 30(3), 307–340. <https://doi.org/10.1525/mp.2012.30.3.307>
- Ellis, R. J., & Simons, R. F. (2005). The impact of music on the subjective and physiological indices of emotion while viewing films. *Psychomusicology: Music, Mind, and Brain*, 19(1), 15–40. <https://doi.org/10.1037/h0094042>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Herget, A.-K. (2021a). *Emotionsquelle, Bedeutungsträger, Taktgeber? Zur Wirksamkeit von Hintergrundmusik in audiovisuellen Medienformaten* [Doktorarbeit, Julius-Maximilians-Universität Würzburg]. OPUS Würzburg. <https://doi.org/10.25972/OPUS-23212>
- Herget, A.-K. (2021b). On music's potential to convey meaning in film: A systematic review of empirical evidence. *Psychology of Music*, 49(1), 21–49. <https://doi.org/10.1177/0305735619835019>

- Herget, A.-K. (2021c). Well-known and unknown music as an emotionalizing carrier of meaning in film. *Media Psychology*, 24(3), 385–412. <https://doi.org/10.1080/15213269.2020.1713164>
- Herget, A.-K., Breves, P., & Schramm, H. (2022). The influence of different levels of musical fit on the efficiency of audio-visual advertising. *Musicae Scientiae*, 26(1), 3–23. <https://doi.org/10.1177/1029864920904095>
- Herget, A.-K., & Albrecht, J. (2022). Soundtrack for reality? How to use music effectively in non-fictional media formats. *Psychology of Music*, 50(2), 508–529. <https://doi.org/10.1177/0305735621999091>
- Hoekner, B., Wyatt, E. W., Decety, J., & Nusbaum, H. (2011). Film music influences how viewers relate to movie characters. *Psychology of Aesthetics, Creativity, and the Arts*, 5(2), 146–153. <https://doi.org/10.1037/a0021544>
- Iwamiya, S.-I. (1994). Interactions between auditory and visual processing when listening to music in an audio visual context. 1. Matching 2. Audio quality. *Psychomusicology: Music, Mind, and Brain*, 13(1-2), 133–154. <https://doi.org/10.1037/h0094098>
- Iwamiya, S.-I. (2013). Perceived congruence between auditory and visual elements in multimedia. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall (Hrsg.), *The psychology of music in multimedia* (S. 141–164). Oxford University Press.
- Kellaris, J. J., Cox, A. D., & Cox, D. (1993). The effect of background music on ad processing: A contingency explanation. *Journal of Marketing*, 57(4), 114–125. <https://doi.org/10.1177/002224299305700409>
- Klimmt, C., Possler, D., May, N., Auge, H., Wanjek, L., & Wolf, A.-L. (2019). Effects of soundtrack music on the video game experience. *Media Psychology*, 22(5), 689–713. <https://doi.org/10.1080/15213269.2018.1507827>
- Kloppenburger, J. (2017). Musik im Film. In G. Rötter (Hrsg.), *Handbuch Funktionale Musik: Psychologie – Technik – Anwendungsgebiete* (S. 429–456). Springer Fachmedien. https://doi.org/10.1007/978-3-658-10219-7_18
- Kock, M., & Louven, C. (2018). The power of sound design in a moving picture: An empirical study with emoTouch for iPad. *Empirical Musicology Review*, 13(3–4), 132–148. <https://doi.org/10.18061/emr.v13i3-4.6572>
- Lalwani, A. K., Lwin, M. O., & Ling, P. B. (2009). Does audiovisual congruency in advertisements increase persuasion? The role of cultural music and products. *Journal of Global Marketing*, 22(2), 139–153. <https://doi.org/10.1080/08911760902765973>
- Lin, C., Yeh, M., & Shams, L. (2022). Subliminal audio-visual temporal congruency in music videos enhances perceptual pleasure. *Neuroscience Letters*, 779, Article 136623. <https://doi.org/10.1016/j.neulet.2022.136623>
- Lipscomb, S. D. (1995). *Cognition of musical and visual accent structure alignment in film and animation* [Unveröffentlichte Doktorarbeit]. University of California, Los Angeles.
- Lipscomb, S. D. (2005). The perception of audio-visual composites: Accent structure alignment of simple stimuli. *Selected Reports in Ethnomusicology*, 12, 37–62.
- Lipscomb, S. D. (2013). Cross-modal alignment of accent structures in multimedia. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall (Hrsg.), *The psychology of music in multimedia* (S. 192–213). Oxford University Press.
- Loo, F. C., Loo, F. Y., & Chua, Y. P. (2019). Congruence in music and movement enhances the perception of sports routine quality. *Revista Música Hodie*, 19, Article e58191. <https://doi.org/10.5216/mh.v19.58191>
- Lykartsis, A., Pysiewicz, A., von Coler, H., & Lepa, S. (2013). The emotionality of sonic events: Testing the Geneva Emotional Music Scale (GEMS) for popular and electroacoustic music. In G. Luck & O. Brabant (Hrsg.), *Proceedings of the 3rd International*

Conference on Music & Emotion (ICME3).

<https://jyx.jyu.fi/bitstream/handle/123456789/41586/1/Athanasios%20Lykartsis%20-%20The%20Emotionality%20of%20Sonic%20Events.pdf>

- Macinnis, D. J., & Park, C. W. (1991). The differential role of characteristics of music on high- and low-involvement consumers' processing of ads. *The Journal of Consumer Research*, *18*(2), 161–173. <https://doi.org/10.1086/209249>
- Marshall, S. K., & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes to geometric figures. *Music Perception*, *6*(1), 95–112. <https://doi.org/10.2307/40285417>
- Meals, C. D., Morrison, S. J., & Confredo, D. A. (2019). The effects of temporal action-sound congruence on evaluations of conductor quality. *Music & Science*, *2*, 1–10. <https://doi.org/10.1177/2059204319891968>
- Mera, M., & Stumpf, S. (2014). Eye-tracking film music. *Music and the Moving Image*, *7*(3), 3–23. <https://doi.org/10.5406/musimoviimag.7.3.0003>
- Millet, B., Chattah, J., & Ahn, S. (2021). Soundtrack design: The impact of music on visual attention and affective responses. *Applied Ergonomics*, *93*, Article 103301. <https://doi.org/10.1016/j.apergo.2020.103301>
- Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS One*, *9*(2), Article e89642. <https://doi.org/10.1371/journal.pone.0089642>
- Reips, U.-D. (2002). Standards for internet-based experimenting. *Experimental Psychology*, *49*(4), 243–256. <https://doi.org/10.1026//1618-3169.49.4.243>
- Rosenfeld, N., & Steffens, J. (2019). Effects of audiovisual congruency on perceived emotions in film. *Psychomusicology: Music, Mind, and Brain*, *29*(4), 200–208. <https://doi.org/10.1037/pmu0000242>
- Rossmann, C., & Rossmann, R. (2018). Die Wirkung von Musik in Informationsfilmen im Gesundheitsbereich. In P. Rössler & C. Rossmann (Hrsg.), *Kumulierte Evidenzen* (S. 165–187). Springer.
- Saltzman, S. A. (2022). *The music of film: Collaborations and conversations*. Routledge. <https://doi.org/10.4324/9781003147749>
- Schaal, N. K., Bauer, A.-K. R., & Müllensiefen, D. (2014). Der Gold-MSI: Replikation und Validierung eines Fragebogeninstrumentes zur Messung musikalischer Erfahrung anhand einer deutschen Stichprobe. *Musicae Scientiae*, *18*(4), 423–447. <https://doi.org/10.1177/1029864914541851>
- Shevy, M. (2008). Music genre as cognitive schema: Extramusical associations with country and hip-hop music. *Psychology of Music*, *36*(4), 477–498. <https://doi.org/10.1177/0305735608089384>
- Spears, N., & Singh, S. N. (2004). Measuring attitude toward the brand and purchase intentions. *Journal of Current Issues and Research in Advertising*, *26*(2), 53–66. <https://doi.org/10.1080/10641734.2004.10505164>
- Steffens, J. (2020). The influence of film music on the moral judgements of movie scenes and felt emotions. *Psychology of Music*, *48*(1), 3–17. <https://doi.org/10.1177/0305735618779443>
- Strobin, A. A., Hunt, J. B., Spencer, F. J., & Hunt, T. G. (2015). The role of music in motion picture advertising and theatrical trailers: Altering music to modify emotional response and genre expectations. *International Academy of Marketing Studies Journal*, *19*(3), 244–260.

- Tan, S.-L. (2017). Scene and heard: The role of music in shaping interpretations of film. In R. Ashley (Hrsg.), *The Routledge companion to music cognition* (S. 363–376). Routledge, Taylor & Francis. <https://doi.org/10.4324/9781315194738>
- Tan, S.-L., Spackman, M. P., & Bezdek, M. A. (2007). Viewers' interpretations of film characters' emotions: Effects of presenting film music before or after a character is shown. *Music Perception, 25*(2), 135–152. <https://doi.org/10.1525/mp.2007.25.2.135>
- Thompson, W. F., Russo, F. A., & Sinclair, D. (1994). Effects of underscoring on the perception of closure in filmed events. *Psychomusicology: Music, Mind, and Brain, 13*(1-2), 9–27. <https://doi.org/10.1037/h0094103>
- Turner, A. (2015). Generation Z: Technology and social interest. *Journal of Individual Psychology, 71*(2), 103–113. <https://doi.org/10.1353/jip.2015.0021>
- Unz, D., Schwab, F., & Mönch, J. (2008). Filmmusik und Emotionen. In S. Weinacht & H. Scherer (Hrsg.), *Wissenschaftliche Perspektiven auf Musik und Medien* (S. 177–191). VS Verlag für Sozialwissenschaften.
- van den Stock, J., Peretz, I., Grèzes, J., & de Gelder, B. (2009). Instrumental music influences recognition of emotional body language. *Brain Topography, 21*(3–4), 216–220. <https://doi.org/10.1007/s10548-009-0099-0>
- Wallengren, A.-K., & Strukelj, A. (2018). Into the film with music: Measuring eyeblinks to explore the role of film music in emotional arousal and narrative transportation. In T. Dwyer, C. Perkins, S. Redmond, & J. Sita (Hrsg.), *Seeing into screens: Eye tracking and the moving image* (S. 65–84). Bloomsbury Academic.
- Wirth, W., Hofer, M., & Schramm, H. (2012). The role of emotional involvement and trait absorption in the formation of spatial presence. *Media Psychology, 15*(1), 19–43. <https://doi.org/10.1080/15213269.2011.648536>
- Wycisk, Y., Kopiez, R., Bergner, J., Sander, K., Preihs, S., Peissig, J., & Platz, F. (2023). The Headphone and Loudspeaker Test – Part I: Suggestions for controlling characteristics of playback devices in internet experiments. *Behavior Research Methods, 55*, 1094–1107. <https://doi.org/10.3758/s13428-022-01859-8>
- Wycisk, Y., Sander, K., Manca, B., Kopiez, R., & Platz, F. (2024). The Headphone and Loudspeaker Test – Part II: A comprehensive playback device screening method in Internet experiments. *Behavior Research Methods, 56*, 362–378. <https://doi.org/10.3758/s13428-022-02048-3>
- Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion, 8*(4), 494–521. <https://doi.org/10.1037/1528-3542.8.4.494>