Johanna N. Wilson

From Trending to Mind-Bending

The Functions and Effects of Music Video Experiences

UNIVERSITY OF JYVÄSKYLÄ

FACULTY OF HUMANITIES AND
SOCIAL SCIENCES

Johanna N. Wilson

From Trending to Mind-Bending

The Functions and Effects
of Music Video Experiences

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

ABSTRACT

Music videos are a popular form of music listening which has received little empirical research from the field of music psychology. This dissertation addresses this gap in the literature using a multimethod approach to address the functions and effects of music video experiences. The first study used an abductive qualitative approach in order to explore adolescent and young adults' experiences with music videos. From this study, the Intention, Attention, Reaction and Retention (IARR) framework was established. This framework was used to inform the studies that followed. The second study extended the findings of Study I by exploring the relationship between IARR categories and the use of music for emotion regulation purposes measured with the Healthy Unhealthy Music Scale. Study III features an online study with items designed based on the main categorical themes within the IARR in order to explore the relationships between music video experience components and retention outcomes. Study IV uses the experience sampling method (ESM) to investigate how subsequent listening outcomes are affected by exposure to music video content. The studies presented in this thesis are the first attempt to understand music video experiences from a music psychology perspective. The results from this dissertation provide insight into how exposure to content in music videos affects music perception and under what circumstances these effects are most salient. These findings are especially important given the mass amount of musical multimedia easily accessible online, and provides a foundation for future research to explore music video phenomena.

Keywords: audio-visual, music perception, everyday listening, music video, emotion, cognition

TIIVISTELMÄ (ABSTRACT IN FINNISH)

Musiikkivideot ovat suosittu musiikin kuuntelun muoto, jota on tutkittu musiikkipsykologian alalla vain vähän. Tämä väitöskirja vastaa tähän tutkimusvajeeseen hyödyntäen monimenetelmäistä lähestymistapaa musiikkivideoiden funktioiden ja vaikutusten tarkastelussa. Ensimmäinen tutkimus käytti abduktiivista laadullista lähestymistapaa tutkiakseen nuorten ja nuorten aikuisten musiikkivideokokemuksia. Tuloksena syntyi Tavoite, Huomio, Kokemus, Seuraus – viitekehys (Intention, Attention, Reaction and Retention framework, IARR). Tätä viitekehystä käytettiin seuraavissa osatutkimuksissa. Toinen tutkimus syvensi ensimmäisen tutkimuksen tuloksia tarkastellen IARR kategorioiden yhteyksiä musiikilliseen tunnesäätelyyn, jota mitattiin HUMS (Healthy Unhealthy Music Scale) mittarilla. Tutkimus III oli verkkokysely, jonka väittämät muotoiltiin IARR:n teemakategorioiden pohjalta, pyrkimyksenä selvittää musiikkivideokokemuksen osa-alueiden yhteyttä koettuihin seurauksiin. Tutkimus IV käytti kokemustenkeruumenetelmää keinona tutkia sitä, miten musiikkivideoiden sisältö vaikuttaa myöhempiin kuuntelukertoihin. Tämän väitöskirjan tulokset tarjoavat tietoa siitä, miten musiikkivideoiden sisällölle altistuminen vaikuttaa musiikin havaitsemiseen, ja millaisten tekijöiden seurauksena nämä vaikutukset ovat voimakkaimpia. Tulokset ovat erityisen tärkeitä ajassa, jossa musiikillinen multimedia on verkossa helposti saatavilla, ja työ tarjoaa lähtökohdan myöhemmälle tutkimukselle musiikkivideoista ilmiönä.

Avainsanat: audiovisuaalinen, musiikin havaitseminen, arjen musiikinkuuntelu, musiikkivideo, emootio, kognitio

Author
Johanna Nancy Wilson
Department of Music, Art and Culture Studies
University of Jyväskylä
jonawils@jyu.fi
ORCID ID: 0000-0003-1130-5507

Supervisors
Suvi Saarikallio
Department of Music, Art and Culture Studies
University of Jyväskylä

Marc Thompson
Department of Music, Art and Culture Studies
University of Jyväskylä

Reviewers
Tuomas Eerola
Department of Music
Durham University

Annabel Cohen
Department of Psychology
University of Prince Edward Island

Opponent
Tuomas Eerola
Department of Music
Durham University

ACKNOWLEDGEMENTS

This thesis is dedicated in loving memory of my father, David Mathew Wilson (1943-2020). My father was a huge influence on my love of music. He was the type of person who would do anything to give his daughters the opportunities to learn and grow as people. He supported every hobby and interest I've had over the years, investing all his spare time and resources into his family so we could live happy, fulfilling lives. There are so many songs that make me instantly think of him because of all the memories I have of us listening, dancing, and goofing off together. He is my strongest, fondest, extramusical association. While he would not be able to see the final product, he never had any doubt I would succeed in completing this dissertation.

Words cannot adequately describe the magnitude my father's love, or the extent of his support for my passions and interests throughout my life. Thank you, Tog, for being the father you were. I could never have done it without you. You'll live on in my heart and in my memory as my adoring father and biggest supporter forever. I know you would have been proud.

Jyväskylä, May 2024
Johanna N. Wilson

FIGURE

CONTENTS

INTRODUCTION

Music media has a valuable role in our lives which has increased as it becomes more accessible than ever before. Technological developments over the course of the 20th and 21st century have provided more access to music, giving individuals the control to choose what they want to listen to, whenever and wherever they want (Krause et al., 2015; Nowak, 2016). This has been made possible thanks to the internet and the variety of music streaming platforms available. One of the most popular platforms hosting music multimedia online is YouTube. Originally a video streaming platform, it has since become one of the most popular platforms for online music streaming. YouTube is especially appealing for music listeners because it allows them to engage with their favourite songs in audio-visual formats, particularly music videos.

Audio-visual methods of consuming music are nothing new. For thousands of years, people have been watching while listening, whether it be the music that accompanied ancient Greek theatre, to classical opera and modern musicals, and contemporary films. In these contexts, we see music accompanying a story, woven into the narrative as a complementary source of information. For example, music in film usually exists as non-diegetic sound providing the viewer with more information about the characters' thoughts and feelings (Cohen, 2001; Cook, 1998). In everyday music listening contexts on the other hand, music is often used as a resource to help us understand our own thoughts and feelings. For this reason, it is widely acknowledged as a powerful tool for self-reflection, emotion regulation, and social bonding (Hargreaves & North, 1999; Schäfer et al., 2013). Music is very much part of the diegesis of our life, and we attribute personal meanings and memories to our favourite songs *because* music has these functions. Therefore, it is never just the music which is affecting us: extramusical information (information which is not physically present in the music itself) influences our experience of the music, even if we are not consciously aware of it. The memories we associate with our favourite songs are a form of extramusical information capable of influencing our emotional responses to music; a mechanism referred to as episodic memory (Juslin, 2013; Juslin & Västfjäll, 2008). Music videos contain a wealth of extramusical information,

whether it be in the form of narratives, interesting visual effects, dance choreography and other performance gestures. How this information influences our perception and affective responding to the music has yet to be explored.

By 2019, YouTube was the most popular online streaming platform, with 47% of all online music streaming occurring on the platform that year (IFPI, 2019; IFPI, 2018). It is no surprise that, of the top 30 most viewed videos on YouTube of all time, 21 are music videos. However, the other nine most viewed videos, including Baby Shark, a children's song released by South Korean education group Pinkfong, are still a form of music video but with a different target audience: young children. While music videos themselves are not a new phenomenon, their popularity and newfound accessibility raises questions about their role and function as a method of engaging with music. Music videos' visual component provides listeners with extramusical information which can potentially alter their perception of the music. If the listener's perception of the music is changed, the way they use a song to fulfil certain psychological functions may be affected as well; music psychology research has yet to account for this possibility.

This thesis has two primary objectives: the first is to gain more insight into the role and function of contemporary music video experiences as a method of musical engagement; the second investigates their effects on music perception, not only during the experience itself but in subsequent everyday listening experiences. Despite having been part of mainstream music listening culture for the last four decades, music videos have been largely overlooked by music psychology research. The thesis addresses this gap in the literature and provides a framework for future research exploring music video phenomena. Music psychology research in general would benefit from a better understanding of how extramusical information from music video experiences affects music listening outcomes. Such a framework is necessary, since music videos and other forms of musical multimedia (for example, livestream concerts, TikToks and short reel formats, etc) are continuing to grow in popularity (IFPI, 2022). A large body of research already exists which investigates the relationship between music video content and anti-social behaviour in young people, such as violence and aggressive behaviour (Collinson et al., 2014; Rich et al., 1998; Smith & Boyson, 2002). This thesis does not expand on or refute these claims. Instead, it explores how visual information from these experiences influences personal music listening outcomes and whether this influence extends to subsequent personal music listening outcomes (that is, audio-only listening experiences). In order to achieve these aims, the thesis explores music video phenomena using a multidisciplinary approach; it draws on existing theoretical models from topical areas of both music psychology and audio-visual perception research and conceptualises them in the context of music video experiences.

The thesis consists of four studies, each of which uses a distinct methodological approach to gain insight into music video phenomena. The thesis was guided by the following research questions:

- What are the functions of music video experiences? Why do individuals choose to watch them?
- How does extramusical content (i.e., visual information) influence music perception in music video contexts? How does this content influence music evoked affective reactions (eg., emotional outcomes)?
- How does exposure to music video content influence music perception and music evoked affect in subsequent, audio-only personal listening experiences?

The summary of this thesis is organised into five chapters. Chapter one provides a background on existing music psychology and audio-visual perception research, including the theories which informed the development of the framework presented in this thesis. Chapter two describes the aims and research questions of each study and the methods implemented to address these questions, including the procedures and materials used. Chapter three gives details of the main findings of each study. Chapter four discusses the main findings of this thesis, its limitations, and suggestions for future research. General conclusions are provided in chapter five.

1 THEORETICAL BACKGROUND

In order to address the questions presented in this thesis we must first explore the existing theories which can contribute to our understanding of music video experiences. Although there has been little research on music videos specifically in music psychology, insight can be drawn from many existing theories and frameworks used to research music in everyday life (especially personal music listening), the influence of context and individual differences on music listening, the mechanisms underlying music evoked affect, music and movement, embodied cognition, audio-visual perception, and the functions of music in film. The aim of the first chapter is to synthesise these various topics as they apply to music videos. In order to accomplish this, the topics in this chapter are based on the two primary objectives of the thesis: the first half of Chapter 1 explores research on the use of music in everyday life, including the functions it fulfils and the role of contextual factors and individual differences, while the second half outlines theoretical background on the cognitive and affective processes involved during music listening and audio-visual perception. The theories outlined in the second section provide a foundation for understanding how extramusical content from music video experiences influence music perception, and by extension, how the music is perceived in subsequent listening experiences. This layout is visualised in Figure 1, which highlights the relevant theoretical background which this thesis is based on. Before diving into the theoretical background, it is first important to provide a very brief history of music videos, the purpose they were meant to serve within the industry, and what they have evolved to become today.
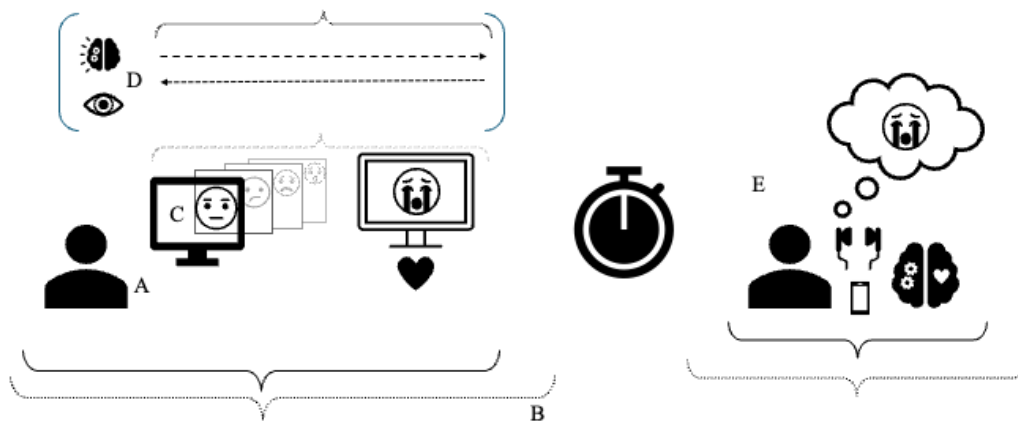
FIGURE 1   Visualisation of music video experience phenomena A: The listener level. B: the context (including the greater macro-context beyond just the situation where the viewer-listener is. C: Content of the music video where attention is directed. D: Cognitive appraisal throughout the experience, directed by emotional content in the music video, influences core affect and culminates to an emotional outcome. This facilitates the joint encoding of music with visual content. E: information encoded from the initial experience manifests in subsequent listens as a result of the processes outlined in level D, and are also affected by variables from level A and B.

Before diving into the theoretical framework behind this thesis, it is first important to provide a very brief history of music videos, the purpose they were meant to serve within the industry, and what they have evolved to become today.

1.1   YouTube killed the MTV star: a very brief history of music videos.

On August 1st, 1981, a new cultural phenomenon found its way to television. The music video for the song "Video Killed the Radio Star '' was debuted on MTV: a television channel dedicated entirely to music, specifically music videos, which would play all day, every day. The channel would bring about some of the most iconic videos in pop music history, including Michael Jackson's "Thriller" (1983), Van Halen's "Jump" (1983), Madonna's "Like a Virgin"(1984), and Peter Gabriel's "Sledgehammer" (1986) , among others. The purpose of music videos at this point was to promote singles off of albums in order to market them to the public (Edmond, 2014). The channel was eventually bought by the entertainment

15

conglomerate Viacom Inc. in 1985, which expanded its programming. However, music videos were still at the core of its production.

The concept of a music television channel continued to spread across the globe, with similar channels sprouting across the Americas, Asia, Australasia, and Europe. By the 1990s, the demand for non-stop music video programming had declined and the channel began to broadcast more content, both musical and non-musical (i.e., reality television) in order to keep their youth audience engaged. By the 2000s, consumers could buy music video packages on DVDs (Straw, 2008). While it was possible to watch music videos on the Internet, selection and accessibility were limited. The music video website Vevo, founded in 2006, became the primary online platform for music videos. Vevo began distributing its content on YouTube in 2013, effectively making YouTube the new music video streaming platform. YouTube has become one of the most popular platforms for watching music videos – so popular the platform developed its own streaming service to accommodate its users' needs.

The ways in which people engage with music content, whether it be music videos or audio formats, continues to evolve at a rapid pace. In the 2018 report released by the International Federation of the Phonographic Industry (IFPI), individuals reported listening to around 17.8 hours of music per week. The 2018 report also states that 52% of all music streaming was done on a video streaming platform, of which YouTube accounted for 47%. By comparison, paid and free audio streaming platforms together accounted for the remaining 48% (28% and 20% respectively). Interestingly, 35% of IFPI's respondents stated that they did not use a paid streaming service because anything they wanted to listen to could be found on YouTube. While this does not mean that individuals are necessarily watching music videos on the platform (that is, they may not be actively attending to the visual content), the trend of using the platform for this purpose is still worth mentioning. By 2022, the amount of time spent listening per week had increased to 20.1 hours l (IFPI, 2022). While the IFPI reports do not always provide the same metrics each year, they still provide key insight into global trends of music engagement. For example, the engagement report from 2022 includes percentages for different streaming formats (paid, ad-supported, and video), radio, purchased formats (CDs, vinyl, DVDs, and digital downloads), as well as social media and short form video apps like TikTok. Although video streaming made up 19% of engagement (compared to 24% for paid subscriptions), it was still ranked first in the top five ways people engage with music. However, it is unclear whether this refers to music videos in the "traditional" MTV sense.

There are more ways to engage with music than ever, and individuals can select based on the format that best suits their needs at the time. For example, during the COVID-19 pandemic, many individuals stated that while the amount of time spent engaging with music had not necessarily changed, the way they engaged (which included both playing and listening formats) had (Carlson et al., 2021). Whether this behaviour continued after the pandemic is unclear, however, the IFPI's 2022 report states that individuals use around six different formats to engage with music on average. This figure is even larger for young people

between the ages of 16 and 24 years-old, for whom the average was between seven and eight formats (the exact figure given is 7.7). Young people are also more likely to use social media platforms like TikTok, which allow them to engage with music in visual formats and even create their own. Although these are not the same as our "traditional" music videos (although TikToks occasionally contain music video clips), these engagement formats are indicative of an ever-changing music listening culture that continues to emphasise music and video.

The amount of "music and video" media available to consumers has increased significantly since work on this dissertation began. It is clear that music and video continue to have a symbiotic relationship, and that relationship is stronger than ever. This justifies the need for the research conducted in this dissertation, which aims to provide meaningful insight into the reasons why individuals (who will be referred to as viewer-listeners throughout this dissertation) choose to watch music videos, what makes them so engaging, and under what conditions they have the most salient influence on music perception.

1.2 Music in everyday life

Why do people watch music videos, and how do these reasons differ from audio-only music listening? Although both engagement styles may have potentially rewarding consequences for the individual, the visual component of music video experiences offers something extra that the audio in isolation cannot. This section has two parts: the first outlines the reasons why people engage with music in general before looking at music video experiences specifically. This section will also consider the potential functions or contextual factors that would make music *video* engagement more suitable than audio-only listening. The second part of this chapter explores the affective and cognitive factors at work during music listening and music video engagement. It will explore the mechanisms behind music-evoked affect before addressing the cognitive processes involved in consuming musical multimedia. This section will draw on existing models and frameworks of music and audio-visual cognition which can assist in understanding of the perceptual phenomenon at work during both the initial music video experience and their carry over effects on subsequent listens.

1.2.1 Psychological functions of music

Music has been an important part of human existence throughout history, and technological advancements and portable devices have made it easier than ever to access and enjoy one's favourite music. It is no surprise considered an important psychological resource to many (Denora et al., 1999; Laiho, 2004; T. Schäfer et al., 2013). In 2023, 71% of global respondents to the IFPI's survey on music engagement stated that music played an important role in their mental and physical well-being (IFPI, 2023). Its ability to move people, helping them change,

maintain, or create a desired mood or emotional state or to relieve boredom, make it a useful tool for affect regulation. Furthermore, music can help us reflect, whether it is about ourselves and our place in the world, making it a useful cognitive resource. It can help people foster social bonds and act as a social surrogate (K. Schäfer & Eerola, 2020). This is especially the case during youth and adolescence, when young people wear their favourite music genres as a type of identity badge (Tarrant, 2002). These affective, cognitive, and social functions of music – referred to as the "Big Three" (see T. Schäfer et al., 2013) reflect the profound ways in which music is part of our lives. It may not always be clear, even to the listener, what function the music experience is meant to fulfil during a particular listening episode, however, the ability to select the music we need to hear at a given time may often seem natural or obvious at the time of the listening event (DeNora, 1999; Saarikallio & Erkkilä, 2007).

Music use begins to serve these psychological functions from adolescence and continues into adulthood. It is an important resource for navigating the typical developmental changes associated with youth, such as identity formation, developing interpersonal relationships, and increased agency. Music is used for regulating emotions, helping young people develop self-control and practice agency as they navigate the stress and emotional challenge associated with this time of life (Laiho, 2004). According to Larson (1995), music stands out compared to other forms of media engagement because it speaks to the issues they are facing at the time, unlike other formats like television, which are produced with adult audiences in mind. Interestingly, however, Larson distinguishes the use of Music Television (MTV) from other television channels. Music and other media also act as a resource for social learning, giving young people a source of information about social and cultural trends as well as exposing them to certain attitudes, beliefs, and behaviours (Arnett, 1995). While music has a distinct functional role during adolescence, it continues to serve these functions into adulthood (Groarke & Hogan, 2016; Saarikallio, 2011). Much of the existing research on the topic of music videos has explored their influence on youth audiences, especially in the 1980s and 1990s (Sun & Lull, 1986). However, that generation has now reached adulthood, and the current trends in music consumer behaviour, especially in respect to streaming platforms, suggest that audio-video methods like music videos are not limited to youth audiences (IFPI, 202).

Research which takes a uses and gratification approach has provided useful insights into the reasons why people engage with media, including but not limited to music engagement. Originally used by sociologists investigating the use of mass media to satisfy different needs the uses and gratification framework posits that individuals are aware of their motives when engaging with certain media and actively use it to fulfil their needs (see Katz et al., 1973). This approach also considers why people select a particular medium over others in specific contexts. Contemporary research which takes this approach suggests building on this theory in order to account for new media phenomena. Ruggiero (2000) posits that there are at least three factors involved in Internet media engagement which

are not necessarily inherent in other forms of media (at least not to the same extent): asynchronicity, demassification and interactivity. Asynchronicity reflects how individuals can save, send, receive and retrieve media at their convenience. While this was also possible with, for example, using blank tapes to record programs on television with a VCR, the amount of information which can be interacted with after it was originally made available has been made exponentially more convenient with the Internet. Another example would be the sending and receiving of emails; two individuals can interact with each other at staggered points in time without affecting communication. Demassification refers to the mass availability of media, which gives the individual more freedom to select the media which meets their needs. Interactivity is a multidimensional concept; it can refer to the amount of choice, control and engagement provided on a particular media platform, as well as the role of communication. Like demassification, interactivity gives the user more control over not only which music they select and in what format, it also allows them to interact by sharing what they enjoy with others or communicating in the comment section.

Lonsdale and North (2011) explored the reasons why people engage with music compared to other leisure activities, such as TV watching and video games, using a uses and gratification approach. Their results revealed six factors describing the reasons why participants listened to music: 1) negative mood management, where music is used to help alleviate negative feelings and enhance mood, 2) personal identity, where music is used to portray a certain image and to develop personal identity; 3) surveillance, where music is used to gain information about the world; 4) positive mood management, where music is used to maintain and optimise a positive mood state; 5) interpersonal relationships, where music is used as part of maintaining and promoting social relationships, and 6) diversion, where music is used to pass time, relieve boredom or as a distraction. Furthermore, their results suggest that music was more versatile at fulfilling the needs of the individual compared to other activities, and that it was particularly useful for helping manage and regulate mood states compared to other media.

Music listening often fulfils several functions simultaneously, especially for individuals who are highly engaged with music in general (Greasley & Lamont, 2011). For example, using music as a temporary substitute for social interaction can influence the listener's current affective state by promoting mood regulation strategies such as mental work and solace (Saarikallio & Erkkilä, 2007; K. Schäfer & Eerola, 2020). Music can facilitate the cognitive function of self-reflection by evoking feelings of nostalgia, and trigger the recall of autobiographical memories (Belfi et al., 2016). This may evoke positive affective outcomes by counteracting emotions such as loneliness and sadness (Barrett et al., 2010). Furthermore, even individuals who are less musically engaged may identify up to five reasons for using it to help meet their needs. For the more musically engaged, music can fulfil up to 10 or 12 functions simultaneously (Greasley & Lamont, 2011).

Music's ability to influence affective states, including moods and emotions, is often regarded as its most important function (Laiho, 2004; Miranda & Claes,

2008; Schäfer et al., 2013). Even if the aim of the musical experience is not consciously directed towards affect regulation, music's emotive qualities have a powerful influence on our affective state. It is important to note that affect consists of more than just moods and emotions; all valanced (i.e., positive and negative) states fall under the umbrella term of affect, including emotions, arousal, motivational impulses, stress responses, interpersonal stances, preferences, moods, attitudes, and affective style (Baltazar & Saarikallio, 2016). Interestingly, in Schäfer and colleagues' (2013) discussion of the Big Three functions of music, the use of music for self-awareness (i.e., its cognitive function) also includes the use of music for self-related thoughts pertaining to emotions and sentiments, as well as affect regulation strategies such as solace, diversion and mental work. However, they conceptualise it differently from music's mood regulating function, which highlights the deliberate use of music for regulating moods and physiological arousal, compared to the self-awareness function, which includes statements about the self and the use of music to find meaning in one's life in addition to those strategies which are commonly associated with affect regulation and music (see Saarikallio & Erkkilä, 2007 for examples). The social function, on the other hand, may be easier to demarcate from music's cognitive and affect regulating functions, because it emphasises the use of music to bond with others as opposed to self-acknowledgement and personal well-being.

While Schäfer and colleagues suggest that the social function of music is less important in today's individualist society, current trends suggest otherwise. Music still has an important function when it comes to social learning, which is especially evident when it comes to social media and its role in today's music listening culture. Social media platforms allow users to discuss and share their favourite musical media, as well as participate in online challenges or trends by creating their own content to be uploaded and shared with others on platforms like TikTok and YouTube. Social learning is facilitated, not only by giving users easy access to sociocultural content in the form of musical media, but also by facilitating discourse about this content on the platforms themselves, where they can exchange ideas and content with other users (Vizcaíno-Verdú et al., 2023). This was evident during the COVID-19 pandemic, where a large volume of musical multimedia emerged in the form of dance challenges, informative songs and parodies that fall under the umbrella term of "Covid music" (Hansen et al., 2021). Furthermore, the pandemic was also a time where individuals found themselves trying new or different formats than they usually would in "regular" times: this included both physical media in the form of vinyl or old CDs, but also virtual concerts, livestreams, or music videos (Carlson et al., 2021). The trend towards music multimedia in particular, however, remains relevant today - they started before the pandemic, and they continue to remain relevant now.

Music listening, regardless of what the underlying motivation for the experience may be, can be very rewarding for the listener. Using music to move from a negative affective state to a positive one is perhaps the clearest example, however using music to bond with others or to reflect about ourselves can have

a positive influence on affective states as well. Music videos can potentially fulfil these functions too, and there may be situations where one method of music consumption is more suited to fulfilling a specific function than the other depending on the individual and their needs. Theoretical insight concerning the mechanisms responsible for evoking affective reactions in music listening (and music video experiences) need to be considered. Before these theories are addressed, the next section will explore the findings of one of the first studies conducted about music video engagement in the early days of MTV.

1.2.2 MTV in everyday life: the functions of music video experiences

Music videos may fulfil the same "Big Three" functions as music listening, even though the specific reason for selecting them may differ. For example, Sun and Lull (1986) found that adolescents' motivations for watching MTV differed from those which motivate audio-only listening and TV watching. While their factor analysis revealed that watching for information and social learning explained a significant amount of the variance, the qualitative component of their study provided more detailed descriptions of motivations for watching music videos. Overall, 32 reasons were extracted from their data, which were grouped into nine categories. The most frequently reported motivation for watching music videos was the individual's interest in the musical content: they wanted to see and hear their favourite music. The second most popular category of reasons reflected the motivation of watching for entertainment and enjoyment, which is a typical motivation for music listening and TV watching in general. The third most popular category reflected the importance of the visual component of music videos. The remaining categories included passing time/habit, information (music, concert and consumer information), emotional response, social learning, escape, and social interaction.

Sun and Lull's findings reflect how music videos can fulfil many of the same functions as music listening (and TV watching), including to relieve boredom/pass time, to escape a negative mood, and as a social activity. However, these experiences still stand out from other media formats. The visual treatment of musical content specifically was what attracted young audiences to MTV. Interestingly, adolescents in Sun and Lull's study reported that this visual component made their experience of the music "better", providing a visual representation which influenced how they interpreted the meaning of the music (thirty years later, similar findings were observed in Study I of this thesis). However, Sun and Lull did not address whether this influence on interpretation carried over to subsequent listens.

The adolescents in Sun and Lull's study considered MTV watching to be different from other television programming. They likened it to watching short films with a strong emphasis on the musical component, or like a visually enhanced version of radio: they might tune in when a good song comes on, and leave it on in the background while doing other activities. This passive engagement style is typical for everyday audio-only music listening(Sloboda et al., 2001). MTV watching could also occur sporadically while watching other

television programming; adolescents were not necessarily continuously watching, but tuning in and out depending on what was playing at the time.

The distinguishing difference between reasons for listening to music or watching music videos is the availability of visual information and the opportunities it affords for understanding the meaning of the music. In focused music listening contexts (as opposed to passive music listening), the listener has the freedom to imagine their own meaning of the music. They can find a meaning in the lyrics that is specific to their situation in life, or let their mind wander (van Goethem & Sloboda, 2011). In music videos, the meaning is potentially made more explicit. Videos that feature the artist performing the music also provide extramusical information in the form of performance gestures, allowing the viewer to experience the music differently, almost similar to a live presentation. It makes sense that motivations that emphasise this visual component, whether it is to enjoy the musical content differently or to appreciate the visual depiction of the musical content, are a key factor motivating music video engagement since this information is not available to the listener in audio-only experiences. How this influences music perception, however, is not yet understood.

Sun and Lull also considered the influence of factors like gender, race, peer group relationships, and attitudes towards school in respect to MTV viewership among adolescents. Their findings revealed that there was no significant difference between boys and girls in respect to how much time was spent viewing MTV during the week, however girls reportedly watched more MTV than boys on weekends. In respect to ethnicity, adolescents who identified as Latino were more likely to watch MTV during the week compared to other ethnic groups in the sample (Asian, African-American, and white). Interestingly, a negative correlation was observed between the amount of time spent viewing MTV during the week and satisfaction with school. No significant relationships were observed between school satisfaction and weekend MTV watching.

Unfortunately, studies on music video experiences and the functions they fulfil have been few and far between since Sun and Lull's original study. Despite the change in technology used to access and engage with this form of musical multimedia, there is no reason to expect today's motivations for watching music videos would be much different. However, one distinct difference between music video experiences now and those from the MTV days of the 1980s and 1990s is the technology used to access them. Everyday music listening contexts, including the devices and platforms used, and the control they afford listeners are important factors determining the outcome of the experience for the viewer-listener (Krause & Brown, 2021; Sloboda, 2013). Sun and Lull's study also sheds light on the importance of musical meaning to the listener, since understanding the meaning of the music was a key factor influencing why adolescents engaged with MTV. How an individual perceives the meaning of the music is integral to their experiences, especially in the context of music evoked emotions and other fluctuations on the individual's affective state (Cespedes-Guevara & Eerola, 2018). This information has an influence on the psychological processes, referred

to as mechanisms, behind music evoked affect and are discussed in the next section.

1.3   Experiencing music multimedia

This dissertation aims to gain knowledge into the effects of music videos, including how their content affects: 1) how the music is perceived and their influence on affective states, including emotional outcomes, as a result; and 2) their influence on subsequent listens. Theoretical models and insight concerned with the cognitive and affective components of musical experiences, and the function of music in film contexts, provides a starting point for answering these questions. A key concept which links these two disciplines lies in what emotions are, both in general and music video experience contexts, and how they are differentiated from other affective phenomena. The difference between emotional content in music video experience contexts and the emotions felt by the viewer-listener (also referred to as emotional outcomes) is addressed. This section will also discuss the psychological mechanisms responsible for music evoked emotion. Of particular interest are those mechanisms which are potentially sensitive to or make use of visual information already stored in long-term memory (LTM). This will lead into the next section about the function of music in film, which will provide additional theoretical insight into the role of these mechanisms and other top-down cognitive processes involved during music video experiences. This research provides a tentative explanation for how these experiences can achieve such salient effects and provides a conceptual foundation from which we can begin to understand music video phenomena.

Before addressing the cognitive and affective phenomena relevant to music videos, it is important to first establish a clear conceptual definition of what a music video experience is. While there is probably not one clear answer to any of the research questions posed in this dissertation, a working definition of what these experiences are and what they entail can help provide a starting point for our conceptual illustration of these experiences. A music video experience involves a viewer-listener who, either deliberately or circumstantially, finds themselves engaged with (as in, involved in an activity and showing interest) and actively attending to (as in focusing and applying their mind's energy) both the audio and visual content of a music video over a period of time (whether the length of the video, or until they are no longer watching it – either deliberately or circumstantially). During this time, the viewer-listener is confronted with different types of information, both visual and auditory[1] . This includes the acoustic features of the music, characters and narrative devices, symbolism and/or iconography (including cultural ones), artistic choices and other forms of emotional content.

---

[1] These two are the most important and relevant, textual information (in the form of subtitles for example) is also a possibility.

Emotional content can be considered a catchall for any information which encapsulates, communicates, or can be recognised as portraying or sharing the features of a particular emotion within a specific music video. For example, research has shown that listeners generally agree on which emotion characterises a piece of music based on its acoustic features (Juslin & Laukka, 2004) and that these features are comparative to those used for emotional expression in speech prosody (Juslin & Laukka, 2003). In the case of the visual channel, emotional content may include the facial expressions and other bodily gestures of the characters or performers of the video, or how these characters, objects and other events portrayed in the video interact with each other. While emotional content exists in both modalities individually in their own right, this dissertation is specifically concerned with how emotional information arises as a result of the features of the music and video interacting with each other. Emotional content does not refer to the affective experience of the viewer-listener, but the features of the music and the video that influence emotion induction and other affective and cognitive processes. Emotional content guides cognition and activates the mechanisms involved in appraisal and memory retrieval that contribute to an emotional outcome. Therefore, what is ultimately considered "emotional content[2]" depends on the subjective evaluation of the individual viewer-listener. While there may be considerable overlap between what viewer-listeners agree is the most salient emotional content guiding their attention and subsequent reactions in a specific music video, it is ultimately affected by the individual (including preferences, familiarity, age, gender, empathy and personality) and other situational factors.

During the music video experience, the viewer-listener is constantly processing and evaluating information, including emotional content. Of interest to this thesis is how emotional content influences the viewer-listener's affective state (also referred to as core affect) and potentially culminates to a full-fledged emotional episode. Russell and Barrett (1999) conceptualise core affect as "the most elementary consciously accessible affective feelings… that need not be directed at anything (p.806)." Core affect fluctuates over time, and while it can be directed at anything, it is not a necessity. Like moods, it need not be directed at anything in particular, however, it is always caused by something, even if the root of that cause is not consciously detectable. Core affect can be measured and described by two dimensions, activation and pleasure. These two dimensions are functionally the same as the valence and arousal dimensions frequently used to measure music evoked emotions in self-reports (see Vuoskoski & Eerola, 2011). On the other hand, emotional outcomes (referred to by Russell and Barrett as prototypical emotional episodes) occur less frequently: these are noticeable changes that occur in an organism (i.e., the viewer-listener) in response to a stimulus object, whether real (physically present) or imagined. Prototypical

_____

[2] An argument can be made that this should be labelled as "affective content" and not "emotional" content. The justification for the choice presented here is that calling it affective content in this context is so vague it is functionally meaningless. Emotional content is the content with the potential to influence which emotion mechanisms are activated both during the experience and in subsequent listens.

emotional episodes require the occurrence of several cognitive processes (including attention to, awareness and appraisal of the object), behavioural responses, and physiological changes. The absence or modification of any of these emotional episodes is conceptualised as a non-prototypical emotional episode, which occur more frequently than their prototypical counterparts (Russell & Barrett, 1999).

This approach to understanding music evoked changes in affect provides a starting point for conceptualising how attention is directed towards a music video's emotional content, allowing for changes in core affect which, under the right conditions, may culminate into a prototypical emotional episode. These changes in core affect occur throughout the experience as a result of the constant feedback loop, where the viewer-listener is constantly evaluating the content of the music video in respect to their own personal values and the goals which motivated them towards the experience in the first place. Appraisal theorists suggest that the type of emotions evoked by music depend on how the listener evaluates the musical stimuli and whether it is compatible with their emotional goals, whether it is conducive of their own standards and values, and how they cope with the consequences of their goal being reached or obstructed (Scherer, 2005; Scherer & Zentner, 2001). While this thesis primarily takes a constructionist approach to understanding music video phenomena, the importance of the individual's subjective feelings and larger-scale cultural norms and societal values during the appraisal process sheds light on the cognitive mechanisms related to associations and memory (which will be discussed in greater detail below). This is especially pertinent in music video experiences, since it is not only the music which needs to be evaluated, but the visual content as well. The viewer-listener's subjective feelings about the objects in the video therefore have the potential to drive what is associated with the music in subsequent listens. Whether this has a positive or negative effect on the viewer-listener's affective state depends on whether or not they are compatible with their previously held expectations about the music.

1.3.1   Regulating affect with music

Music's ability to influence affective states and evoke powerful emotions has been established as one of its most important functions. Different musical activities may be more useful for helping the individual meet certain needs over others, depending on their affect regulation goals. Individuals often have an idea of what kind of music they need to listen to at a particular moment, even in the absence of a specific affect regulation goal (Saarikallio & Erkkilä, 2007). Research examining music's ability to regulate affective states, including emotional reactions, and research concerned with the mechanisms underlying these processes provides a starting point from which music video's influence on perception and affective outcomes can be explored.

Van Goethem and Sloboda (2013) outlined four levels of analysis for understanding music's use for affect regulation: goals, tactics, strategies, and mechanisms. Music listening and music video watching are examples of tactics,

which are used to meet a particular affect regulation goal, for example: to relax after a long day. The individual then uses a strategy like revival or diversion. Mechanisms are the processes which enable regulation, and there are two different types: those that are related to the content of the music itself (intrinsic mechanisms) and those which are based on the memories and experiences of the listener, and other situational features (extrinsic mechanisms). The mechanisms at work in a particular listening activity depend on the regulation goal and strategies employed.

Music video experiences are an example of a music listening tactic which employs many of the same or similar strategies and mechanisms as audio-only listening. For example, Sun and Lull's (1986) study found emotion regulation goals were associated with music video watching, including those related to escaping negative thoughts and emotions, and boredom relief. Categories of reasons related to entertainment, emotional responses, passing time, and escape are similarly reflected in Saarikallio and Erkkilä's (2007) strategies of music and affect regulation (entertainment, revival, strong sensation, diversion, discharge, mental work and solace). However, the viewer-listener is constantly appraising and reacting to the emotional content of the music video, which may cause changes in core affect that culminate into prototypical emotional responses.

Baltazar and Saarikallio (2019) suggest a model of music's use for affect regulation which explores combinations of strategies and mechanisms across three components of affect: cognition, feelings and bodily reactions. Two groups of strategies emerged: analytical and change strategies and repairing and pleasure strategies. Mechanisms were also grouped according to whether they are dependent on the individual or the features of the music. Analytical and change strategies were associated with individual-dependent mechanisms, whereas repairing and pleasure strategies are related to feature-dependent mechanisms. For example, an experience which emphasise the bodily reaction component of affect regulation (as opposed to focusing on the task at hand) is associated with repairing strategies like revival and individual-dependent mechanisms such as contagion: the process of inducing emotions whereby the listener mimics the emotion they perceive as being expressed in the music (Juslin & Västfjäll, 2008). This research provides insight into the possible strategies being used by viewer-listener's during music video experiences and their relationships to affective outcomes. The cognitive, affective and bodily components of music video experiences may be more salient, depending on how its emotional content is appraised, and whether that appraisal is influenced by the viewer-listener's pre-existing perceptions about the music.

While Sun and Lull's (1986) study found that music video experiences were used to fulfil similar affect regulation goals in respect to entertainment and aesthetic experiences, whether the music video was successful at fulfilling their needs was not addressed (it was not the aim of their study). Of particular interest to this dissertation are the mechanisms involved in influencing changes in core affect during music video watching and their relationship to the mechanisms activated in subsequent listens. These mechanisms may also be related to specific

intentions for engaging with music video experiences. Current research on the psychological mechanisms involved in music evoked emotion, particularly those involved in top-down cognition and imagination processes are of particular interest. These mechanisms, including BRECVEMA mechanisms like visual imagery and episodic memory (Juslin et al., 2013; Juslin & Västfjäll, 2008), provide a starting point for understanding how music video information affects music perception and listening outcomes.

1.3.2   The influence of extramusical information

Music videos contain extramusical information, which research has shown has the potential to influence how a song's emotional quality (or meaning) is perceived (Vuoskoski & Eerola, 2015). Music video experiences may have the potential to influence the viewer-listener's affective reaction to the music, including emotion induction, during both the initial music video experience and in subsequent listens by affecting how musical information is encoded and retrieved from long-term memory. An understanding of the underlying affective and cognitive mechanisms at work during music listening experiences and how they may be affected by visual information in music videos provide a starting point for understanding these phenomena.

As it was defined above, a music video experience involves the viewer-listener attending to and engaging with the audio and visual content of a music video over a period of time. The events of the music video (which includes emotional content) can be conceptualised as a micro-context within the larger situational context of the viewer-listener who, by using their imagination, is able to experience the "other world" offered in its content. The extent to which perception is influenced, however, is determined by interactions between the individual viewer-listener, the content of the music and video (the events of the "other world"), and the context where the experience takes place. Hargreaves (2012) refers to these as the determinants in his reciprocal feedback model (RFM) of music processing, where any one determinant exerts an influence on the other two. For example, the preferences of the listener influence what type of music they will select depending on its appropriateness in a particular environment. According to this model, it is the interactions between networks of association between these determinants which influence the type of cognitive, affective and physiological responses experienced by the listener. The imagination is the operating system through which these interactions occur (Schubert, 2013). Hargreaves (2012) describes imagination as the "cognitive basis of musical perception and production (p. 540)." Imagination has a pivotal role in how associations at the personal, cultural, and structural (i.e., the music's features) are retrieved from memory when processing information during music listening. As the cognitive basis for music perception, imagination has a key role in activating the cognitive and affective mechanisms involved in processing different sources of information (whether contextual, feature based, or memories) during music video experiences. Imagination has an important function in respect to memory processes (including formation, storage and retrieval), and cognitive processes

involved in fantasy and make-believe, which have an important function in respect to the formation ideas, concepts or mental images, even in the absence of external stimuli. These mental phenomena also play a key role in the emotion induction process.

The underlying psychological mechanisms responsible for music induced emotions outlined in the BRECVEMA framework listening (Juslin et al., 2014; Juslin & Västfjäll, 2008) shed light on the role of these mental processes. This framework outlines eight mechanisms: brainstem reflex, rhythmic entrainment, evaluative conditioning, contagion, visual imagery, episodic memory, musical expectancy, and aesthetic judgement. These mechanisms, in addition to cognitive appraisal, provide an explanation for the processes by which emotions are induced in response to musical features in a given listening context.

While all of these mechanisms can be applied to music video experiences in some form, of particular interest are the mechanisms: visual imagery, episodic memory and aesthetic judgement, due to their relationship with imagination (i.e., fantasy and memory) processes and the viewer-listener's subjective evaluations. Visual imagery refers to the process of emotion induction as a result of the internal images experienced in the listener in the absence of any external stimuli. These images reflect a metaphorical mapping of the music's emotional characteristics (Juslin & Västfjäll, 2008). The experience of visual imagery also involves the activation of other cognitive mechanisms involved in emotion, memory, and motor control (Kosslyn et al., 2001). The experience of vivid mental imagery for the purpose of regulating emotional arousal has been associated musical training (Küssner & Eerola, 2019). Furthermore, musically trained students have been found to perform better at cognitive tasks which require creating, maintaining and transforming visual imagery (Commodari & Sole, 2020). Episodic memory refers to emotion induction as a result of the music triggering the recollection of specific memories associated with the music. These memories may evoke feelings of nostalgia depending on their autobiographical salience and the individual's proneness for nostalgia (Barrett et al., 2010). Episodic memory reflects imagination processes related to memory storage and retrieval, whereas visual imagery reflects its fantasy and make-believe processes. Aesthetic judgement reflects the subjective feelings of the individual towards the music based on factors such as preferences and other criteria (Juslin et al., 2014).

There are many overlaps between visual imagery and episodic memory mechanisms, and the differences may be particularly ambiguous for those who have a congenital lack of visual imagination (a phenomenon referred to as aphantasia, see: Keogh & Pearson, 2018)[3]. For the purpose of this thesis, the key difference between these two mechanisms is in the personal component of episodic memory. Visual imagery that reflects the content of the music video may be more accurately conceptualised as episodic memory if these visual images are

---

[3] This is, at least in part, due to my own anecdotal and practical experience with these phenomena as someone with a strong visual memory. If a piece of music reminds me of an event in my life, I experience that memory as visual mental imagery. I have a close friend with aphantasia (who gave me permission to mention this in this footnote) who can remember the source of their associations but never as visual imagery.

associated with the specific memories of the individual. The effect of the video's visual representation of the music's meaning being congruent with the associations already attached to the music influence the song's personal significance to the listener. Music video content may be retrieved from memory and manifested as visual imagery in subsequent listens without contributing to the personal importance the individual places on that specific piece of music (the importance of this distinction will become clearer in the next section of this chapter, which addresses the audio-visual perception literature).

1.4   Music in film

The visual component of music video experiences is the key factor which differentiates them from other forms of music engagement. Existing research concerned with musical multimedia, including the perceptual and emotional effects that occur as a result of pairing music with visual stimuli provides a starting point for understanding music video experience effects. This research sheds light on the cognitive processes at work during multimedia experiences, including how attention is directed towards information from multiple modalities, and how information from one modality contributes to the perception of the other. This section will also outline how the same or similar mechanisms involved in music listening function in music video experiences, particularly those involved in imagination processes.

Existing models and theories of the cognitive and affective processes at work during music listening and film-viewing provide a starting point for understanding how music videos influence music perception. The exact effect a music video may have in subsequent listens is contingent on many factors, many of which are beyond the scope of this dissertation. However, existing research concerned with the psychological (including affective and cognitive processes) and physiological processes involved in consuming bimodal stimuli like music videos provides a foundation from which we can begin to generate hypotheses about these effects and the conditions under which they are most salient. This section will conclude by discussing the cognitive processes involved in audio-visual perception and their effect on which psychological mechanisms are activated in subsequent listening experiences.

1.4.1   Congruence-Associationist Model

Music's ability to direct attention and influence core affect is also an integral component of its function in film. The word film in this context refers to any multimedia which pairs music and moving pictures (i.e., video) for entertainment purposes, including music videos, advertisements, and Hollywood films. One of the most cited frameworks concerned with the function of music in film contexts is Annabel Cohen's Congruence-Associationist Model (CAM; Cohen, 2013; Marshall & Cohen, 1988). The cognitive processes outlined

in CAM provide a starting point for exploring the possible effects of music video experiences on music perception. Additional models and research which relate to or build on CAM concepts are also highlighted.

The CAM framework outlines how music guides attention and supplies emotional information during film experiences. Both music and film consist of structural and meaning aspects (Cohen, 2013). Attention is directed towards shared structural features in both the music and the film, such as shared accent patterns and synchronicity, as they unfold over time. This gestalt principle of grouping by similarity reflects the congruence component of the model. The music's emotional qualities provide an emotional context for understanding the events of the film as a result of associations brought to mind by the music. CAM outlines how the perception of sensory information (bottom-up cognition) and information retrieved in long-term memory (top-down cognition) are used to construct a working narrative of the film. Cohen (2013) posits that the concept of a "working narrative", which is constructed by combining the relevant audiovisual information currently available at the level of conscious experience, is loosely related to that of working memory. CAM has useful applications in respect to understanding how music video content is processed. The outcome of these processes, however, depends on the viewer-listener's perception of the emotional content reflected in the structural features of the music and video and any pre-existing associations they have about the music.

Evidence from existing research suggests that perceived congruencies between the music and the visual channel will have a key role in directing attention and influencing core affect during music video experiences (Marshall & Cohen, 1988). Research has also shown that music can decrease the time it takes for the viewer's gaze to fixate on particular objects in the film, as well as accentuate the emotional impacts of those objects (Millet et al., 2021). Iwamiya (2013) refers to this temporal alignment of structural features as formal congruence, which has an important role in combining audio-visual information into a unified percept. Formal congruency has also been shown to enhance viewers' evaluations of multimedia, suggesting that synchronising audio and visual structures can create a more impactful experience (Kim & Iwamiya, 2008). On the other hand, semantic congruence reflects the perceived similarity between the affective components of the audio and visual modalities. For example, when pairing a computer-generated version of a piano performance under different colours of light, semantic congruence perceived between the music and the colour influenced the impression of the music. Brighter colours (for example, green, yellow, and cyan) influenced the affective impact of pieces in a major key and fast tempo (Iwamiya, 2013). This sheds light on how subjective congruence can influence the overall impact of an audio-visual scene. While the relationship between specific patterns of structural features and their influence on music perception, enjoyment, and affective reactions are not specifically addressed in the studies presented in this thesis, the role of subjective congruence in directing attention and core affect processes is integral to its goals.

The association component of CAM highlights how information stored in long-term memory (LTM) influences audio-visual perception. For example, extramusical concepts related to the genre of a piece of music playing during a movie trailer can influence expectations about the genre of the movie, as well as the feelings or ideas presented in the movie (Strobin et al., 2015). Extramusical associations about specific genres of music are activated in the form of cognitive schemas; these reflect networks of ideas, feelings, attitudes, and knowledge about concepts related to a particular genre. Cultural stereotypes related to genre are an example of a negative form of cognitive schema (Shevy, 2008). These schemas also apply to the visual content as well and may play a role in the subjective evaluation of a music video's content based on its alignment with their personal values and beliefs (Manolika et al., 2021).

Music can clarify the emotional tone of film events and provide viewers with an interpretative framework for understanding the characters' behaviours and motivations. It may also affect the viewer's interpretation of what occurred in the film, including the personality and motivations of the characters, in a way that is congruent with the mood of the music (Boltz, 2005). Furthermore, music can bias attention towards elements of a visual scene which are coherent with or embellish the emotional tone of the music, resulting in better memory performance when recalling the events of the film (Boltz et al., 1991). However, visual information can also influence how music is perceived. Boltz (et al., 2009) found that visual information may bias music perception in such a way that it is congruent with the emotional tone of that visual information. Participants falsely identified a piece of music as being faster or slower than the original presentation depending on the valence of the visual content: music presented with low valence visual content were identified as slower, and music presented with high valence visual content were identified as faster. This occurred regardless of whether the visual information was presented in the form of a video or as a montage of still images.

Boltz suggests that the valence of the music accompanying the film biases what the viewer remembers and how they evaluate the events of the film in a mood-congruent way. Furthermore, it can guide attention to certain objects in the film that are consistent with their interpretative framework; this attentional highlighting facilitates subsequent recall because the information from both modalities has been integrated into a single memory code (Boltz, 2004; Boltz et al., 1991; Boltz et al., 2009). This insight sheds light on how music video content may potentially music perception in subsequent listens as well, depending on whether the perceived emotional impact of the audio-visual pairing.

1.4.1.1   Music video content

The visual content video and its perceived structural and semantic relationship to the music reflect the feature-dependent determinants of the experience. This content can be categorised as narrative or corporeal; narrative content reflects information in the video that provides context about the emotional content or meaning of the music, whether it is in the format of storylines or through audio-

visual iconicity (see Kendall, 2005), whereas corporeal content reflects human gestures. Corporeal information may also be further categorised according to whether this corporeal gesture reflects musical movement such as scenes of the musician or musicians performing the music and choreography, or expressive gestures which, although not related to the production of the music, provide emotional information through bodily expression.

One of the most popular reasons for music video watching includes watching in order to understand the meaning of the music (Sun & Lull, 1986). Videos which depict narrative content, such as storylines or the use of culturally symbolic and/or iconic imagery, can provide viewers with a source of meaning which they can use as a framework for understanding what the artist is trying to convey. Storylines and narratives are a strong example of how music videos can provide additional context about the meaning of the music. The story, its characters, and the outcome of the music video may then provide a framework for understanding how the artist or artists intended the meaning to be understood. One example would be to provide a story which visually depicts the content of the lyrics; however, this is not necessary for a music video to be impactful.

1.5   Complementing existing models

Both CAM and the reciprocal feedback model shed light on the role of associative mechanisms involved during music and music video experiences: information stored in LTM activates cognitive networks associated with that information which, through the cognitive processes of the imagination, guides attention and influences changes in core affect. Schubert's (2014) spreading activation theory explains how, in a given context (for example, music listening or music video watching) associated meanings activated in the form of mental representations, referred to as nodes, influence the experience. Through exposure and learning, individuals develop their own networks of nodes. Spreading activation theory provides an explanation for aesthetic pleasure experienced in response to musical activities: greater hedonic outcomes occur as a result of activating larger networks of association. This also explains the importance of preferences in experiencing hedonic outcomes from music and music videos.

The importance of imagination and its relationship to memory was first applied to movie-going experiences in the earliest theories on the psychology of film: "The Photoplay" by Hugo Münsterberg (1916). Imagination transports the viewer into the diegesis of the film. However, while their subjective thoughts, feelings and associations about the content of a film influence their perception of the film's events, these existing associations are only supplementary information. It is the content of the film which ultimately exerts the most control over what ideas the viewer is meant to associate with its content. The psychological processes involved in attention, perception, emotion and memory contribute to how the imagined world of the film is experienced as a result of top-down

mechanisms providing supplementary information that can assist the viewer's comprehension of the events that unfold on screen. In doing so, the film can free the imagination from the constraints of real-world expectations and provides gratification in the form of aesthetic pleasure, emotional outcomes, learning, and changes in arousal (Tan, 2018).

The models of multimedia put forth by Cook (1998) provides a framework for theoretical discourse concerned about audio-visual interactions. This framework provides both an inventory of the typical relationships between the different media within a multimedia context, as well as a set of terminology for describing them. This framework consists of three models of multimedia, conformance, complementation and contest. These models describe the semantic (meaning) relationships between the different constituent media in respect to their differences and similarities. The conformance model, for example, reflects when the meaning conveyed in one media can be seen as a projection of the meaning conveyed in another, and vice versa. This does not mean to say the audio and the visual share all of the same structural similarities, only that they are compatible with the same meaning. Cook illustrates conformance using the constituent media of Kandinsky's Der gelbe Klang as an example:

"...one might speak equally well of the coloured lights … projecting the stage action, or the stage action projecting the music, or the music projecting the colours; that is what it means to say that all constituent media of Der gelbe Klang are conformant with one another, and that Der gelbe Klang is an instance of conformance overall… each medium of Der gelbe Klang is congruent with each of the others; it embodies the same spiritual content (pg.100-101)."

This type of unitary conformance is rare in actual multimedia. For this reason, Cook posits that conformance is less applicable for modelling the relationships within an entire multimedia context, but instead represents a model of the relationships between different constituent media within the larger multimedia context. Conformance occurs when the constituent media pass what Cook refers to as the similarity test. However, since complete conformity is rare, what is more relevant (especially in music video contexts) is the difference test.

The difference test and the models that stem from it provide a conceptual framework for understanding how audio and visual information interact to influence perception and affective outcomes in music video experiences. The outcome of the difference test establishes whether the relationship between the constituent media (i.e., within a music video) is a relationship of contrariety or contradiction. If the content of the music and video are perceived as similar in their character or connotation, then it is a relationship of contrariety: it fails the difference test. Contrariety occurs when the same semantic meaning can be inferred from the underlying structural content of the constituent media; they blend together to form a single semantic structure. Contradiction, on the other hand, suggests that there is a direct conflict between constituents. If the relationship is one of contradiction, then it passes the difference test and falls under the contest model.

In the contest model, the different channels of media (the music and the video) each embody their own separate semantic structures. Each media, through its own structural characteristics, tries to impose its own meaning upon the other. Interestingly, Cook asserts that multimedia where one media (for example, the music) already existed before the production of the other makes for particularly interesting examples of contest, using the music video for Madonna's Material Girl as an example. By adding a cinematic context to the meaning already recognised in the music, a third meaning is formed: this is the outcome of contest. However, there are instances where the constituent media are neither entirely consistent nor contradictory. In complementation, each constituent media represents its own meaning with a unique role within its larger multimedia context. Cook equates this to the role of music in the Hollywood film, especially in the earlier days where the score would be composed to match the final visual product. Although the role of the music may be considered a supporting one, the final meaning of the film is not complete without the additional meaning provided by the music. While Cook views the audio-visual interactions that occur in respect to meaning generation during music videos as representative of the contest model of multimedia, there may be contexts where the complementation model better reflects the relationship between the music and visual components of the experience. If the music and video are perceived as adding something to the experience that the other cannot, then complementation applies. If the viewer-listener has a previous association with the music which is at odds with the content of the video, a third meaning is formed. The appraisal of the visual content relative to the original meaning associated with the music forms the basis of this new meaning.

Music's ability to promote feelings of pleasure and eudaimonic enjoyment are part of what makes it such a versatile psychological resource(Lamont, 2011; Mas-Herrero et al., 2013). A defining component present in all the models and frameworks outlined in the sections above is the importance of associations and the construction of meaning. These meanings contribute to bottom-up and top-down cognitive processes involved in directing attention and constructing a narrative. The viewer-listener then appraises this content in respect to their situational context and the reason for watching (goal). The extent to which this experience is perceived as pleasant or unpleasant, the potency of that (dis)pleasure, and the relative importance of any pre-existing associations the viewer-listener may have with the music determines the affective response. In this sense, music listening can be conceptualised as a skill with the potential for adaptive or maladaptive consequences on well-being (Reybrouck & Eerola, 2022). The same may be said about music video experiences, particularly among young people for whom music media may become a source for self-comparison and social learning (Kistler et al., 2010). A better understanding of music videos experiences and their effects may help mitigate the potential negative effects that arise as a result of maladaptive music listening (and general media engagement) habits.

2   RESEARCH QUESTIONS AND METHODOLOGIES

This dissertation investigates music video engagement and perceptual phenomena by addressing the following research questions:

- What are the functions of music video experiences?
- How does extramusical content (i.e., visual information) influence music perception in music video contexts?
- How does exposure to music video content influence how music is perceived in subsequent, audio-only personal listening experiences?

One of the distinguishing features of this dissertation is its use of qualitative and quantitative methodologies. This mixed methods approach allows for a more thorough investigation of music video experiences. All four studies were data-driven, with each study building off the findings of the last. Specific research questions for each study are outlined in their respective sections below.

Each study uses a different methodological approach, with the first study being the most exploratory and the fourth being hypothesis driven. Study I features a qualitative analysis of data collected from an open-ended questionnaire designed to get an understanding of participants' experiences with music videos. It broadly addresses all three research questions presented above, in order to gain insight into the nature of music video experiences and the scope of the perceptual phenomena they elicit.

The second study explores these phenomena further and explores their relationship to individual differences in music listening for emotion regulation behaviours. Study III consists of a survey about music video experiences whose items were developed based on the key findings from the study I. Finally, Study IV analyses data from real-time, everyday listening episodes in order to determine the effects of music video exposure on listening outcomes in personal listening experiences. This final study uses an experience sampling method (ESM) in order to explore the theoretical insights brought forth in the previous studies in a more ecologically valid setting.

2.1 Study I – Qualitative questionnaire

The first study was exploratory and had no specific research question. The purpose of the study was to gain as much qualitative insight as possible into music video experiences, including why participants watch them, what kind of thought processes occur during these experiences, their affective reactions, and whether their future listening experiences were influenced by what they saw. The findings from this study set the foundation for the subsequent studies in this thesis by identifying the most salient conceptual phenomena which could be explored further using different methodologies.

2.1.1 Participants

Participants were recruited from a university e-mail list of approximately 100 music theory students and via social media. This resulted in a convenience sample of 34 participants between the ages of 15 and 27 (M = 22.4, SD = 2.79; 53% female). The majority of participants identified as Canadian (53%), eight chose not to identify their nationality and the remaining three identified as American, Australian and Korean. Two participants disclosed having a second nationality in addition to being Canadian (Chinese and Dutch).

2.1.2 Study Design and Procedure

The study featured an online questionnaire, which consisted of 12 open-ended questions and took approximately 25 minutes to complete. The study was hosted on an online platform so that participants could complete the study in an environment where they would usually watch music videos (MVs), and to minimise experimenter influence. It was important to create an authentic environment, since they were also asked to watch a MV before they filled out the questionnaire; this elicitation technique promotes discussion and elaboration of ideas and reduces recall bias (Barton, 2015).

The questions were designed to elicit descriptive responses from participants about all aspects of their music video experiences. Participants also indicated which devices they use to watch MVs on, including Smartphone, computer/laptop, tablet, and "other". If the participant selected "other", they were prompted to type the name of the device.

2.1.3 Coding procedure and analysis

The study used an abductive grounded theory approach in order to gain insight into music video experience phenomena. This type of qualitative approach is especially useful for exploring topics which have been underexplored or where understanding is currently limited. Abductive analysis requires the researcher to limit their preconceptions about the problem being investigated (Charmaz, 2006). Grounded theory starts inductively; however, it evolves into abductive reasoning

as the researcher attempts to interpret and understand the data using alternative theoretical explanations (Timmermans & Tavory, 2012) . This process of defamiliarizing and revisiting the data from different theoretical perspectives was appropriate, since many components of MV experiences have previously been explored, even if they have not been applied directly to MVs in an empirical research setting. Topical areas of research which informed the analysis include: personal music listening in everyday life, the psychological function of music listening, theories of music evoked affect (including emotions), and the cognitive encoding of audio-visual stimuli.

The first step of the coding process was to identify 1) the main reasons why participants choose to watch MVs, 2) the cognitive and affective phenomena that occurred during the MV experience itself, and 3) the carry-over effects imposed on future listening experiences. Once the data were organised into these three temporal-categories, they were sub-coded further using the in vivo method. In vivo coding preserves the language used by the participants, keeping the analysis grounded in the data (Charmaz ,2008; 2006). If a code described multiple aspects of the experience or inferred multiple meanings, simultaneous coding was used (Saldana, 2016). The second step of the analysis was to perform axial coding, which involves grouping codes with similar meanings together. It was during this phase that the main categories and subcategories within each temporal level were formed. All coding was performed on NVivo (version 12).

2.2 Study II – Mixed method exploratory study

The second study builds on the first study's findings by triangulating the main qualitative categories established in the first study with quantitative data (previously validated psychometric scales) collected from the same participants. In doing so, a preliminary investigation into the potential relationships between specific aspects of music video experiences and personality traits, as well as between experience aspects and healthy and unhealthy music-listening behaviours could be explored. The study did not have any specific hypotheses in respect to these possible relationships.

2.2.1 Design

The second study was an extension of the first: it used the same data and qualitative analysis, however it also included additional quantitative data which were not explored in Study I. This was in order to explore the potential relationships between music listening for emotion regulation behaviours, personality, and the categorical data established in Study I. The same participant sample was used; however one participant was omitted from the analysis. This was because only select categories were transformed into nominal data, and this participant did not have any data coded in those categories.

Select thematic categories from the four temporal levels established in Study I, labelled Intention, Attention, Reaction and Retention (IARR) framework, were chosen for this analysis. At the Intention level, internal and external goals were selected. At the experience level, which consists of Attention and Reaction levels, only the Reaction questions were selected, since this study was primarily interested in how MV experiences reflected emotion outcomes. Analyses were performed on the Strong Affect sub-category. An additional sub-category that was not explored in the first study was included, labelled Unimpacted in order to compare the scale measure data for participants who experienced strong affective outcomes from MV experiences with those participants who reported that MV experiences were less emotionally salient than audio-only listening experiences. The individual Retention categories (New Interpretation of Meaning, New Affect Perception, and Visual Mental Imagery) were not analysed in this study; instead, the duration of effects (long-term or unaffected) was analysed.

2.2.2 Materials

Participants were instructed to complete two psychometric scale measures prior to completing the open-ended questionnaire. Personality traits were measured using the Ten-Item Personality Index (TIPI; Gosling et al., 2005). However, since the sample was small and the study was specifically interested in the relationship between emotional use of music and MV outcomes, only emotional stability scores were analysed. Use of music for emotional health and well-being were measured using the Healthy Unhealthy Music Scale (HUMS; Saarikallio et al., 2015). Qualitative categories selected for analysis were internal and external goals (Intention level), strong affect and unimpacted (Experience-Retention level) and duration of effects (Retention). These categories were transformed into nominal data for the purpose of analysis. All statistical analyses were performed on Jamovi (v2.1).

2.2.3 Analyses

The study triangulated qualitative findings from Study I with the scale measure data in order to explore the potential relationships between individual differences in these traits and MV listening experience behaviours. Independent sample t-tests, Mann Whitney u-tests and one-way ANOVA were performed in order to explore the relationships between both sets of data.

2.3   Study III – Online survey

The third study investigates the characteristics of music video experiences and their effects on subsequent listens further using quantitative methods. The study features an online survey with items based on the main thematic categories of each level of the IARR framework in order to gain more insight into the potential

relationships between thematic concepts at different temporal levels. The study used principal component analysis (PCA) to reduce the dimensionality of items. Correlations and linear regression analyses were conducted in order to explore the possible relationships between principal components.

2.3.1 Research questions and hypotheses

The study had two research questions:
- What are the key characteristics of music video experiences (Experience components) and their carry-over effects (Retention outcomes)? (RQ1)
- How are Experience components related to Retention outcomes? (RQ2)

For RQ1, it was hypothesised that Experience components would reflect themes related to the content of the music video, such as narratives and performance content, and intention goals related to emotion regulation, cognitive goals related to understanding the meaning of the music, and social goals related to social media engagement. It was hypothesised that Retention outcomes would reflect effects describing how personal associations with the music and the visual imagery mechanism were influenced by the music video's content in subsequent listens. RQ2 was mostly exploratory, however, a strong relationship between experiences emphasising the importance of visual content that outlines the meaning of the music, whether by providing context about the meaning of the lyrics or via performance gesture, was anticipated.

2.3.2 Participants

Participants were recruited via Survey Circle (www.surveycircle.com), Twitter, Facebook, and JYU e-mail lists. The first 25 participants to complete the survey were offered a €5 Amazon gift card voucher, and all participants who provided a valid email address were entered in a raffle for a chance to win an Amazon gift card valued at €25.

The final sample consisted of 155 participants after screening for careless responding and removing all those who did not complete the entire survey. Participants were between the ages of 15 and 44 (M= 27.3, SD = 5.88). A total of 39 countries were represented in the sample. 62.3% of participants reported they lived in the EU. The countries with the highest number of participants were Germany (19 participants), followed by Finland (18 participants) and the UK (18 participants). 66% of the sample identified as female. 44.5% of participants identified as music-loving non-musicians.

2.3.3 Design and Procedure

Questionnaire items were based on the main thematic categories within each temporal of the IARR framework identified in Study I. The questionnaire was divided into two sections: the first section included items describing intention, attention and reaction-level themes, while the second section included items

describing retention-level themes. Participants answered each question by indicating how much they agreed with each statement on a five-point Likert scale (1 = disagree strongly, 5 = agree strongly). The questionnaire was hosted on the online survey platform webporol.

All participants indicated their consent to participate on the study's landing page, which also contained all instructions and details about the purpose of the study. The study took approximately 20 minutes to complete.

2.3.4   Analyses

Two separate principal component analyses were performed to reduce the dimensionality of the items. The first analysis included all items describing intentions and experiences (referred to as Experience components), while the second   analysis contained all items describing Retention outcomes. Since correlations were expected between components, Promax rotation was used to increase interpretability of the loadings. Sampling adequacy was evaluated using the Kaiser-Meyer-Olkin (KMO <.60; Kaiser, 1974). Once the dimensionality of items was reduced for both surveys, correlation and linear regression analyses were performed in order to explore the relationships between variables at both levels.

2.4   Study IV – Experience Sampling Study with MuPsych

Study IV was designed with the aim to explore whether the perceptual effects of music videos identified in the previous studies affected music listening experiences in everyday life. The analysis featured here is from data collected during the pilot phase of the study. Study IV was hypothesis driven: its primary purpose was to examine whether music video content influence the subsequent enjoyment of and emotional reactions to the music in audio-only listening experiences, and whether these outcomes could be explained by changes in personal significance attached to the music (Personal Significance), how the individual perceives the music's meaning (Interpretation), and the extent to which visual content is recalled as visual imagery (Image Recall).
The study also explored interactions between Image Recall with both Personal Significance and Interpretation predictors, in order to explore any possible moderating effects of these variables on the strength of each outcome. Findings from this study were discussed in respect to how visual information in the form of performance gestures and narratives depicting the meaning of the music can be applied in music education practices.

2.4.1   Research questions and hypotheses

This study had two research questions:

- How do changes in personal significance, interpretation of meaning, and visual imagery from the music video influence enjoyment in personal music listening experiences?

- How do changes in personal significance, interpretation of meaning, and visual imagery from the music video influence emotional reactions in personal music listening experiences?

For the first research question, it was hypothesised that Personal Significance would have an effect on enjoyment, and that this effect would be stronger as Image Recall increased (H1). For the second research question, it was hypothesised that changes in Interpretation and Personal Significance would have a stronger effect on emotional reactions as Image Recall increased (H2).

2.4.2   Participants

Data from 15 participants between the ages of 18 and 37 (M = 21.7, SD = 4.56) were collected. The most frequently reported nationality of participants was Finnish (N = 4), followed by Canadian (N= 2) and American (N = 2). Participants were recruited via email lists for doctoral students and music students at the University of Jyväskylä, as well as through Google (AdSense). Two participants did not complete the demographic information survey. All participants who completed the demographic information portion of the study had stated having at least a high-school level education.

2.4.3   Design and Procedure

The study was completed entirely on the MuPsych application for Android smartphones. This allowed participants to complete the study while they were listening to music as they normally would in everyday life. The application works by prompting the participant to answer questions directly on their device at three time points: the first prompt occurs at the start of their music listening episode, the second prompt occurs at the five-minute mark, and the third prompts after approximately 30-minutes, if they are still listening to music. At least the first and second prompts need to be completed: this set of questions represents one experience sampling report (ESR). Participants needed to complete at least 20 ESRs over the course of a two-week period in order to be entered into a draw to win one of two €25 Amazon gift cards.

The first screen consists of standard MuPsych items, which collect self-report data about the participant's current affective state. Two seven-point Likert scales appear on the first screen (titled "How do you feel?") which assesses initial valence (labelled at points 1, 4 and 7: very negative - neutral - very positive) and arousal (labelled at points 1 and 7: very low energy - very high energy). The second and third screens were designed specifically for Study IV. The second screen, which prompts at the five-minute mark, assesses valence and arousal for a second time with the same 7-point Likert measures as the first screen. After this,

a screen is presented which asks the participant if they have seen the music video for the song they were just listening to when the ESR was prompted. If the participant indicates that they were not familiar with the music video, they are asked if they are familiar with the music video for the song they were listening to at the start of the music listening episode (i.e., the song that was playing when they filled out the screen with the first set of questions). If they had seen a music video for either of these songs, the next screen consisted of the following questions: "How has the music video influenced your current listening experience of <Track Name>?", with five questions on 7-point slider scales: "How has your enjoyment of the music changed?" (Enjoyment; Enjoyed less - No change - Enjoyed more); "How has your emotional reaction to the music changed?" (Emotional reaction; Not at all - Very much); "Has your interpretation of the lyrics/ music changed?" (Image Recall; Not at all - Very much); 'Has the music video made the song more personally significant?' (Personal Significance; Not at all - Very much). (Interpretation; Not at all - Very much); "Did you remember images/ visual content from the video?" Responses to these MV-ESRs were used in the analysis, where enjoyment and emotional reaction were entered as dependent variables, and image recall, personal significance, and interpretation were entered as predictors. If participants had not seen the music video for either song, they were presented with an alternate screen, with three 7-point Likert-scale sliders relating to the music. This screen was added in order to keep the number of screens presented to participants consistent regardless of whether they had seen the music video or not. This was done in order to remove any incentive to indicate that they had not seen the music video even if they had in order to complete the number of ESRs necessary to be entered into the prize raffle.

2.4.4 Analyses

A total of 152 MV-ESRs were collected from participants. MV-ESRs consist entirely of experiences for songs that participants had seen the music video for. Two Linear Mixed Models (LMMS) were performed which incorporate the participants as random intercepts; this allows for individualised intercepts for each participant, thereby considering individual differences between participants.

Each predictor variable (Image recall, Personal Significance, and Interpretation) was entered into each model separately, as well as interactions between Image Recall and Personal Significance, and Image Recall and Interpretation. Simple effects analyses were performed in the event any significant moderations were observed in order to detect the slope of each predictor (Personal Significance and Interpretation) and its significance level at different levels of Image Recall.

3   SUMMARY OF STUDIES

The main findings from each of the four studies presented in this dissertation are summarised in this chapter. Studies I, II and III were the most exploratory; Studies I and II identified the key conceptual phenomena, while Study III explored the relationships between these concepts further. Study IV was hypothesis driven and used more sophisticated data collection and statistical analyses in order to provide insight into when, how and for whom MVs have the most salient influence on everyday music listening experiences. The findings from these studies are presented in greater detail in their original publications

3.1   Exploring music video experiences and their influence on music perception

The aim of the first study was to create a preliminary framework for understanding MV experiences. The data categorised according into four temporal levels: the first level, Intention, describes the reasons or contexts for watching MVs in the first place; the second level, Attention, includes the type of content which absorbs attention during the experience, the factors that distract them from enjoying their experience of the music, and the circumstances where they direct more attention to certain features of the music or video (i.e, one modality over another); the third level, Reaction, describes affective outcomes evoked during MV experiences, the contingent factors that influence these responses, and the mechanisms that make these reactions possible; the fourth level, Retention, describes the carry-over effects of MV experiences on future listening episodes. The discussion of these results are focused on how MV experiences can bridge the gap between theories concerning audio-visual cognition and everyday music listening experiences.

3.1.1 Results

Intention-level categories provided insight into the factors which motivated participants to watch MVs and were further categorised depending on whether these reasons were self-motivated (Internal) or motivated by peers or other outside influences (External). Internal goals included affect-regulation goals such as to relieve boredom, and to change or maintain their emotional state or energy levels. Cognitive goals reflect wanting to learn something about the music, whether it be the artist's personal interpretation of the song, as well as learning how to perform the music or dance choreography. External goals reflect the social function of music listening and includes sharing MVs with peers or watching MVs peers have shared with them. Media influences were also included as a separate sub-category within the External heading in order to reflect participants' experiences watching MVs which were trending in the media or recommended to them by YouTube. An additional category was included in order to cover reasons that reflect reasons such as new releases by a favourite artist or watching due to personal preferences for an artist (Preference-driven).

The actual experience of watching MVs was divided into two levels: Attention and Reaction. The purpose for this division was to differentiate between codes describing the content guiding attention during the experience, and the affective outcomes these experiences evoked for the participant. According to the data, participants were aware of how their attention was directed towards semantic features and structural features of the MV. Semantic features describe attending to the narrative content of the MV (labelled interpretation-focused) and visuals which draw attention to the music's emotional quality (labelled affect-focused). Structural features describe attending to how the music and video's structural components were aligned with or otherwise complementing each other, as well as focusing on performance gestures, whether it be the artist performing the song or the presence of dance and choreographed movement. Reactions were more difficult to categorise, since participants overwhelmingly stated that the reaction evoked by the experience depended on a number of factors (labelled contingencies in the study), such as their personal associations with the music, the content of the video, and the characteristics of the song. However, just under half of participants stated that MVs had the potential to elicit especially strong affective reactions, either in general or under particular conditions. These affective reactions could be of either positive or negative valence. Only four participants stated that MVs generally had no impact on their affective state.

The final category, Retention, is perhaps the most important in respect to the current study and this dissertation as a whole. The findings from this level highlight and synthesise current theories of audio-visual cognition and everyday music listening outcomes and shed light on how malleable and flexible music listening experiences are over time. Overall, 76% of participants experienced long-term effects on their perception of the music as a result of having seen the MV. Three categories were derived from this data: New Interpretation of Meaning (NIM), which reflects how MV content influenced participants'

perception of the music's meaning in subsequent listens; New Affect Perception (NAP), which reflects how participants perception of the music's emotional quality changed as a result of the MV clarifying its emotional tone, which in turn would influence emotional reactions; and Visual Mental Imagery (VMI), which reflects how images from the MV were automatically triggered in the mind's eye of the listener. These categories often overlapped with each other: for example, MVs with a story could provide new interpretation of the music, causing a recall of scenes from that story in subsequent listens, and/or causing them to have a substantially different emotional reaction compared to their listening experiences pre-MV exposure.

3.1.2 Conclusions

The study provided novel insight into modern MV listening experiences by providing new insight into why and how individuals engage with music in this type of format. The study found results similar to those of Sun and Lull (1986), including watching to see a visual interpretation of the music, for social learning purposes, and for affect regulation related reasons. The results from the study extend their findings by providing an updated model which accounts for sociocultural factors which did not exist at the time of Sun and Lull's original study, such as social media and streaming platforms like YouTube. Attention and Reaction-level categories highlight how CAM mechanisms function in MV contexts, which may be especially nuanced in cases where the listener has previously held associations with the music. Furthermore, the ways in which MVs affect the perception of the music's meaning are also in line with Cook's (1998) contest and complementation models of multimedia. MVs could provide a meaning for the music which was in contest with the participant's previous associations with and perceived meaning about the music. Although understanding the music's meaning was frequently stated as an important factor which motivated the individual to watch the MV, there were cases where participants stated they would explicitly avoid MVs for songs which held significant personal meanings for them in order to avoid this conflict. On the other hand, there were cases where the MV complemented the participants initial associations, which afforded more salient affective reactions in the future. Furthermore, Retention-level results support the findings from previous research which found that providing extramusical, contextual information about a piece of music can influence the emotions it evokes, and the mechanisms involved, especially the visual imagery mechanism (Vuoskoski & Eerola, 2015). This study's findings also support Schubert and colleagues (2014) spreading activation theory. This theory posits that aesthetic pleasure occurs by triggering a network of mental representations associated with the stimulus (ie., the music).

In conclusion, the results from this study support and extend the findings from previous topical areas related to music and audio-visual cognition research and synthesise them in the context of MV listening. The IARR model provides a new framework for exploring these experiences which considers the new platforms, devices and technologies providing accessibility and control over

when and how individuals engage with this music in this multimedia format. The study addresses the potential attentional and emotional mechanisms involved during these experiences and the influence this has on how the music is encoded in long-term memory. Overall, the IARR framework sheds light on how much extramusical information affects music perception, and the consequences this can have for the listener in certain situations. The study's findings are limited on their own however, since it relies on a small pool of participants and their subjective reports about previous experiences with MVs. Future research which looks at individual differences is required, as well as studies which explore the effects of context on listening outcomes However, the IARR framework provides a promising first step towards understanding MV and other musical multimedia by integrating existing knowledge across relevant disciplines and addressing the gaps within them. This research highlights the greater need for existing theoretical models to be revised and expanded upon when they do not adequately consider or reflect the variables influencing everyday listening experiences.

3.2   Study II: The function of music videos in everyday listening experiences

The second study analysed quantitative data collected from participants in Study I to explore the relationships between the IARR's qualitative categories and individual differences in music listening for emotional health purposes.

3.2.1   Results

T-test and u-test analyses revealed a significant difference in unhealthy and emotional stability scores and external goals. Participants who watched MVs for this reason had significantly lower unhealthy use scores (M = 12.9, SD = 4.13) compared to those who did not (M = 19.5, SD = 3.27): $t(32)$ = -4.48, $p < .001$. Participants with external goals also had higher scores in emotional stability (M = 4.88 , SD = 1.35) compared to those who did not (M = 3.4, SD = 1.45): $t(32)$ = 2.85, $p = .008$. No significant relationships were observed for internal goals.
At the experience level, significant differences in unhealthy scores were observed for both strong affect and unimpacted outcomes. Interestingly, unhealthy scores were significantly higher for individuals who experienced strong affect (M = 17.2, SD = 4.26) compared to those who did not (M = 12.8, SD = 4.58): $t(32)$ = 2.89, $p = .007$. Unhealthy scores were also significantly lower for individuals whose experiences were categorised as unimpacted (Md = 9.0), compared to those who were not (Md = 15.86): U = 12.5, $p = .004$. No relationships were found between either category and healthy or emotional stability scores.
    At the duration level, an independent samples t-test revealed that individuals who experienced long-term outcomes, regardless of whether they were perceived as positive or negative, had significantly higher unhealthy scores

(M = 16.0, SD = 4.8) compared to those who did not experience long-term outcomes (M = 12.2, SD = 4.27): t(31) = 2.09, p = .044. Participants who experienced long-term outcomes also had lower emotional stability scores (M = 4.04, SD = 1.43) compared to those who did not (M = 5.56, SD = 1.31): t(31) = -2.77, p = .009. A one-way ANOVA was performed to compare scale measure scores for individuals whose experiences had long-term positive effects, individuals whose experiences had long-term negative effects, and unaffected individuals. This analysis revealed no significant effects for healthy or unhealthy scores; however a significant difference was observed for outcome duration groups and emotional stability scores: F (2, 30) = 5.56, p = .009. Tukey post-hoc comparisons revealed that unaffected individuals had significantly higher emotional stability scores (M = 5.81, SD = 1.36) than both positive (M = 4.13, SD = 1.36) and negative (M = 3.67, SD = 1.13) outcome groups.

3.2.2 Conclusions

These results extend the findings of Study I by exploring the relationship between IARR categories and individual differences in use of music for emotional health and emotional stability. While these results are from a small participant sample, they provide a starting point for research to explore the relationship between different patterns of music video uses and their relationship to emotional health and well-being which could be particularly useful among youth for whom music is an important part of their everyday life. Of particular interest is the relationship between being Unaffected by music video content and emotional stability; this may indicate that emotionally stable individuals are able to exert some agency over which associations they attribute to music, which is why they are less likely to experience salient long-term changes in their perception of the music. However, one of the limitations these results present is that they do not account for different evaluations of different music videos: even individuals who disclose having experienced negative outcomes from having seen a music video in the past does not mean this is the case for every music video they see, even if they do generalise it to most music video experiences. Further research is required in order to examine the nature of this relationship in everyday listening experiences.

3.3 Study III: The characteristics of music video experiences and their influence on music perception

Studies I and II provided an exploratory overview of the functions of music video experiences, the influence of their extramusical content on music perception, and the influence of that exposure on subsequent listening experiences. However, there were several limitations. First, the data from these studies were extracted from a small participant pool. Second, while many participants provided descriptive accounts of their music video experiences and the perceptual

phenomena that occurred during initial exposure and in subsequent, audio-only listens, it was not possible to ascertain whether any relationship existed between their reasons for watching and these perceptual effects. Furthermore, since participants were asked to reflect about their experience with a music video they enjoyed, which they had already seen before, it was unclear whether their accounts of these perceptual effects were common to all music video experiences, or only in specific circumstances with songs that they already enjoyed. Therefore, the second study explored whether relationships exist between Intention, Attention and Reaction-level themes of the IARR framework and outcomes observed in Retention level themes.

The results from the study are presented in two parts: the first part outlines the findings from the PCAs conducted on both sets of survey questions (Experience components and Retention outcomes); the second part outlines the relationships observed in the linear regression analyses performed between Experience components and Retention outcomes.

3.3.1 Principal component analyses

3.3.1.1   Experience components.

The first PCA was performed on the items derived from the first three levels of the IARR. It was hypothesised that the extracted components would reflect the main reasons for engaging with music videos originally observed at the intention-level of the IARR in Study I, including watching for emotional reasons, for cognitive reasons related to understanding the meaning of the music, and social media engagement. Furthermore, components which reflect the narrative and performance elements of music video content were also anticipated. For items describing Retention outcomes, extracted components which highlight the influence of music video content on associations with the music, including the influences on the visual imagery mechanisms, were anticipated.

In the first PCA, seven Experience components were extracted. The first four components had the strongest loadings for items describing the experiential aspects, including the type of content being attended to and the nature of the affective reactions evoked, in addition to some intention-level items. The last three components primarily reflected intentions for watching and had comparatively weaker loadings for attention and reaction level themes compared to the first four components.

The first component, labelled Narrative Content, consisted of intention, attention and reaction-level themes which were primarily cognitive in nature, reflecting interest in and attention to the narrative components of the music video, as well as reactionary effects indicative of a deep emotional understanding of the music. Narrative Content items derived from the Intention level included cognitive reasons such as watching music videos that are like short films, in order to see the artist's interpretation, and to experience the music differently. The emotional reason of watching in order to relax was also loaded on this component, however it was comparatively weaker than the other intention level

items. Narrative Content items derived from the Attention and Reaction levels of the IARR also reflect this primarily cognitive function: Attention-level items had the highest loadings for this component, including attention to the storylines, interest in alternative interpretations of the music (as opposed to one that is similar to the listener's personal interpretation), and attending to how the music is represented in a visually interesting way. Reaction level items for this component included strong emotional reactions, a deeper understanding of the music's meaning on an emotional level and relating to the artist as a person. The second component, Visual Performance, had high loadings for items reflecting attention to performative gesture and dance choreography, and seeing the features of the music highlighted in visually creative ways. High loadings for items describing interest in and attention to musical performance features were also observed in both intention and attention-level items. The cognitive intention of watching in order to understand the artist's interpretation was also observed, however this is connected to interpretations of how it is performed as opposed interpretations related to narrative meaning. Interestingly, the Visual Performance component also reflected the use of music videos for modulating arousal levels: the emotional intention of watching in order to gain a boost of energy, and reaction items describing strong emotional reactions as well as feeling energised and motivated were observed at this level. The third and fourth components, labelled Meaning of Lyrics and Personal Interpretation respectively, both reflect experience characteristics with an emphasis on understanding the music's meaning, however, with nuanced differences. The Meaning of Lyrics component features strong loadings for intention items describing a desire to understand and reflect on the meaning of the music and its lyrics and has strong loadings on reaction-level items describing strong emotions responses, relating to the artist as a person, and a deeper emotional understanding of the music. The Personal Interpretation component primarily reflects items highlighting the importance of the music video featuring an interpretation similar to that of the viewer-listener's own personal interpretation, and reactions describing emotional contagion and relating to the artist as a person.

The last three components consist mainly of intention-level items. Emotional Use includes strong loadings for all emotional function intention items, especially "I watch music videos to distract myself when I am in a bad mood or emotional state." There were no experience-derived items with particularly strong loadings for this component, even items describing reaction level-phenomena. The last two components, Social Use and YouTube Recommender, consist primarily of intention-level items: Social Uses were characterised by items that describe the influence of other people and factors such as the media (e.g., "I watch music videos that are hyped in the media, have a reputation for something"). Relatively strong loadings for intention and experience-level items describing interest in and attention to dance choreography were observed for the Social Use component. Given that many music videos that have gone "viral" (that is, have been heavily circulated online) feature signature dance moves, this was not a surprise. Finally, the YouTube Recommender component highlights how

YouTube itself influences which music videos an individual is exposed to as a result of the algorithm recommending content based on their previous video streams.

3.3.1.2 Retention outcome components.

At the Retention-level, two components of interest were extracted: Visual Mental Imagery, and Personal Significance Increase. A third component, labelled "Unaffected", which reflects an absence of carry-over effects in subsequent listens, was also extracted, but was not explored to the same extent as the Visual Mental Imagery and Personal Significance Increase components. Both the Visual Mental Imagery and Personal Significance Increase components are made up of items derived from New Affect Perception (NAP), New Interpretation of Meaning (NIM) and Visual Mental Imagery (VMI) categories described in Study I.

The Visual Mental Imagery component reflects how visual information from the music video experience manifests itself as visual imagery in subsequent listens. The core items loading on this component are those that describe cognitive events such as picturing the performance gestures, dance, or specific scenes as mental images in subsequent listens. This outcome also has a strong loading on the item "It is hard not to think of the music video," which suggests that imagery may be triggered automatically in subsequent listens. Both the Visual Mental Imagery and Personal Significance Increase outcome include items derived from the NIM category described in Study I. The Personal Significance Increase outcome, however, also includes items that describe that song becoming more meaningful and personal to the viewer-listener. Personal Significance Increase also includes items derived from the NAP from having a different purpose of mood and emotion regulation, which do not load on to the Visual Mental Imagery component, and the affective quality of the song being perceived differently in subsequent listens.

3.3.2 Correlation and linear regression results

To address the second research question posed in this study, correlation and linear regression analyses were performed to explore the relationships between characteristic Experience components and Retention outcome components.

The analyses revealed positive correlations between all Experience components and both Visual Mental Imagery and Personal Significance Increase outcomes, however only a handful of these were moderately strong. The strongest correlations for Visual Mental Imagery were observed with Visual Performance $(r(153)= .53, p <.001)$ and Personal Interpretation $(r(153)= .45, p <.001)$ Experience components. Moderately strong, positive correlations were observed between Personal Significance Increase and the Visual Performance $(r(153)= .49, p <.001)$, Emotional Use $(r(153)= .47, p <.001)$, Meaning of Lyrics $(r(153)= .47, p <.001)$, and Narrative Content $(r(153)= .46, p <.001)$ Experience components. A moderately strong correlation was observed between both Visual Mental Imagery and Personal Significance Increase outcomes $(r(153)= .57, p <.001)$.

In order to explore the predictive relationship between Experience components and Retention outcomes, linear regression analyses were performed. Bonferroni corrections were applied to all regression results (p = .007). Although these regressions were exploratory, it was expected that relationships would be observed between experience characteristics emphasising narratives, performance gestures, and content outlining the meaning of the music; therefore, it was expected that the first five Experience components in particular would have strong positive relationships with both Visual Mental Imagery and Personal Significance Increase Retention outcomes.

Visual Mental Imagery was positively predicted by all seven Experience components to some extent, however, only two predictors explained over 20% of the variance. Visual Performance explained 27% of the variance (R^2= .27, $F(154)$ = 59.25, p <.001, 95% CI [0.39, 0.66]), and Personal Interpretation explained over 20% of the variance (R^2= .20, $F(154)$ = 40.16, p <.001, 95% CI [0.31, 0.60]). Three Experience components explained at least 20% of the variance for the Personal Significance Increase outcome. The strongest predictor was also Visual Performance, which explained 23% of the variance (R^2= .23, $F(154)$ = 47.28, p <.001, 95% CI [0.35, 0.36]). Meaning of Lyrics explained 21% of the variance (R^2= .21, $F(154)$ = 42.13, p <.001, 95% CI [0.32, 0.61]) and Narrative Content explained 20% of the variance (R^2= .20, $F(154)$ = 40.06, p <.001). These findings were in line with the expectation that experiences emphasising content with meaning and movement elements would be the strongest predictors of Retention outcomes.

3.3.3 Conclusions

The purpose of Study III was to explore the conceptual categories outlined in the IARR framework from Study I further to explore music video experience characteristics and their influence on subsequent listening outcomes. Items from the first PCA were derived from the thematic categories presented in the first three levels of the IARR (i.e., intention, attention and reaction) in order to establish these characteristics, which are referred to as Experience components. Items from the second PCA consisted of Retention-level categories of the IARR, and reflect the effects of music videos on subsequent listening outcomes.

The first four Experience components are an amalgam of intention, attention and reaction-level items. These components reflect patterns of uses that account for the initial reason for watching the music video, the content that absorbs attention, and typical affective reactions. Narrative Content, Meaning of Lyrics, and Personal Interpretation shed light on how semantic content can fulfil cognitive goals related to understanding the meaning of the music. There are nuanced differences between these experiences depending on factors like the importance of the viewer-listener's personal interpretation of the music, the presence of storylines, or whether they want to reflect on specific musical features like the lyrics. Visual Performance experiences fulfil cognitive goals pertaining to how the music is performed and emotional goals related to arousal modulation. The last three reflect emotion regulation-specific phenomena like watching in

order to distract from a negative mood, and social uses, depending on whether the experience is brought about as a result of social factors (such as watching with friends or because the video has a reputation in the media) or because of YouTube's recommender algorithm.

Visual Mental Imagery and Personal Significance Increase Retention outcomes reflect how content from the music video influences cognitive and emotional mechanisms in subsequent listens. Findings from this study were in line with the expectation that experiences emphasising content with meaning and movement elements would be the strongest predictors for Retention outcomes. The relationship between visual imagery and recalling music video content of performance gestures, whether it is of the artist performing or dancing, merits further exploration.

This effect may be the product of a joint memory code having been formed during the original music video experience. Mirror neuron systems (MNS) may also contribute to the joint encoding process in these cases: these neural systems activate the brain regions responsible for performing an action when observing another person performing the same action (Overy & Molnar-Szackacs, 2006). This may also explain the relationship between Visual Mental Imagery and other Experience components which highlight narratives and meaning-related content. The motor and emotional mechanisms activated during the experience result in a strong association being formed between the song and the video's content.

3.4   Study IV: The influence of music video content on listening outcomes, an experience sampling method study

The purpose of this study was to explore the influence of music video content in everyday personal music listening experiences using the experience sampling method (ESM). Data collected using the MuPsych smartphone application were analysed in order to investigate whether exposure to a song's music video affected subsequent enjoyment of and emotional reactions to the music. The study also explored whether visual imagery (Image Recall) of the music video content moderates the effects of other associative mechanisms (Interpretation and Personal Significance) on changes in emotional reactions and enjoyment. The analyses conducted for this study are from the piloting phase of the study; therefore, these findings reflect only a preliminary exploration of the first set of data collected with MuPsych. Two linear mixed models (LMMs) were performed with participants entered as random intercepts.

The study had two hypotheses:

1) As Image Recall increases, the effect strength of Personal Significance on musical enjoyment is expected to increase.

2) As Image Recall increases, the effect strength of both Interpretation and Personal Significance on emotional reactions is expected to increase.

The results presented in the manuscript for this study reflect the preliminary analyses on data collected from the first 15 participants recruited for the study. Data collection is still ongoing.

3.4.1 Results

The first hypothesis was not supported in the data. Image Recall and Personal Significance each have a significant effect on enjoyment individually, however, Image Recall did not moderate the strength of Personal Significance's influence on enjoyment.

The second hypothesis was supported in these analyses. Interpretation had a significant effect on emotional reaction (b=0.94, SE=0.25, p<.001), and these effects were significantly moderated by Image Recall (b=-.0.11, SE=0.05, p=0.28). As the amount of imagery from the music video increased, the strength of Interpretation on emotional reactions was weaker. The effect of Interpretation on emotional reactions was significant at all levels of Image Recall (p<.001), however this effect was weakest at higher levels (i.e., Mean +1SD; b= 0.31, SE= 0.08, p<.001) and strongest at lower levels of Image Recall (i.e.,Mean – 1SD; b= 0.68, SE = 0.14, p<.001).

While no significant effect was observed for Personal Significance on emotional reaction, an interaction effect was observed between this variable and Image Recall (b= 0.14, SE=0.05, p=.006). These results suggest that as the amount of visual content remembered from the music video increased, the effect strength of Personal Significance on changes in emotional reactions increased. This effect was only significant for higher levels of Personal Significance (i.e., (Mean +1SD; b= 0.33, SE= 0.09, p<.001).

3.4.2 Conclusions

The results from the pilot study presented in this paper provide insight into the influence of visual content from music videos on subsequent listening outcomes using ecologically valid data collected during personal listening experiences. While these results are preliminary, these analyses provide evidence which suggests that not only does music video content affect the visual imagery mechanism, the extent to which images are remembered influences the effects of other mechanisms (specifically those sensitive to semantic information) on emotional outcomes.

While the first hypothesis was not supported, these results only suggest that Image Recall does not moderate the effect that the other two variables have on enjoyment. They do not necessarily suggest that any of these three variables on their own do not have an influence on how music is enjoyed in subsequent listens. Further analysis is required to establish whether changes in perception of musical meaning or in how much personal significance is attributed to the music are

affected by music video content in general. In addition, it is possible that other variables, including individual differences or situational factors related to where, why and with whom the individual is listening may moderate the strength of these mechanisms on musical enjoyment. Results from this study suggest that visual imagery from music videos can influence changes in emotional reactions in subsequent listens by moderating the effects of Interpretation and Personal Significance mechanisms. Interpretation and Personal Significance both reflect how semantic information from music video experiences affects subsequent emotional reactions by influencing what kind of information is associated with the music. The difference between these two mechanisms lies in whether these associations reflect something appraised as being of high personal significance or value or is merely representational. Interestingly, as the amount of imagery recalled increased, the strength of influence of these representational interpretations on emotional reactions decreased. On the other hand, the influence of personal significance on emotional reactions increased. These findings reflect what Cook (1998) refers to as complementation: the change in the personal significance attributed to the content of the music video is complemented by the visual images conjured of this content in subsequent listens, creating an emotional experience.

While the findings from this study are only preliminary, there are still limitations to address. The study did not address the influence of listening context or individual differences: variables which may also act as moderators. Musical training may also influence the ways in which music video content affects subsequent outcomes. For example, studies have shown that highly musically trained individuals use visual imagery in order to modulate arousal during personal listening (Küssner & Eerola, 2019). It is possible that, for musicians, music videos which feature performance elements in particular may have a stronger influence on subsequent listens as a result of the imagery providing additional context about the emotional tone of the music. Performance gestures in particular may be more easily encoded in memory during the initial music video experience for musically trained individuals due as a result of their music education. While these variables were not considered in the present analyses, additional data collection and analyses are planned which take these variables into account. Musical training and individual trait variables, including empathy, personality, and listening style will be measured using scales such as the Healthy-Unhealthy Music Scale (Saariakallio et al., 2015), MUSEBAQ (Chin et al., 2018), and the Interpersonal Reactivity Index (IRI; Davis, 1980). Further investigation into whether musical background in particular influences the relationship between the prevalence of certain mechanisms and listening outcomes may have useful benefits in music education settings in particular.

4 GENERAL DISCUSSION

The studies presented in this thesis approach the topic of music videos and the perceptual phenomena they incite from different methodological angles in order to create a theoretical foundation for exploring this popular method of music engagement. It aimed to fill this gap in music psychology research by investigating the following research questions:

- What are the functions of music video experiences; why do individuals choose to watch them?
- How does extramusical content (i.e., visual information) influence the perception of and affective reactions to the music in music video contexts?
- How does exposure to music video content influence the perception of and affective reactions to the music subsequent, audio-only personal listening experiences?

Study I and Study II explored all three of these research questions using a qualitative approach in order to achieve a general understanding of music video experiences and the perceptual phenomena they elicit. Study I resulted in the IARR framework, which stands for Intention, Attention, Reaction, and Retention. This framework outlines the characteristic phenomena of music video experiences across three points in time: antecedents which lead to the experience, the experience itself, and effects on future listens. Intention-level phenomena include the reasons why individuals choose music videos over other music formats. The Attention- and Reaction-levels reflect the experience itself: Attention-level phenomena shed light on how attention towards audio and visual content influences music perception, while Reaction-level phenomena describe how these perceptual changes influence affective outcomes. Retention-level phenomena describe how mechanisms like visual imagery and personal associations may potentially change as a result of music video content.

Study II complements the findings of Study I by exploring which individual differences are associated with different categories of music video use and effects. Study III extends these findings by exploring the characteristic patterns of music

video use and the visual content associated with particular functions. Results from Study III shed light on how associative mechanisms related to personal associations and interpretations, and emotional mechanisms like mental imagery are influenced by the video's content. Study IV explores the influence of these associative and emotional mechanisms on the enjoyment of and emotional reactions to the music in subsequent listens.

Study III and Study IV address the second and third research questions using quantitative methodologies. Study III features a survey study with items derived from the qualitative categories in Study I in order to establish the characteristic patterns of music video use and their relationship to subsequent outcomes. Study IV focuses exclusively on the third research question; it uses experience sampling methodology and more sophisticated quantitative analysis procedures to investigate how music video content influences subsequent listening experiences. Findings for each research question, as well as limitations and directions for future research, are discussed.

4.1   RQ1: The functions of music video experiences

Why do people watch music videos? What type of experience does this multimedia format offer that cannot be fulfilled by audio-only listening experiences? Prior to this dissertation, Sun and Lull's (1986) investigation of adolescents' reasons for watching MTV was the only study concerned with the function of music video experiences. In Study I, the categories at the Intention-level of the IARR framework are equivalent to the types of functions these experiences fulfil. In Study II, these categories were explored in relation to individual differences in personality and music use for emotional health. In Study III, relationships between themes from these categories were explored further using quantitative analyses. These findings are discussed in light of both the functions music videos fulfil and the affordances that new technological developments (i.e., personal devices such as smartphones, and online platforms like YouTube) offer viewer-listeners seeking to gratify their psychological needs through music engagement.

The studies presented in this thesis provide insight into the cognitive, emotional and social functions music videos fulfil, including the strategies that facilitate these functions and the mechanisms they employ. Cognitive, emotional and social reasons for watching music videos were collapsed into two categories in the Intention level of the IARR in Study I: internal factors and external factors. The decision to amalgamate emotional and cognitive uses together was justified on the grounds that considerable overlap exists between these uses in respect to affect regulation and self-reflection (T. Schäfer et al., 2013). Therefore, reasons which highlight these functions were labelled together under the broader categorical label of internal factors. These reasons were considered self-motivated since the participant reports purposefully seeking out music videos in order to fulfil their specific psychological needs. This need may reflect emotion

regulation goals like boredom relief, or more cognitive- oriented goals, such as understanding the music's meaning, or seeing the music performed. Emotional reasons not only include regulation goals like changing or maintaining one's affective state, but also those characterised by wanting to experience the music differently. For example, in Study I, visual appreciation was categorised as an emotional reason since the visual component is perceived as having an additive effect which provides a different aesthetic experience of the music. On the other hand, external factors describe experiences which reflect, not just the social function of music, but the influence of outside sources on music video engagement behaviour. In Study I, these sources were labelled as peer influence or media influence. In the case of the former, the viewer-listener happens upon a music video as a result of other people: friends and peers. In the case of the latter, music video engagement occurs as a result of a music video's pop culture relevance or as a result of the viewer-listener's engagement with social media and streaming platforms like YouTube. The additional preference driven category reflects reasons for watching based on the individual's musical preferences, such as watching new releases by favourite artists or for songs they really enjoy.

Experience components outlined in Study III highlight the different patterns of music video engagement; the structure of these components is derived from survey items describing both the reasons for watching and the experience itself, including the content of the video associated with fulfilling that function (for the purpose of this section, only insights relative to RQ1 are discussed). These items were established based on the IARR themes outlined in Study I; the patterns which emerged highlight the relationships between specific goals, content types, and affective outcomes. These Experience components also highlight the many conceptual overlaps that exist between music's cognitive, emotional and social functions. For example, Narrative Content is characterised by wanting to see music videos that are like short films, or for the artist's interpretation of the music (both of which are considered cognitive functions). Visual Performance characteristics also include wanting to see the artist's interpretation, as well as for the purpose of regulating arousal or increasing motivation (which was considered an emotional function). Furthermore, the differences between the Personal Interpretation component, and Meaning of Lyrics and Narrative Content components highlight how the viewer-listener's current interpretation or personal associations of the music are reflected in their motivations for watching. This has valuable applications for future research interested in the influence of extramusical information on music perception by providing a foundation for understanding why such information would be sought in the first place and its relationship to an individual's psychological goals and needs. If the viewer-listener watches a music video in order to distract themselves from a bad mood, they may be more likely to watch music videos that fulfil certain criteria over others, such as a preference for familiar music videos over discovering new ones. This may be particularly useful for adolescent and young adult populations, since music has an important function at this stage of

life in respect to psychological development (Greenwood & Long, 2009; Laiho, 2004; Miranda et al., 2015).

Two intention-specific Experience components are highlighted in Study III: Emotional Use and Social Use. These components reflect broader patterns of engagement where the importance of the music video's content depends on additional context variables. The Emotional Use component consisted primarily of items detailing intentions with clear mood or emotion regulation goals, such as to distract from negative thoughts and feelings, to relax, or to feel more energised. The Social Use component highlights the external factors category outlined in Study I, such as watching because of peer or media influences. Watching in order to see dance choreography was associated with Social Use. This is perhaps not surprising, given the relationship between dance and popular music: many of the most viewed YouTube videos of all time feature iconic dance moves, such as PSY's "Gangnam Style" (see the Discussion Section in Study III). There was also one platform specific component, YouTube Recommender, which reflects music video experiences that come about, not necessarily as a part of music-specific media engagement but as part of engaging with YouTube in general. While this component was named after YouTube specifically (since the survey featured other questions which were specific to the participant's use of that platform), social media and music streaming platforms alike both generate recommendations to users based on their previously watched content. Users are recommended content based on what other users with similar tastes watch, or based on what is currently trending on the platform in general (Khan, 2017; Oh & Choeh, 2022). As a result of these recommenders and autoplay systems, individuals can end up watching music videos without seeking them out.

The importance of performance gestures, whether in the form of the artist performing or choreography, and its relationship to music video functions were highlighted. Seeing how the musician performs the music was identified in Study I as a factor motivating music video experiences. For individuals with musical training, music videos were an opportunity to gain information about how the music is produced. Similarly, individuals with dance training stated an interest in videos with choreography. Sun and Lull also reported dance as a reason for watching MTV, categorising it under the label emotional response. Individuals, regardless of whether they have any dance background, watch music videos with dance for the simple reason that dance is enjoyable to watch. This can be considered an emotional reason and a cognitive one, regardless of whether they are motivated by the desire to learn the choreography or not. However, as one participant pointed out in Study I, it was a result of their dance background that certain music videos (presumably ones which they did not actively seek out for their dance, however this is purely conjecture) stood out more to them than others. Dance can provide insight into the meaning of the music through the dancer or dancers' expressive movements to the music, and additional information in their facial expressions, clothing, and the environment where they are dancing (Van Dyck et al., 2017). This may also explain the relationship between cognitive reasons related to watching performance gestures and

emotional reasons related to arousal modulation, which was identified in the Visual Performance component established in Study III. Individuals may be aware of the additive effect that seeing dance choreography, or the artist perform the music has on their affective state and therefore seek out music videos which they know contain this type of content in order to have their psychological needs met.

The findings from these studies relative to RQ1 suggest that, while the motivating factors behind music video engagement previously identified in Sun and Lull's study have remained more or less consistent, the contexts where they are suitable for fulfilling these functions have expanded due to technological developments. This is supported by recent research concerned with music use and gratifications and their relationship, which suggests that the different formats are associated with different types of functionalities (Krause & Brown, 2021). While the aspect of control afforded to music listeners has generally had positive consequences in respect to using music to meet their psychological needs (Krause et al., 2015; Krause & Hargreaves, 2013; Krause & North, 2016; Skånland, 2013), the extent to which this holds true for music video experiences specifically needs to be explored further.

4.2 RQ2: How does content from music videos influence music perception and music evoked affect

The studies presented in this thesis apply insight from existing research across multiple related disciplines in an attempt to explain the cognitive and affective phenomena of music video experiences. These phenomena were explored using a combination of both qualitative and quantitative datasets. The findings complement previous research concerned with music and audio-visual perception and extend it to a modern and continuously evolving music format.

In Study I and II, this question is addressed within the Attention and Reaction-levels of the IARR framework. These levels, which reflect the experience itself, provide insight into content which absorbs attention and the affective reactions that occur as a result. The Attention and Reaction levels illustrate the bottom-up and top-down cognitive processes highlighted in CAM; the viewer-listener directs attention towards information in the video and generates inferences about the meaning of the music based on that information. These findings expand our understanding of CAM by highlighting how emotional responses to the music are potentially affected when the associations already attributed to the music by the viewer-listener are contested by content of the video. These findings provide a starting point for future research to explore the effects of visual information on music perception in personal music listening contexts. This is especially pertinent in today's listening culture, where 31 % of all online music streaming is done on a video streaming channel (IFPI, 2023).

Study I and II explored experience phenomena by analysing qualitative accounts of music video experiences. At the Attention-level of the IARR, two main categories of content were established: Semantic features and Structural features. Semantic features reflect visual content which contributes to the meaning of the music, either in the form of narrative stories, or by drawing attention to the music's affective qualities. Structural features highlight how attention is directed towards the perceived congruencies between the audio and the visual channels, whether in the form of performance gestures (i.e., dance and musician movements) or the ways in which the audio and visual are synchronised with each other. The ways in which these features influence the viewer-listener's reactions, however, depends on their appraisal of the video relative to their appraisal of the music. At the Reaction-level of the IARR, the data revealed three factors which affective responses were contingent on: musical factors (i.e., artist, familiarity, genre, emotional tone), visual factors (i.e., storylines, performance elements, quality of production) and personal (previous appraisal, existing interpretation of the music, current affective state). While nearly half of participants (47%) reported a salient emotional response (categorised as Strong Affect in Study I and II), it is clear that affective reactions to music videos do not exist in a vacuum: there are a number of variables that need to be accounted for, and participants are aware of the interactions between these contingencies and their influence on both their subjective appraisal of and affective responses to the music video's content. Previous research has shown that music can accentuate moments in the film by providing emotional information that increases the emotional saliency of the film and increases positive sentiments for the film's story (Millet et al., 2021). Findings from these studies suggest that the reverse is also true: the video can potentially increase positive sentiments about the music. However, this depends on a number of factors which were not explored in this thesis. This includes contingent factors related to the individual, the music, and the video, as well as contextual (i.e., situational) variables.

The type of content influencing attention and reactions during music video experiences and its relationship to intention themes was explored in the greatest depth in Study III. The first three Experience components outlined in this study (Narrative Content, Visual Performance, and Meaning of Lyrics) outline the functionality of specific visual content in relation to certain listening goals. What is especially interesting about these three components, however, is how each embodies a specific type of content, and how that content functions as a mechanism for expressing musical meaning. Narrative Content demonstrates the relationship between watching music videos to see the artist's interpretation of the music and for their narrative components, and experience items describing attention towards visually interesting content towards, visual elements which creatively highlight the features of the music (for example, with special effects), storylines, and ambiguous endings. The video reflects a visual interpretation of the music's events in a narrative form. These narratives influence music perception by giving context about the music's meaning, evoking feelings of

relatedness with the artist, and potentially strong emotional responses. The Visual Performance component demonstrates how gestural content in the form of singing, instrument playing, and dance choreography influence the perception of the music's emotional meaning and on the individual's affective state, especially their energy levels/arousal. The third, Meaning of Lyrics, is similar to Narrative Content, however more emphasis is placed on using the video as a mechanism to understand the meaning behind the music's lyrics and to see how this meaning is expressed in the artist's performance. Interestingly, all three of these components are associated with strong emotional experiences, relating to the artist (i.e., parasocial bonding), and a deeper understanding of the music's emotional meaning.

The findings from Study III complements existing research exploring the relationship between strategies of listening and affect regulation mechanisms. For example, Visual Performance experiences exemplify the same combination of revival strategies (watching music videos to get a boost of energy) and repairing and pleasure mechanisms (Baltazar & Saarikallio, 2019). While the neurocognitive processes involved in music video experiences are beyond the scope of this thesis, the aspect of seeing the movements of the artist or another human (i.e., dancer) and the influence this has on emotional outcomes could provide support for the role mirror neuron systems (MNS), regions of brain activity involved in both the perception and production of motor functions (see Bonini et al., 2022), in music evoked emotional outcomes. The importance of MNS in musical experiences is present, even in the absence of visual stimuli, and has been linked to cognitive capacities such as empathy, theory of mind, and self-other discrimination. Music is, at least at a subconscious level, associated with the movement required to produce it, whether it be the arm movements required for drumming or the larynx for singing (Molnar-Szakacs & Overy, 2006). Watching these performances may activate mechanisms such as emotional contagion or rhythmic entrainment as a result of MNS (the relationship between music videos depicting human gestures and their influence on subsequent listens will be discussed in greater detail below). In sum, Visual Performance experience components may contribute to music perception by activating the different neural and cognitive mechanisms involved in both the perception of emotions and the motor activity associated with producing those emotions.

The findings from the first and third studies may be of particular interest to researchers interested in embodied cognition and the psychological mechanisms behind affective reactions to dance (and human movement in general). Further music video research may provide insight into what makes dance such an effective tool for social bonding and displaying emotion. These findings may also be relevant to researchers interested in the motivations and functions for watching dance over other aesthetic experiences. This research could shed light on the possible relationships between the perception of music and dance, and the role of different mechanisms (for example, entrainment, contagion, and MNS), music listening behaviours and traits such as personality and empathy. Emotional outcomes in response to watching dance may be a result of kinesthetic

empathy, where the viewer experiences a physical reaction, such as relaxing or tightening their muscles, in response to the emotion communicated by the dancer (Reason & Reynolds, 2010). This type of kinesthetic empathy may also occur in response to emotional gestures associated with singing or playing an instrument. Music video research provides a unique opportunity to apply this conceptual insight to a popular method of music engagement which has yet to be fully understood from a music psychology and cognition perspective.

4.3   RQ3: Influence on subsequent listens

Music videos can have a profound effect on how an individual perceives and responds to the music, even in subsequent listens. The results from all four studies provide evidence to support this. Three conceptual categories of effects from music videos on subsequent music perception are outlined: New Interpretation, Personal Significance, and Visual Mental Imagery. Although conceptually different, these effects may occur independently or concurrently with each other depending on the individual, the situation, and the features of the music and video. These findings are discussed in respect to how these outcomes reflect our understanding of music video experiences in everyday life, as well as the underlying perceptual mechanisms responsible for these carry-over effects.

The broadest category of effects is New Interpretation; this was highlighted in the Retention categories from Study I and II, Retention outcomes from Study III (excluding the Unaffected outcome), and the Interpretation variable measured in Study IV. This category reflects how semantic information in the form of narrative structures and emotional imagery contribute to the subsequent listening experiences by influencing how the meaning of the music is perceived. New Interpretation may contribute to changes in Personal Significance as well in cases where the viewer-listener has a strong personal connection with the themes, characters, objects or events depicted in the video. Changes in Personal Significance reflect how semantic information from the music video influenced affective reactions by virtue of this information causing a change in their personal relationship or associations with the music. Therefore, Personal Significance may not reflect a change in semantic interpretation, but the reinforcement or abolishment of an existing one. The constructs associated with this effect are outlined in the Personal Significance Increase outcome (see Study III).

The last category of effects, Visual Mental Imagery, reflects the influence of music video content on the type of mental imagery conjured during subsequent listening experiences. These results suggest that, when music and video are perceived as being both structurally and semantically congruent, an integrated memory code is formed (Boltz et al., 1991; Boltz et al., 2009). This may explain how images from the music video are retrieved from long term memory and experienced as mental images in subsequent listens, almost automatically. Visual Mental Imagery may reflect a type of conditioning at the level of higher-order

cognition, where the presence of one modality (the music) facilitates the recall of the other (the video). This may explain why some participants remembered specific images at particular moments in the video or found it difficult not to think of the music video in subsequent listens (see Retention category quote examples in Study I).

In Study IV, Visual Mental Imagery from music videos was explored, not only as an effect experienced in subsequent listens, but as a potential moderating variable capable of influencing emotional outcomes. Visual Mental Imagery, labelled Image Recall in this study, had a different moderating effect on emotional reactions when Interpretation and Personal Significance were considered. Interestingly, Interpretation had a significant influence on emotional reactions, however, the strength of this effect was weaker in cases with higher levels of music video-related mental imagery (labelled as Image Recall in Study IV). On the other hand, no significant effects on emotional reaction were observed for Personal Significance, however, an interaction effect was observed between this variable and Image Recall. The influence of Personal Significance on emotional outcomes was stronger as Image Recall increased. This may be a result of the music video affirming or creating a strong personal association which intertwines their own autobiographical narratives with that of the music, resulting in what Green (2016) refers to as peak music experiences. The increased emotional impact of the music reflects how the construction of musical meaning may evolve as new inspiration and motivation are procured.

These findings complement existing research on audio-visual interactions and their influence on music perception. In Study I, participants were asked whether the music video had a long-term influence on their perception of the music. Since participants were asked to watch a music video which they had already seen before and enjoyed, the majority of responses were biased towards outcomes for music videos they had enjoyed. However, several participants still provided an example of a music video which had a negative influence on their subsequent listening experiences, and why that was the case. The data revealed that whether a music video context was perceived as having a positive or negative influence on subsequent listening outcomes could be determined by whether the video complemented the meaning understood within the music or contested it. These findings were described from the perspective of Cook's (1998) complementation and contest models of multimedia: when the audio and visual channels were perceived as having an additive effect on the experience of the music, the complementation model applies. On the other hand, when a new interpretation was generated as a result of the semantic characteristics of both the audio and visual information imposing their own meaning on each other, the contest model applies. However, the way Cook's content model applies to these situations is determined by the original meaning assumed to be present in the music from the perspective of the listener contends with the meaning put forth in the video. If the original meaning is appraised as being more personally significant and important to the viewer-listener compared to what is presented in the video, then the meaning generated tends to be evaluated more poorly. This

effect can potentially extend to future listens, creating a more negative appraisal of the music, making it less impactful than before and potentially ruining its ability to regulate affective states compared to how it did before. If the video provides a source of meaning which is perceived as having an additive effect on the music's emotional qualities, this appraisal extends to subsequent listens in a positive way.

This research provides a starting point for understanding the effects of music video content on subsequent listening outcomes. Music videos give viewer-listeners an engaging, multisensory experience of music capable of absorbing focus and eliciting strong affective reactions. The findings from these studies shed light on how extramusical information from a music video's emotional content contributes to music perception by providing additional context about the meaning of the music.

4.4   Limitations and directions for future research

In the literature section of this dissertation, a definition was provided for music video experiences: "A music video experience involves a viewer-listener who, either deliberately or circumstantially, finds themselves engaged with and actively attending to both the audio and visual content of a music video over a period of time." While I have provided a definition for the components of the music video, namely "emotional content", no definition for a music video was provided. While this may seem like an oversight, it was intentional. The definition of what a music video is has changed significantly over the years since work on this dissertation began. Music videos were supposed to be defined as a short film accompanying a song that was released for the purpose of promoting a song, artist or album. Music videos which fit this description that were not considered at the start of this dissertation were videos from concerts, behind-the-scenes footage, and fanmade videos. This was addressed in the design of studies III and IV, where a more specific definition of a music video was provided.

Once the COVID-19 pandemic occurred, however, the problem of how to define a music video became even more pronounced. The pandemic created a surge in online content that could be defined as music videos, including new songs about the pandemic and lockdown, parodies of existing songs, dances, livestreams, and virtual concerts (see Hansen et al., 2021 for a database of COVID19 videos). In addition, the new social media app, TikTok, has continuously grown in popularity over the last few years - a relatively short period of time which overlaps with the period where this dissertation was being written. Music engagement on social media platforms was considered a method of music engagement according to the 2022 and 2023 IFPI reports. While it currently only makes up a small percentage of music engagement on a global scale, the way social media platforms like TikTok have boomed in popularity over such a short period of time need to at least be on the radar of music

psychology researchers who are interested in everyday music engagement behaviours.

The functions of music (and music videos) as they are understood by music psychologists (i.e., the Big Three proposed by Schäfer et al., 2013) are becoming increasingly difficult to disentangle, especially given the importance of social media platforms and their influence on music engagement behaviours (Vizcaíno-Verdú et al., 2023). The characterisation of internal versus external suggests the existence of an overlap between music's emotional and cognitive functions, and the lack of such an overlap between these functions and music's social functions . While this initial decision provided an adequate conceptualisation of the themes observed in the data, in hindsight, it may be overly reductive. While clear conceptual boundaries are certainly necessary, it may be more helpful in the case of music video functions to explore when these overlaps appear in the context of specific experiences.

These studies did not account for differences between first time views and rewatches of the same music video. It is possible that meaning-related motivations are a stronger influencing factor for first time views; how the meaning is represented in the video will then determine the function of subsequent experiences, depending on how it is appraised by the viewer-listener. For example, if the music video adds emotional depth and meaning to the song, the viewer-listener may choose to watch it to regulate affective states in the future or share it with friends during social activities. While some qualitative accounts from Study I suggest that one view is enough to completely change their experience of the music, how this influences subsequent engagement with the same music video was not explored. While the results suggest that individuals do rewatch music videos (see Study II: Supplementary Data), differences in first time views compared to rewatches were not explored. Future research should investigate these differences, as well as how individual differences and contextual factors motivate individuals to rewatch music videos. The reasons why a person chooses to rewatch a particular music video is likely to depend significantly on the outcome of the initial watching experience, including the cognitive processes and affective reactions that occurred during the original experience.

Future research should consider the possibility that certain types of content, such as storylines or performance gestures, are associated with specific functions. In addition, future research should also consider the relationships between individual differences and music video engagement behaviours. Existing research on the use of music in everyday life has shown how individual differences are related to music listening engagement patterns, including differences in age (Chamorro-Premuzic et al., 2012; Groarke & Hogan, 2016; Miranda & Claes, 2008) gender, personality (Chamorro-Premuzic & Furnham, 2007) and other traits such as empathy (Vuoskoski et al., 2012). While some individual measures were used in Study II, the studies in this thesis ultimately do not take individual differences into account. For example, individuals who are high in trait empathy and are highly engaged with music may seek out music

videos they are familiar with when implementing affect regulation strategies like strong sensations, revival or entertainment.

Initial mood and psychological health are also important factors which need to be considered when considering the types of functions music video experiences fulfil for certain individuals. For example, if they are in a positive affective state, they may be more likely to watch music videos that have been recommended by the platform compared to if they are in a negative mood. In the later circumstance, they may be more likely to select a music video for a familiar song in order to distract themselves from or cope with their negative affect (Randall & Rickard, 2017). Future research should consider the relationships between individual differences, current mood, contextual variables with different patterns of music video use, including the functions they fulfil, and the specific content associated with gratifying that function. Larger datasets are necessary which include not only self-reports, but more detailed qualitative perspectives as well. A larger-scale multi-method approach which incorporates self-reports, physiological data, and controlled experiments could provide more insight into the conditions under which music videos have the most salient effects and for whom. On the other hand, the same technological developments which have made music videos more accessible provide the opportunity to implement these types of designs with the potential to be less intrusive for the participant.

While the findings from Study IV do not take individual differences into account, the study was set up in a way that these variables could be addressed in the future. Although they were not accounted for in this study as it is presented in this thesis, the ESM study (which is still ongoing) also has a survey component; participants can fill out surveys about their weekly music and music video consumption habits, demographic information, and psychometric scales measuring personality, empathy, and other traits with potential relationships to music video experience functions and effects. Furthermore, MuPsych also collects data about the music they were listening to, allowing for further investigation into the specific music videos which participants reported as having had an effect on their subsequent listening experiences. Although these analyses could not be performed as part of this dissertation, a more detailed analysis on the features of the music video may be performed in the future.

Due to the exploratory nature of this thesis, these findings are only generalisations of music video experiences. The categories and principal components outlined in these studies are based on retrospective accounts and self-reports of experiences - often positive ones - but do not directly measure any of the perceptual or affective phenomena that occur during these experiences due to methodological constraints. More research is necessary in order to investigate the scope of these effects. Despite this limitation, however, this type of research is necessary when investigating underexplored phenomena since it provides a foundation for future research to consider for future study designs. Future research investigating audio-visual perception should consider the use of experimental studies, where the music videos features can be controlled and

measured prior to exposing them to participants. An experimental design which combines self-reports and physiological data (for example, eye-tracking and galvanic skin response) could provide more concrete insight into the perceptual phenomena evoked during these experiences. Eye-tracking studies may be especially useful, since eye movements and pupillary responses provide a physiological measure of attention and emotional response (Gülçay & Cangöz, 2016; Mera & Stumpf, 2014). It may also be beneficial to conduct the same experiment using both self-selected and experimenter-selected music videos, since individuals tend to have stronger affective responses to preferred, familiar music than unfamiliar music (Szpunar et al., 2004).

5 CONCLUSIONS

The purpose of this thesis was to explore music video experiences and provide a foundation for future research to explore the topic. The need for a better empirical understanding of how visual information influences music perception in personal music contexts such as music video experiences has only become more pertinent since work on this thesis began in 2019. The findings from this dissertation provide a starting point for future research to explore this topic by providing insight into the functions these experiences fulfil, and their influence on music perception – both during the initial music video experience and in future, audio-only listening experiences.

These studies draw on existing research insight from multiple disciplines and theoretical perspectives. The IARR framework developed in Study I was explored and refined throughout the rest of the studies using a combination of qualitative and quantitative self-report datasets. The IARR provides a conceptual illustration of the different functions of these experiences (Intention), the cognitive (Attention) and affective (Reaction) phenomena that occur during the experience, and the influence of their content on subsequent listens (Retention). Studies I and II were exploratory, while Studies III and IV were hypothesis driven.

The findings from this thesis complement existing frameworks and theoretical models aimed at explaining musical experiences or their cognitive and affective components. The IARR applies insight from Juslin's BRECVEMA mechanisms (Juslin, 2013; Juslin & Västfjäll, 2008), Hargreaves' (2012) reciprocal feedback model, Schubert's (2014) spreading activation, models of musical multimedia such as CAM (Cohen, 2013; 2001; Marshall & Cohen, 1988) and Cook's (1998) models of multimedia, and research which investigates the perceptual effects of audio-visual stimuli. Music video experiences and the effect they have on viewer-listener's depend on interactions between the situation where the listening event takes place, the content of the music video, and the individual. These interactions determine the influence of variables related to the music and the video's perceived structural and semantic congruencies, the emotional salience of the experience, and the extent to which information is stored in memory and associated with the music in subsequent listens.

Music videos are an artefact of today's music listening culture and have been for over four decades. They reflect the symbiotic relationship that has always existed between these two modalities and continues to evolve as new forms of musical multimedia emerge. While these results are exciting and informative, there is still more work to be done.

SUMMARY IN FINNISH

Musiikkivideot ovat olleet osa musiikin kuuntelua ja populaarikulttuuria jo neljänkymmenen vuoden ajan. Nykyään musiikkivideot ovat helposti saatavilla YouTuben kaltaisilla alustoilla, mikä mahdollistaa oman lempimusiikin audiovisuaalisen kokemisen käytännössä missä ja milloin tahansa. Suosiostaan huolimatta, musiikkivideoiden käyttöä musiikinkuuntelun muotona tai niiden vaikutusta musiikin havaitsemiseen ei juuri ole tarkasteltu musiikinpsykologisessa tutkimuksessa. Vastatakseen tähän vajeeseen, tämä väitöskirja tutkii musiikkivideoita ilmiönä, selvittäen seuraavia tutkimuskysymyksiä: 1) Mitä merkityksiä/funktioita musiikkivideoilla on? 2) Millä tavoin ulkomusiikillinen sisältö vaikuttaa musiikin havaitsemiseen ja musiikin herättämiin tunteisiin musiikkivideoiden yhteydessä? Ja 3) Kuinka musiikkivideoiden sisällölle altistuminen vaikuttaa musiikin havaitsemiseen ja musiikin herättämiin kokemuksiin myöhemmillä kuuntelukerroilla? Kysymyksiin vastaamiseksi suunniteltiin ja toteutettiin neljä, eri menetelmällisiä lähestymistapoja hyödyntävää, tutkimusta. Tutkimusten laadinnassa hyödynnettiin aiempaa tutkimustietoa musiikkipsykologian eri aloilta, mukaan lukien elokuvamusiikin funktiot, musiikin käyttö tunnesäätelyssä, musiikin herättämät tunteet, musiikki ja muisti, sekä audiovisuaalinen havaintokokemus. Tutkimus I oli laadullinen selvitys musiikkivideoihin liittyvistä kokemuksista. Sen tuloksena syntyi Tavoite, Huomio, Kokemus, Seuraus – viitekehys (Intention, Attention, Reaction and Retention framework, IARR), joka toimi perustana muille osatutkimuksille. Tutkimus II tarkasteli ensimmäisessä tutkimuksessa tunnistettujen teemojen välisiä yhteyksiä ja yksilöiden välisiä eroja persoonallisuudessa ja musiikin käytössä tunne-elämän hyvinvoinnin tukena. Tutkimus III oli verkkokysely, jonka väittämät pohjautuivat ensimmäisessä tutkimuksessa havaittuihin sisällöllisiin kategorioihin. Tutkimus IV hyödynsi kokemuksenkeruumenetelmää (experience sampling method, ESM) tarjotakseen tietoa musiikkivideoiden ulkomusiikillisen sisällön vaikutuksista arjen kuuntelutilanteissa. Näiden tutkimusten tulokset tukevat kiistatta sitä havaintoa, että musiikkivideot eivät vain palvele monia samoja kognitiivisia, tunteisiin liittyviä ja sosiaalisia funktioita kuin musiikin kuuntelu, vaan kuulijat päämäärätietoisesti etsivät niitä niiden ulkomusiikillisen sisällön vuoksi. Tämä informaatio antaa musiikille lisää kontekstia ja sisältöä, joka puolestaan vaikuttaa myöhempään musiikin havaitsemiseen. Tutkimustulokset viittaavat siihen, että musiikkivideoiden ulkomusiikillinen sisältö vaikuttaa siihen, millaiset tunnemekanismit aktivoituvat kokemuksen aikana, erityisesti visuaalinen muisti, perustuen mielleyhtymiin jotka syntyvät ulkomusiikillisen sisällön pohjalta. Tutkimukset tarjoavat syvällisen katsauksen musiikkivideokokemuksiin ja niihin mahdollisiin havaintovaikutuksiin, joita ulkomusiikilliseen informaatioon liitetyt mielleyhtymät aiheuttavat. Tulokset ovat erityisen tärkeitä nykyisen kuuntelukulttuurin valossa, jossa musiikin striimauksen audiovisuaaliset käyttöliittymät ovat musiikinkuuntelun valtavirtaa.

REFERENCES

Arnett, J. J. (1995). Adolescents' Uses of Media for Self-Socialization. In *Journal of Youth and Adolescence* (Vol. 24, Issue 5).

Baltazar, M., & Saarikallio, S. (2019). Strategies and mechanisms in musical affect self-regulation: A new model. *Musicae Scientiae*, 23(2), 177–195. https://doi.org/10.1177/1029864917715061

Barrett, F. S., Grimm, K. J., Robins, R. W., Wildschut, T., Sedikides, C., & Janata, P. (2010). Music-Evoked Nostalgia: Affect, Memory, and Personality. *Emotion*, 10(3), 390–403. https://doi.org/10.1037/a0019006

Belfi, A. M., Karlan, B., & Tranel, D. (2016). Music evokes vivid autobiographical memories. *Memory*, 24(7), 979–989. https://doi.org/10.1080/09658211.2015.1061012

Boltz, M. (2004). The cognitive processing of film and musical soundtracks. *Memory & Cognition, 32* (7). 1194-1205

Boltz, M. G., Ebendorf, B., & Field, B. (2009). Audiovisual interactions: The impact of visual information on music perception and memory. *Music Perception*, 27(1), 43–59. https://doi.org/10.1525/mp.2009.27.1.43

Boltz, M., Schulkind, M., Kantra, S., & Bitnet, M.-B. (1991). Effects of background music on the remembering of filmed events. *Memory & Cognition*, 19(6), 593-606.

Bonini, L., Rotunno, C., Arcuri, E., & Gallese, V. (2022). Mirror neurons 30 years later: implications and applications. *Trends in Cognitive Sciences, 26*(9), 767-781.

Carlson, E., Wilson, J., Baltazar, M., Duman, D., Peltola, H. R., Toiviainen, P., & Saarikallio, S. (2021). The Role of Music in Everyday Life During the First Wave of the Coronavirus Pandemic: A Mixed-Methods Exploratory Study. *Frontiers in Psychology*, 12. https://doi.org/10.3389/fpsyg.2021.647756

Cespedes-Guevara, J., & Eerola, T. (2018). Music communicates affects, not basic emotions — A constructionist account of attribution of emotional meanings to music. *Frontiers in Psychology*, 9(FEB). https://doi.org/10.3389/fpsyg.2018.00215

Chamorro-Premuzic, T., & Furnham, A. (2007). Personality and music: Can traits explain how people use music in everyday life? *British Journal of Psychology*, 98(2), 175–185. https://doi.org/10.1348/000712606X111177

Chamorro-Premuzic, T., Swami, V., & Cermakova, B. (2012). Individual differences in music consumption are predicted by uses of music and age rather than emotional intelligence, neuroticism, extraversion or openness. *Psychology of Music*, 40(3), 285–300. https://doi.org/10.1177/0305735610381591

Cohen, A. (2001). Music as a source of emotion in film. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 249–272). Oxford University Press.

Collinson, L., Judge, L., Stanley, J., & Wilson, N. (2014). *Portrayal of Violence, Weapons, Antisocial Behaviour and Alcohol: Study of Televised Music Videos in New Zealand*.

Commodari, E., & Sole, J. (2020). Music education in junior high school: Perception of emotions conveyed by music and mental imagery in students who attend the standard or musical curriculum. *Psychology of Music, 48*(6), 824–835. https://doi.org/10.1177/0305735619832962

Cook, N. (1998). *Analysing Musical Multimedia*. Oxford University Press.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10,* 85.

Denora, T., Webster, J., Zolberg, V., Balfe, J., Macdonagh, J., Stone, T., & Scott, S. (1999). *Music as a technology of the self "Human-music interaction-music's 'effects' on feeling, embodi-ment and temporality*. www.elsevier.nl/locate/poetic

Edmond, M. (2014). Here We Go Again: Music Videos after YouTube. *Television & New Media, 15*(4), 305–320. https://doi.org/10.1177/1527476412465901

Greasley, A. E., & Lamont, A. (2011). Exploring engagement with music in everyday life using experience sampling methodology. *Musicae Scientiae, 15*(1), 45–71. https://doi.org/10.1177/1029864910393417

Green, B. (2016). 'I Always Remember That Moment': Peak Music Experiences as Epiphanies. *Sociology, 50*(2), 333–348. https://doi.org/10.1177/0038038514565835

Greenwood, D. N., & Long, C. R. (2009). Mood specific media use and emotion regulation: Patterns and individual differences. *Personality and Individual Differences, 46*(5–6), 616–621. https://doi.org/10.1016/j.paid.2009.01.002

Groarke, J. M., & Hogan, M. J. (2016). Enhancing wellbeing: An emerging model of the adaptive functions of music listening. *Psychology of Music, 44*(4), 769–791. https://doi.org/10.1177/0305735615591844

Gülçay, Ç., & Cangöz, B. (2016). Effects of emotion and perspective on remembering events: An eye-tracking study. *Journal of Eye Movement Research, 9*(2). https://doi.org/10.16910/jemr.9.2.4

Hansen, N. C. (2021). A Crowd-Sourced Database of Coronamusic: Documenting Online Making and Sharing of Music During the COVID-19 Pandemic. *Frontiers in Psychology, 12*.

Hargreaves, D. J. (2012). Musical imagination: Perception and production, beauty and creativity. *Psychology of Music, 40*(5), 539–557. https://doi.org/10.1177/0305735612444893

Hargreaves, D. J., & North, A. C. (1999). *The Functions of Music in Everyday Life: Redefining the Social in Music Psycholog*

IFPI. (2018). Music Consumer Insight Report. International Federation of the Phonographic Industry.

IFPI. (2019). Music Consumer Insight Report. International Federation of the Phonographic Industry.

IFPI. (2022). Engaging with Music. International Federation of the Phonographic Industry. https://www.ifpi.org/ifpi-releases-engaging-with-music-2022-report/

IFPI. (2023). Engaging with Music. International Federation of the Phonographic Industry. https://www.ifpi.org/resources/

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Juslin, P. N., Harmat, L., & Eerola, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, *42*(4), 599–623. https://doi.org/10.1177/0305735613484548

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Juslin, P. N., & Laukka, P. (2004). Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. *Journal of New Music Research*, *33*(3), 217–238. https://doi.org/10.1080/0929821042000317813

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5). https://doi.org/10.1017/S0140525X08005293

Katz, E., Haas, H., & Gurevitch, M. (1973). On the Use of the Mass Media for Important Things. *American Sociological Review*, *38*(2), 164–181.

Kendall, R. A. (2005). Music and Video Iconicity: Theory and Experimental Design. *Journal of PHYSIOLOGICAL ANTHROPOLOGY and Applied Human Science*, *24*(1), 143–149. https://doi.org/10.2114/jpa.24.143

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex*, *105*, 53–60. https://doi.org/10.1016/j.cortex.2017.10.012

Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior*, *66*, 236–247. https://doi.org/10.1016/j.chb.2016.09.024

Kim, K.-H., & Iwamiya, S.-I. (2008). Formal Congruency between Telop Patterns and Sound Effects. *Music Perception*, *25*(5), 429–448. https://doi.org/10.1525/mp.2008.25.5.429

Kistler, M., Rodgers, K. B., Power, T., Austin, E. W., & Hill, L. G. (2010). Adolescents and music media: Toward an involvement-mediational model of consumption and self-concept. *Journal of Research on Adolescence*, *20*(3), 616–630. https://doi.org/10.1111/j.1532-7795.2010.00651.x

Kosslyn, S., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature*, *2*, 635–642.

Krause, A. E., & Brown, S. C. (2021). A uses and gratifications approach to considering the music formats that people use most often. *Psychology of Music*, *49*(3), 547–566. https://doi.org/10.1177/0305735619880608

Krause, A. E., & Hargreaves, D. J. (2013). MyTunes: Digital music library users and their self-images. *Psychology of Music*, *41*(5), 531–544. https://doi.org/10.1177/0305735612440612

Krause, A. E., & North, A. C. (2016). Music listening in everyday life: Devices, selection methods, and digital technology. *Psychology of Music*, *44*(1), 129–147. https://doi.org/10.1177/0305735614559065

Krause, A. E., North, A. C., & Hewitt, L. Y. (2015). Music-listening in everyday life: Devices and choice. *Psychology of Music*, *43*(2), 155–170. https://doi.org/10.1177/0305735613496860

Küssner, M. B., & Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 90–99. https://doi.org/10.1037/pmu0000238

Laiho, S. (2004). The psychological functions of music in adolescence. *Nordic Journal of Music Therapy*, *13*(1), 47–63. https://doi.org/10.1080/08098130409478097

Lamont, A. (2011). University students' strong experiences of music: Pleasure, engagement, and meaning. *Musicae Scientiae*, *15*(2), 229–249. https://doi.org/10.1177/1029864911403368

Larson, R. (1995). Secrets in the bedroom: Adolescents' private use of media. In *Journal of Youth and Adolescence* (Vol. 24, p. 535).

Lonsdale, A. J., & North, A. C. (2011). Why do we listen to music? A uses and gratifications analysis. *British Journal of Psychology*, *102*(1), 108–134. https://doi.org/10.1348/000712610X506831

Marshall, S. K., & Cohen, A. J. (1988). Effects of Musical Soundtracks on Attitudes toward Animated Geometric Figures. *Music Perception*, *6*(1), 95–112. https://doi.org/10.2307/40285417

Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R. J., & Rodriguez-Fornells, A. (2013). Individual Differences in Music Reward Experiences. *Music Perception*, *31*(2), 118–138. https://doi.org/10.1525/mp.2013.31.2.118

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*. https://doi.org/10.1016/j.apergo.2020.103301

Miranda, D., Blais-Rochette, C., Vaugon, K., Osman, M., & Arias-Valenzuela, M. (2015). Towards a cultural-developmental psychology of music in adolescence. *Psychology of Music*, *43*(2), 197–218. https://doi.org/10.1177/0305735613500700

Miranda, D., & Claes, M. (2008). Personality Traits, Music Preferences and Depression in Adolescence. *International Journal of Adolescence and Youth*, *14*, 277–298.

Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons: From motion to 'e'motion. *Social Cognitive and Affective Neuroscience*, *1*(3), 235–241. https://doi.org/10.1093/scan/nsl029

Münsterberg, H. (1916). *The Photoplay*. Appleton.

Nowak, R. (2016). Music Listening Activities in the Digital Age: An Act of Cultural Participation through Adequate Music. *Leonard Music Journal*, *26*, 52. https://doi.org/10.1162/LM

Oh, Y. K., & Choeh, J. Y. (2022). Social media engagements of music videos on YouTube's official artist channels. *Convergence: The International Journal of Research into New Media Technologies*, *28*(3), 804–821. https://doi.org/10.1177/13548565211054906

Randall, W. M., & Rickard, N. S. (2017). Reasons for personal music listening: A mobile experience sampling study of emotional outcomes. *Psychology of Music*, *45*(4), 479–495. https://doi.org/10.1177/0305735616666939

Reason, M., & Reynolds, D. (2010). Kinesthesia, Empathy, and Related Pleasures: An Inquiry into Audience Experiences of Watching Dance. *Dance Research Journal*, *42*(2), 49–75. https://doi.org/10.1017/S0149767700001030

Reybrouck, M., & Eerola, T. (2022). Musical Enjoyment and Reward: From Hedonic Pleasure to Eudaimonic Listening. *Behavioral Sciences*, *12*(5), 154. https://doi.org/10.3390/bs12050154

Rich, M., Woods, E. R., Goodman, E., Emans, S. J., & DuRant, R. H. (1998). *Aggressors or Victims: Gender and Race in Music Video Violence*.

Ruggiero, T. E. (2000). Uses and Gratifications Theory in the 21st Century. *Mass Communication and Society*, *3*(1), 3–37. https://doi.org/10.1207/s15327825mcs0301_02

Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music*, *39*(3), 307–327. https://doi.org/10.1177/0305735610374894

Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. In *Music and Psychology Research* (Vol. 35, pp. 88–109). Zillmann and Gan. http://pom.sagepub.com

Schäfer, K., & Eerola, T. (2020). How listening to music and engagement with other media provide a sense of belonging: An exploratory study of social surrogacy. *Psychology of Music*, *48*(2), 232–251. https://doi.org/10.1177/0305735618795036

Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Frontiers in Psychology*, *4*(Aug). https://doi.org/10.3389/fpsyg.2013.00511

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

Skånland, M. S. (2013). Everyday music listening and affect regulation: The role of MP3 players. *International Journal of Qualitative Studies on Health and Well-Being*, *8*(1). https://doi.org/10.3402/qhw.v8i0.20595

Sloboda, J. A., O'neill, S. A., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the Experience Sampling Method. In *Muslcae Sclentlae Spnng* (Vol. 1, pp. 9–32).

Sloboda, J.A. (2013). Music in everyday life: the role of emotions. In Juslin, P. & Sloboda, J.A. *Handbook of Music and Emotion: theory, research and applications*. Oxford University Press. pg. 493-512

Smith, S. L., & Boyson, A. R. (2002). Violence in Music Videos: Examining the Prevalence and Context of Physical Aggression. *Journal of Communication*.

Straw, W. (2008). Music Videos. In *The International Encyclopedia of Communication*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405186407.wbiecm098

Sun, S. -W, & Lull, J. (1986). The Adolescent Audience for Music Videos and Why They Watch. *Journal of Communication*, *36*(1). https://doi.org/10.1111/j.1460-2466.1986.tb03043.x

Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and Memory for Musical Stimuli as a Function of Exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 370–381. https://doi.org/10.1037/0278-7393.30.2.370

Tan, E. S. (2018). A psychology of the film. *Palgrave Communications*, *4*(1). https://doi.org/10.1057/s41599-018-0111-y

Tarrant, M. (2002). Adolescent Peer Groups and Social Identity. *Social Development*.

Van Dyck, E., Burger, B., & Orlandatou, K. (2017). The Communication of Emotions in Dance. In M. Lesaffre, P.-J. Maes, & M. Leman (Eds.), *The Routledge Companion to Embodied Music Interaction* (1st ed., pp. 122–130). Routledge. https://doi.org/10.4324/9781315621364-14

van Goethem, A., & Sloboda, J. (2011). The functions of music for affect regulation. *Musicae Scientiae*, *15*(2), 208–228. https://doi.org/10.1177/1029864911401174

Vizcaíno-Verdú, A., De-Casas-Moreno, P., & Tirocchi, S. (2023). Online prosumer convergence: Listening, creating and sharing music on YouTube and TikTok. *Communication & Society*, *36*(1), 151–166. https://doi.org/10.15581/003.36.1.151-166

Vuoskoski, J. K., & Eerola, T. (2011). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, *15*(2), 159–173. https://doi.org/10.1177/1029864911403367

Vuoskoski, J. K., Thompson, W. F., McIlwain, D., & Eerola, T. (2012). Who Enjoys Listening to Sad Music and Why? *Music Perception*, *29*(3), 311–317. https://doi.org/10.1525/mp.2012.29.3.311

Zhou, R., Khemmarat, S., & Gao, L. (2010). *IMC '10: Proceedings of the 2010 ACM Internet Measurement Conference: November 1-3, 2010, Melbourne, Australia*. 404–410.

ORIGINAL PAPERS

STUDY I

EXPLORING MUSIC VIDEOS AND THEIR INFLUENCE ON MUSIC PERCEPTION

by

Johanna N Dasovich-Wilson, Marc Thompson & Suvi Saarikallio 2022

Music and Science vol 5

Exploring Music Video Experiences and Their Influence on Music Perception

Johanna N. Dasovich-Wilson [iD], Marc Thompson
and Suvi Saarikallio [iD]

Abstract

Streaming music videos on the internet is an increasingly popular music listening activity that has remained virtually unexplored within music psychology. Studies of the role of music in film, as well as empirical research investigating the influence of audio-visual media and memory, have shown that visual information can have a profound effect on how music is perceived and remembered. The current study aimed to create a framework for understanding music video (MV) experiences by finding out when and why individuals choose to engage with this form of media, how these experiences contribute to the perception of musical meaning and influence affective outcomes, and whether these effects carry over to subsequent listening experiences. An online questionnaire study was designed, and data were collected from 34 participants with a mean age of 22.4 years (SD = 2.79). Abductive analysis of the qualitative data was conducted based on theories derived from topical areas of music psychology research. A framework was devised which illustrates MV listening experiences over four temporal stages: Intention, Attention, Reaction, and Retention (IARR). The IARR framework provides novel insights into MV listening experiences and outcomes by shedding light on how extra-musical information can have a long-term influence on the perception of music's meaning and affective quality.

Keywords

Affect, audio-visual, framework, music video, perception, qualitative

Introduction

Music videos (MVs) offer a unique musical experience that allows listeners to engage with songs in an audio-visual format. Research has shown that pairing music with visuals can have a significant influence on the perception of the music's meaning and affective quality (Boltz, 2004; Boltz et al., 2009; Cohen, 2001). However, this research has focused almost exclusively on music in the context of film, where the music is meant to complement the film and guide the viewer's attention. In MVs, however, the video serves an entirely different purpose: to promote new singles and showcase the artist. Global reports on music consumer behaviour have shown YouTube, a platform on which MVs can be easily accessed from any personal device from virtually anywhere at any time, is a leading resource for music streaming. According to the International Federation of the Phonographic Industry's (2019) report, 47% of all music streaming occurred on a video platform such as

YouTube, and 77% of people surveyed reported using YouTube to stream music in the last month. Furthermore, 95 out of 100 of YouTube's most viewed videos are MVs, all of which have over a billion views (see YouTube, 2021). This highlights the need for music psychology research to examine MVs, including the reasons for which individuals choose to engage with them, and their potential effects on listening experiences and outcomes.

In this study, we investigated MV listening experiences from an everyday music listening perspective and collected qualitative responses via an online questionnaire study. Our

Department of Music, Arts and Culture Studies, University of Jyväskylä Jyväskylä, Finland

Corresponding author:
Dasovich-Wilson, Department of Music, Arts and Culture Studies, University of Jyväskylä, P.O. Box 35, FI-40014, Jyväskylä, Finland.
Email: johanna.n.wilson@jyu.fi

aims were to uncover why individuals choose to engage with this form of music media, how these experiences influence their perception of the music's meaning and affective quality, and whether the extra-musical information offered in the video influences subsequent, audio-only listening episodes. The study draws on insight developed from theories of music-evoked emotion, the use of music in everyday life, and the cognitive processing of music and film, and highlights how these theories work together during MV experiences. In this introductory section, we first examine current theories on the cognitive processing of music in film. Second, we examine findings from previous research concerning everyday music listening experiences and music's influence on affective states. Theoretical knowledge from these areas of research informed the data analysis and development of the theoretical framework outlined in this article.

Cognitive Processing of Audio-Visual Information

MV content is highly variable, often featuring scenes of musicians performing, using narratives to portray musical meaning, or a combination of both. Theoretical models of multimedia and the cognitive processing of multimodal stimuli provide the foundational knowledge necessary for understanding audio-visual interactions across different multimedia contexts. According to the Congruence-Association model (CAM; Cohen, 2001, 2013; Marshall & Cohen, 1988), music supports the emotional meaning of film by guiding the viewer's attentional and inference processes. Film music (which refers to any multimedia context, such as television, advertisements, movies, etc.) contributes to the perception of emotional meaning, establishes mood, and evokes emotions in the audience (Cohen 2001, 2013). Film also provides two kinds of information: structural (temporal) and semantic (meaning). The viewer's attention is directed towards structural features shared by the film and the music, such as accent patterns and synchronicity. The meaning of the music, which Cohen defines as "associations the music brings to mind" (2013, p. 23), is assigned to the object or events in the film to which the viewer's attention has been directed.

These associative elements shed light on the role of long-term memory (LTM) in the cognitive processing of complex audio-visual material. More recent versions of CAM highlight the importance of the viewer's LTM in the processes of drawing emotional inferences from sensory information derived from the narrative of the film (Cohen, 2013). How the listener constructs the meaning of music – as well as that of visual scenes, sound effects, and speech, whether presented in the soundtrack or across the screen as text – depends on their past experience. Semantic associations are elicited by these sources of information, which are then retrieved from LTM. The viewer's attention is directed towards information in the film which matches those associations, creating an interpretation of the film's narrative.

Cohen's model illustrates how structural and semantic elements of each modality guide attention and generate inferences about the multimedia being consumed. However, what happens when the individual's perceived meaning of a song or piece of music is not in line with what is being depicted visually? This is particularly problematic for MVs, especially in cases where the song is associated with, for example, autobiographical memories or other personalized meanings. MV viewing potentially ends with the viewer generating a new meaning about the song, even if this effect is accidental (or undesired). This sheds light on what Cook (1998) refers to as the *contest* model of multimedia. Contest occurs when each medium (the song and video) is its own, independent source of meaning, even if the meanings they represent are similar. New meaning is generated through a dynamic process in which each medium tries to impose its characteristics on the other. On the other hand, the relationship between the song and the video may have a more complementary function. The conflicting meanings indicative of contest are avoided because each medium serves a separate role within the MV context. Instead, the video complements what is already present in the music, and vice versa. Both Cohen and Cook's models provide insight into how audio and visual information interact with each other and the viewer's subjective experience in multimedia contexts to generate meaning.

Research on MVs can provide a better understanding of how visual information influences the interpretation, perception, enjoyment, and recall of musical material in an ecologically valid context. Laboratory studies have demonstrated how music can influence the recall of film events (Boltz, 2004), as well as how visual information can bias the perception of the music's acoustic properties (Boltz et al., 2009). MVs, however, provide an opportunity to better understand how music and visual information interact with each other in a more ecologically valid context, and the influence this interaction has on music perception and enjoyment. Boltz (2013) draws attention to these issues, highlighting the different components of MVs which have been investigated empirically by drawing on research in music education, musical performance gestures, and theories of audio-visual interactions. For example, in addition to (or in lieu of) a storyline, MVs often show musicians performing or include "live" scenes from concerts. Previous studies have analysed how the performer's movements contribute to the viewer's appraisal and perception of emotion during a musical performance (e.g., Davidson, 1993; Vines et al., 2011). Being able to see the performer's movements can either lessen or intensify the perception of tension in the music, thus providing a more complex or nuanced interpretation of the music's emotional quality (Vines et al., 2006). Unlike experimental stimuli, however, MVs are more complex. It is common for MVs to include scenes of the artist performing, as well as dancing. How this contributes to the perceived and felt emotional quality of the music has yet to be investigated in MV contexts specifically, however.

Considering the popularity of MVs and their role in today's music listening culture, further research which explores their effects is imperative.

Affective Responses in Everyday Music Listening Experiences

In recent years there has been increased interest in studying the function of music in everyday life. Music is recognized for its ability to fulfil several psychological functions, including to regulate affective states such as moods and emotions, for self-reflection and for social bonding purposes (e.g., Hargreaves & North, 1999; Schäfer et al., 2013). Music begins to adopt these functions during youth, a time when it is used as a tool for psychosocial development (see Laiho, 2004 for an overview); however, it continues to serve these functions into adulthood (Saarikallio, 2010). Throughout this literature, MV viewing has received little attention. While MV experiences may fulfil the same functions as "audio-only" music listening, their contribution to these functions, and whether they help or hinder them, is not yet understood.

One early study to explore MVs was conducted by Sun and Lull (1986), who investigated adolescents' reasons and motivations for watching MTV (the Music Television Channel), which exclusively aired MVs at the time.[1] Their findings suggest that the reasons why youth engaged with MTV went beyond those usually identified for regular TV viewing or music listening, since MVs allow the audience to discover the "true" meaning behind popular songs. This contrasts with other studies on everyday music listening, which posit that music is frequently listened to in the background during another, primary activity such as doing chores, commuting, or exercising (Sloboda et al., 2001).

These findings shed light on why it is important to understand the role of context when evaluating the functions of music listening. For example, while music's emotion regulating function has been regarded as one of its most essential functions (e.g., Baltazar & Saarikallio, 2016; Rentfrow, 2012; Schäfer et al., 2013), an ESM (Experience Sampling Method) study by Randall & Rickard (2017) found that music listening for emotion regulation purposes only occurs in around 32% of listening episodes, and in 59% of cases where the listener is in a negative mood. Furthermore, whether music properly fulfils this function depends on the individual's general emotion regulation tendencies, since individuals who use more maladaptive regulation strategies tend to experience more negative effects from music listening than individuals who use healthier, more adaptive strategies (Chin & Rickard, 2014; Saarikallio et al., 2015; Saarikallio & Erkkilä, 2007).

Music listening strategies for mood regulation, such as listening to be entertained, to distract from negative thoughts, or to find solace (see Saarikallio & Erkkilä, 2007), can also apply to MV experiences. For example,

the motivational factors for watching MVs reported by Sun and Lull (1986) are in line with those of Saarikallio and Erkkilä's (the equivalent strategies are listed in parentheses), including: boredom relief (entertainment), to relieve tension (revival), for distraction (diversion), and to feel less alone (solace). While Sun and Lull posit that MTV watching goes beyond regular music listening, these motivations they outline reflect the same functions as music listening. For example, other motivational factors for watching MVs are in line with the self-awareness and social relatedness functions highlighted by Schäfer et al. (2013), including: information/social learning ("learn more about self/others", "understand the world", "supports my ideas"), and social interactions ("conversation topic", "do with friends").

Music listening experiences and the outcomes they elicit also vary depending on the music, the situation and the listener (Juslin et al., 2008; Liljestrom et al., 2012). Individuals who are more engaged with music overall listen to music or participate in musical activities for more hours a day (compared to less musically engaged individuals) and use music to fulfil several functions simultaneously (Greasley & Lamont, 2011). Situational variables are also important to consider, since research has shown that the activity accompanying music listening is the most important in determining function, followed by control over what is being listened to and attention paid to the music (Greb et al., 2018). For example, for the function "Intellectual Stimulation" the activities with the strongest, positive relationships were: "making music", "pure music listening", and "working and studying". Certain functions are more affected by individual differences than others, with musical taste and strength of preference being the strongest individual differences predicting music's function. Overall, the importance of music to the individual is an important variable which influences music listening behaviour (Krause & North, 2017). Furthermore, devices which allow for personal input and control over what was being listened (such as smartphones and MP3 players) yield more positive affective outcomes on the listener, such as contentment and an increase in motivation (Krause et al., 2015). This is particularly important in the case of modern MV listening experiences: unlike the early days of MVs, where they were watched on television channels dedicated to music (for example, MTV), individuals today can watch virtually any MV, anywhere, at any time on their personal devices on YouTube. This gives the listener more control and choice over which MVs they engage with.

Interactions between the listener, the acoustic features of the music, and the situation where the music is being heard are important to consider when analysing how meaning, including emotional meaning, is attributed to music. Cespedes-Guevara and Eerola (2018) suggest a constructionist approach to the perception of affect in music in order to account for the interactions between the listener's knowledge, their listening goals and current psychological

state, the features of the music, and the context where the listening experience takes place. They draw on Barrett's (2006) Conceptual Act Theory of emotion, which posits that experiencing an emotion (or observing it in somebody else) occurs when top-down knowledge from past emotional experiences is combined with sensory information from our bodies, or from witnessing another person's behaviour. In the case of music, Cespedes-Guevara and Eerola posit that music can afford specific meanings, including emotional ones, because of the cognitive processes that occur when the listener combines top-down knowledge from past musical (and emotional) experiences with information about their current affective state, and the context where the musical event takes place. Associative mechanisms allow the listener to integrate information from the music's acoustic cues with other sources of information available in the listener's mind allowing them to construct meaning and perceive emotions in music. These associative mechanisms and their role in the construction of meaning are also present in CAM (Cohen, 2001, 2013; Marshall & Cohen, 1988). MVs may influence the types of associations the listener attributes to the music, or even replace them with new ones, if the information they contain confirms or violates their expectations about the music's meaning.

Theoretical insights concerning the mechanisms of music-evoked affect may also provide insight into MV effects. Of particular interest to the current study is the *Visual Imagery* mechanism component of the BRECVEMA model of music-evoked emotions (see Juslin, 2013; Juslin et al., 2014; Juslin & Västfjäll, 2008). Visual imagery (or visual *mental* imagery) refers to emotional outcomes evoked by music as a result of the inner images conjured by the listener. Providing contextual information about a piece of music, such as narrative descriptions, may influence or enhance this mechanism (Vuoskoski & Eerola, 2013). Since MVs often contain visual information about the music's meaning, there is the potential that this imagery becomes associated with the music and conjured as visual imagery in future listens. Preferred music and familiarity also contribute to affective responses to music (Schubert et al., 2014; Szpunar et al., 2004). For example, *spreading activation theory* describes how preference and familiarity predict affective reactions to music. As an individual becomes increasingly familiar with a certain genre, artist, or song, they develop more mental representations (referred to as *nodes*) and a network of associations connected to the music is formed. When familiar music is heard, the network is activated, resulting in aesthetic pleasure (Schubert et al., 2014). MVs may potentially increase the number of nodes activated during listening by providing new associations with the music. This may also explain how MV experiences can help slow down wear-out from over-listening (Goldberg et al., 1993).

Mental representations and associations reflect how the listener's long-term memory influences affective responses to music. This is particularly important in the case of MV experiences since there is the potential for more

associations to form as the viewer-listener attends to the visual and musical content. These frameworks, in addition to CAM, can provide a rich insight into how MV listening experiences influence the perception of and affective responses to music.

The Current Study

The current study explores MV experiences and their effects on listening outcomes. The current study is part of a larger survey study, which also consisted of quantitative data and from which a Master's thesis (Wilson, 2018) and a less developed proceedings paper (Wilson, et al., 2020) have previously been published, but focusing on different perspectives. The study features an online, open-ended questionnaire designed to gain insight into participants' experiences with this form of musical multimedia, including the circumstances that lead to the experience, the experience itself, and the perceptual consequences for future listens. A preliminary framework was devised which outlines the reasons and situations in which MV experiences occur, the cognitive and emotional outcomes elicited during the experience, and the carry-over effects they impose on subsequent, audio-only experiences. The study specifically asked participants whether the MV had any influence on how they perceive the meaning of the music going forward, as well as whether they experience visual mental imagery related to the content of the MV in future listens.

Method

Participants

Qualitative data were collected from a convenience sample of 34 participants. Participants were recruited from a university music theory class email list of approximately 100 students and via social media. All participants were adolescents and young adults between the ages of 15 and 27 ($N = 34$; M = 22.4, $SD = 2.79$): 53% identified as female, 41% identified as male and 6% chose not to disclose their gender. Most of the sample (23 participants) identified as Canadians, eight participants preferred not to disclose their nationality, and the remaining three identified as American, Australian, and Korean. Two participants identified as having a second nationality in addition to being Canadian (Chinese and Dutch). Participants who provided an email address were entered in a raffle to win an Amazon gift card (value of $25 CAD) as incentive to participate. The study received ethical approval from the Ethics Committee of the University of Jyväskylä.

Study Design and Procedure

The study was available online on the platform Qualtrics and took approximately 25 min to complete. The study was designed to be completed in an environment where

the individual would usually find themselves watching MVs naturally, such as at home, in order to create the most authentic experience possible and minimize experimenter influence. Furthermore, participants were asked to watch an MV they had already seen before that they enjoyed prior to starting the questionnaire, which participants could then reference in their response. This type of elicitation technique was implemented in order to limit recall bias as well as to promote discussion and elaboration of ideas on the topic (Barton, 2015). The study consisted of an analysis of 12 open-ended questions designed to elicit descriptive responses from participants about their experiences with MVs. A total of 10,520 words were analysed and the average number of words provided per participant was 309 (min = 111, max = 1,288).

The open-ended questionnaire addressed where and why participants would watch MVs, how these experiences compare to audio-only listening experiences, and the extent to which this multimodal listening format influenced the perception of the music's meaning and affective quality, both during the initial MV experience and in future, audio-only listening experiences. In addition, participants were asked to indicate from a list of items which electronic devices they use to watch MVs.

Method of Analysis

In order to establish a preliminary framework for understanding participants' experiences, an abductive grounded theory approach was adopted. According to Charmaz (2008), grounded theory is a method for analysing qualitative data that is well suited for investigating phenomena that have been underexplored or where understanding is currently limited. This method calls for the researcher to remain open to various explanations and limit their preconceptions about the problem they are investigating (Charmaz, 2006). While research that takes a grounded theory approach starts inductively, it moves into abductive reasoning as the researcher attempts to interpret and understand the phenomena observed in the data, allowing them to consider all the possible theoretical explanations while also remaining open to alternative interpretations. This approach allows the researcher to revisit and defamiliarize themselves with the data after exploring the phenomena at hand from different theoretical perspectives (Timmermans & Tavory, 2012).

This abductive method to qualitative analysis was considered suitable for exploring MV experiences in the current study since many facets or components of this type of listening style have been explored empirically. For example, there already exists a large body of research dedicated to understanding the psychological function of music in everyday life, the cognitive encoding of audiovisual stimuli, and theories of music-evoked emotion, all of which are relevant to the study of MVs but have yet to be applied to them explicitly.

Coding Procedure

All coding was performed by the first author; however, all authors were consulted throughout the coding process to discuss any potential ambiguities or theoretical explanations for phenomena encountered in the data. The first step of the coding process was directed towards identifying (1) the antecedent factors that motivated individuals to watch MVs; (2) the MV listening experience itself; and (3) carry-over effects on subsequent listening experiences in the data. Once data were organized according to these three temporal stages, they were sub-coded using the in vivo method. In vivo coding was used in order to segment passages into individual codes when they described different phenomena. This method is useful as it also preserves the language used by the participants and keeps the analysis grounded in the data (Charmaz 2008; 2006; Saldaña, 2016). If codes inferred more than one meaning or described more than one aspect of the MV listening experience, simultaneous coding was used. Once the data were saturated and all codes were identified, axial coding was performed.

During the axial coding phase, codes describing similar phenomena were grouped together, creating the main categories and subcategories within each temporal level. The frequency of each category and subcategory is reported as a percentage in the results, and each participant is considered an individual case. It is important to note, however, that since the data were collected using open-ended questions on an online platform and not, for example, with semi-structured interviews, there was no way to follow-up with participants about their answers. Therefore, the analysis was limited to the amount of data provided by each participant, which varied in length and substance. The percentages and number of cases provided in the results section are only representative of their occurrence *within the current study*. The authors acknowledge this limitation; however, the purpose of this study is to create a preliminary framework to serve as a starting point for future studies investigating these phenomena. Suggestions for future research are outlined in the discussion.

Results and Discussion

Four primary categories were established from the data which revealed new insights into MV listening experiences. Themes emerged in the data which described the contexts and personal goals individuals aimed to achieve, how attentional processes are directed during the experience, the circumstances that evoke and the nature of affective outcomes experienced from watching MVs, and the potential carry-over effects that manifest in subsequent listening experiences. These four levels, referred to as Intention, Attention, Reaction, and Retention (IARR) are defined as follows (see Figure 1):

1. **Intention:** Describes the circumstances that lead to MV listening experiences, including the goals the

participant intends to accomplish, the contexts where these experiences take place, and whether the participant is alone or with others.

2. **Attention:** Describes how participants direct attention, what factors distract or detract from their enjoyment of the music, and the conditions that direct attention towards one modality over the other.

3. **Reaction:** Describes the affective outcomes of MV watching and the contingent factors that influence these outcomes. This level describes the mechanisms responsible for MV evoked affect, as well as provides insight into when and why MV evoked affect is considered more salient compared to outcomes evoked during audio-only listening experiences.

4. **Retention:** Describes how the MV watching experience influenced subsequent, audio-only listening experiences, and the mechanisms responsible for evoking affective outcomes in these subsequent listens.

Intention

The Intention level outlines the reasons and motivations for watching MVs. Participants also selected from a list of items which devices they usually watch MVs on, with smartphones and laptops being selected the most frequently (see Figure 2). Devices listed in the event they selected "other" included one participant who stated using a gaming console, and three who stated they watch them on

Figure 1. The four temporal levels of MV listening experiences: Intention, Attention, Reaction, Retention (IARR). Intention refers to all antecedent factors leading to MV engagement. The experience of watching MVs is divided into two levels (Attention and Reaction). Retention refers to all carry-over effects imposed on future listening experiences.

TV (for example, via MTV). On average, participants watched four MVs a week (min = 0, max = 20).

Three categories were established in the data describing the psychological goals, social factors, and preference related reasons for watching MVs (see Figure 3). Affect-related (i.e., emotional reasons) and cognitive goals were observed in 67.7% of cases and are categorized under the label Internal factors. Social motivations, categorized as External factors, were observed in 50% of cases. A third factor that motivated individuals to watch MVs is illustrated in the category labelled Preference-Driven, which includes the 44% of cases where MV watching is motivated by listener's enjoyment of specific songs or artists. Example codes for Intention categories and subcategories are found in Table 1.

Internal factors were subcategorized depending on whether the Intention reflected an *emotional goal* or *cognitive goal*. *Emotional goals* were observed in 47% of cases and describe MV engagement for the purpose of regulating, maintaining, or enhancing affective states, as well as engaging with MVs for aesthetic enjoyment or in order to relieve boredom. Boredom relief was considered an emotional goal since it illustrates how engaging with music for entertainment functions as an affect regulation strategy (Saarikallio & Erkkilä, 2007). *Cognitive goals* were observed in 41% of cases and include using MVs to reflect on the content of the music or video, such as the true meaning of the lyrics, as well as watching in order to engage with or learn any physical performative elements featured in the video (such as musical technique or dance choreography).

External factors reflect the role of social influences as an antecedent to MV watching. Two subcategories emerged, labelled *peer influences* and *media influences*. *Peer influences* were observed in 26.5% of cases, these cases describe in-person or online exchanges with friends as an antecedent factor motivating MV engagement, as well as watching MVs with friends as a social activity. *Media influences* were observed in 23.5% of cases and describe antecedents related to the "hype" surrounding a song, for example because it is trending on social media, as well as cases where YouTube auto-played or recommended the MV to the participant.

The Preference-Driven category consists of reasons for watching MVs that were motivated by the participant's familiarity with and preference for certain artists or songs. Watching newly released MVs for singles by their favourite artists are also included under this category. Preference-Driven codes were homogenous enough to not require further subcategorization.

Results at the Intention level provided insight into the motivating factors (goals) and contexts that led participants to engage with MVs. All participants provided at least one reason for engaging with MVs or choosing them over audio-only listening; however, others provided several depending on factors related to who they were with at the time, what mood they were in, or what they were doing at the time. These reasons complement existing research
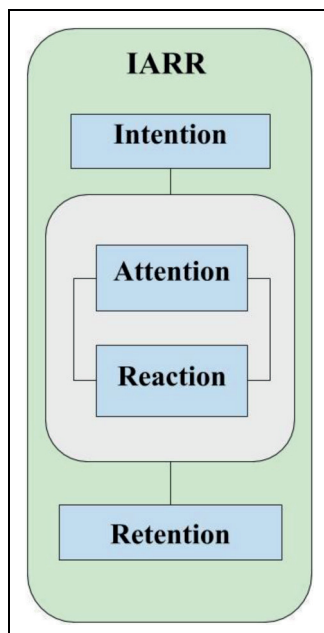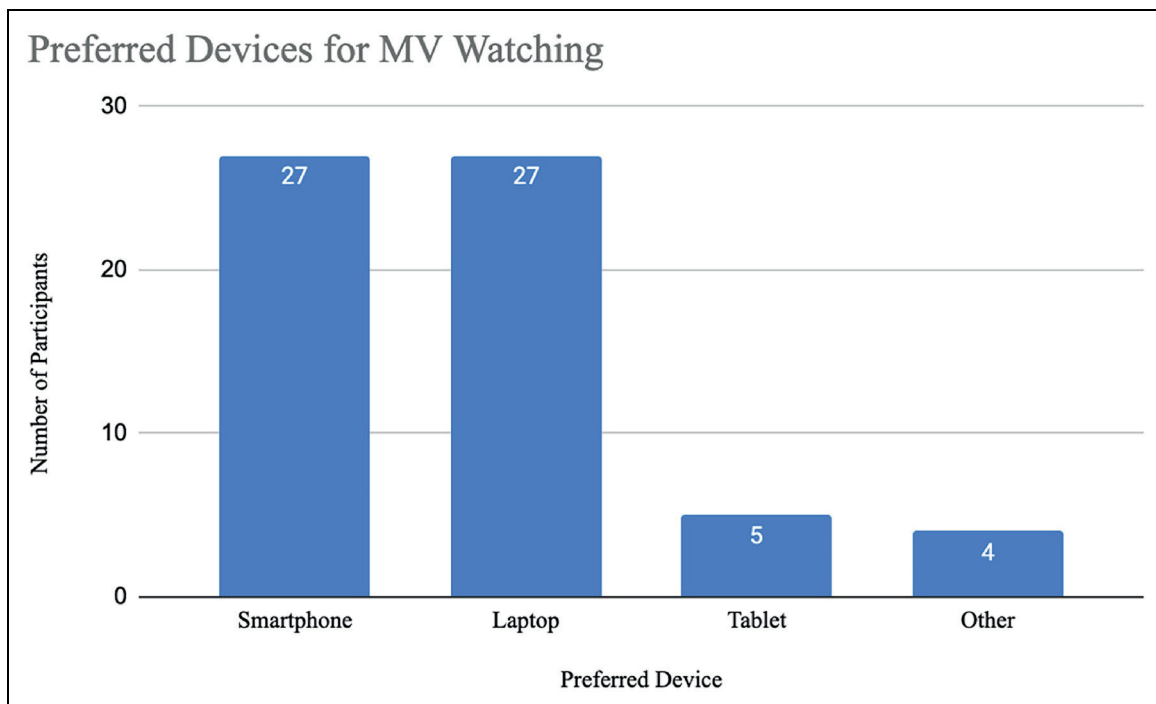
Figure 2. Preferred devices for watching MVs. Participants were allowed to select one of the three options provided, or specify the device if selecting "Other". Other devices were non-portable, such as watching on television, or via YouTube on a gaming console.
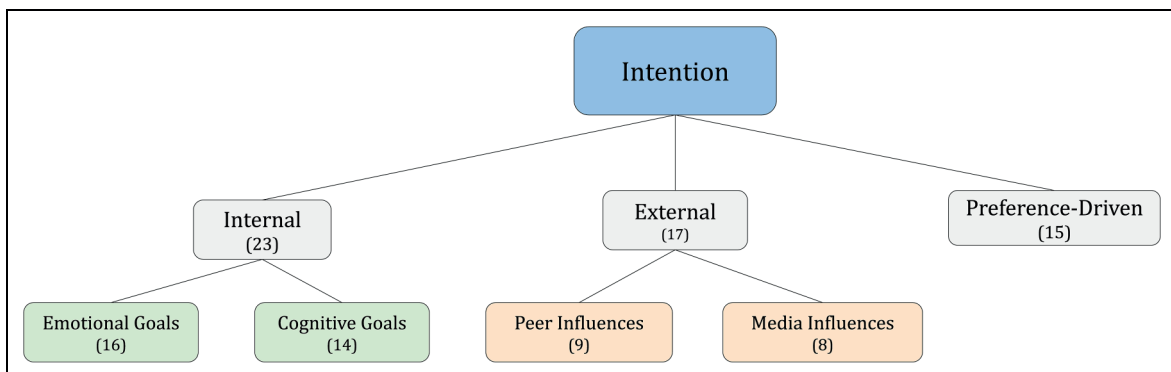
Figure 3. Intentions for watching MVs. The number of participants coded under each category and sub-category are represented in brackets.

highlighting the use of music to fulfil different emotional, cognitive, and social needs, especially during youth (Laiho, 2004; Tarrant et al., 2000). Although the data from the current study were collected over 30 years after Sun and Lull's (1986) study similar findings were observed: in both studies, learning the "true" meaning of the music was a significant motivating factor for MV watching. Although the age demographic in the current study consists of both high school students and young adults, our findings complement the social, emotional, and information-seeking factors originally highlighted by Sun and Lull. On average, participants in Sun and Lull's study watched approximately two hours of MTV a day, whereas participants in the current study watched an average of four videos per week.

However, participants in the current study had more control over which specific MVs they wanted to see and when they wanted to see them. These findings provide more insight into the ways modern technology affords more complex interactions with music and the functions it can fulfil in different listening situations (Greb et al., 2018; Krause et al., 2015; Randall & Rickard, 2017).

Individual differences in musical engagement can provide further insight concerning the frequency and effectiveness of MV listening, particularly when the experience is aimed at fulfilling specific psychological or social functions. For example, Preference-Driven reasons may fulfil Internal and External goals concurrently by enabling mental work, regulating affective states, or contributing to

Table 1. Description of intention categories and subcategories with examples.

Intention Categories and Subcategories

Category/ Frequency	Subcategory, Frequency and Themes	Example (Participant number)
Internal Factors (Self-motivated) 67.7% of cases	*Emotional Goal* 47% of cases • Watches for aesthetic enjoyment • Watches for boredom relief • Watches to change or maintain their current emotional state	*When I want to fully experience the music instead of playing [it] in the background while I'm working. (P11)* *[I watch music videos] during my down time. When I'm procrastinating. (P36)* *When I am relaxing at night or having an anxiety attack. (P34)*
	Cognitive Goal 41% of cases • Watches to understand the meaning of the music • Watches to see the artist perform and/or to learn the music • Watches to or see and/or learn choreography	*[To] discover what message [the artist] was trying to get through in the song. (P10)* *…if it is a performance style music video where I could gain visual information [as to] how they produced the music. (P21)* *I like to have choreography (e.g. dance) in music videos as [I] find it so fascinating how some moves to this particular song with their body. (P04)*
External Factors ("Other"-motivated) 50% of cases	*Peer Influence* 26.5% of cases • Watches with others as part of social activities • Watches MVs shared by friends	*When [I] am with friends. (P30)* *If someone had sent me a link to the song on YouTube, I'm more likely to end up watching the video. (P25)*
	Media Influence 23.5% of cases • Watches during social media engagement, browsing YouTube • Watches MVs that have received media attention	*If I've already been browsing YouTube. (P01)* *Mostly just watch music videos when a new one comes out that gets some buzz. (P35)*
Preference-Driven 44% of cases	• Watches MVs put out by their favourite artists (including new releases) and/or for favourite songs.	*I'm also more likely to watch video from my favourite artists or for songs that are really catchy. (P19)* *…or it's a new music video that I have not seen before to artists and songs that I like. (P01)*

social interactions between peers. While the current study succeeded in highlighting the diverse reasons for MV engagement, the frequency and effectiveness of this listening method for fulfilling specific functions requires further investigation.

The Intention level of the IARR framework provides insight into the reasons and contexts for MV engagement, however, there are limitations to consider. Individual differences such as personality, musical engagement, musical expertise, and mental health have been acknowledged in previous studies as mediating factors that influence music listening outcomes (e.g., Chamorro-Premuzic & Furnham, 2007; Chin & Rickard, 2014; Saarikallio et al., 2015). Future research using quantitative methods, such as surveys, can help provide a clearer picture of which functions MV experiences can fulfil and for whom, since previous research has identified how individual differences in people's level of engagement with music (highly engaged compared to less engaged) influences the reasons, contexts, and outcomes experienced in response to music (Greasley & Lamont, 2011). These reasons do not only highlight potential differences in musical engagement at the participant level, but also the role of MVs in today's popular

culture and the dissemination of music in general. The External factors outlined in this study suggest that many individuals do not restrict their engagement with MVs to songs that match their listening preferences: they also watch out of curiosity if the MV has received attention or notoriety in the media. This may be the case for MVs shared with them by their peers; however, it is unclear in the current set of data whether this reflects sharing between peers based on shared preference, or for other reasons related to the content of the *video* (for example, cameos from other celebrities, socio-political commentary, or controversial themes) and not necessarily the music.

Interestingly, MVs were also selected in cases where the song was not available on other streaming platforms. This suggests the video component was not always attended to or even desired by the listener. Findings related to attentional processes are discussed in greater detail in the next section. However, in respect to Intention, it is important to note that in some cases MVs are selected because of their being freely and easily accessible on YouTube. How exposure to this visual content influences listening outcomes, including those elicited in subsequent, audio-only listening experiences, is discussed in the Reaction and Retention sections of the results.
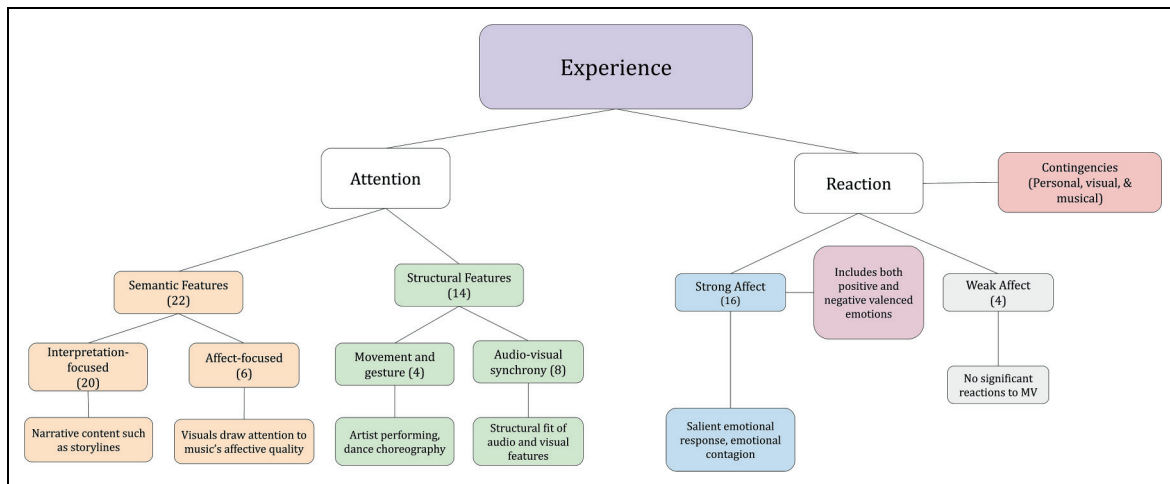
Figure 4. Experience is divided into two levels: Attention and Reaction. The number of participants coded under each category and subcategory are represented in brackets.

MV Experiences: Attention and Reaction

Data describing the experience of watching MVs were split into two categories: Attention and Reaction. These top-level categories provide more insight into how attentional processes are directed during the experience and the affective outcomes that arise as a result. Attention and Reaction codes reflect the psychological processes that occur *during* the MV experience, and the *immediate* affective response evoked by their content. The qualitative results for Attention and Reaction are reported separately for clarity. They are briefly discussed individually as well as cumulatively, as the content that absorbs attention influences the types of affective outcomes elicited by the MV experience (see Figure 4).

Attention

The Attention level describes the specific content or features of the MV which participants describe as having absorbed their attention during the MV experience. Two categories were established in the data that describe the features absorbing participants' attention, labelled here as Semantic and Structural features. Example codes and definitions of categories and subcategories are found in Table 2.

Participant reports that describe being focused on how the content of the MV influences the meaning or affective quality of the music were categorized under the label Semantic features. These features were reported in 65% of cases and describe how the MV provided more context about the meaning of the music, as well as features that directed their attention towards and enhanced their perception of the music's affective quality. Two subcategories, *interpretation-focus* (IF) and *affect-focus* (AF) were created to reflect the qualitative differences between experiences and occurred in 58.8% and 17.6% of cases, respectively. The primary difference between these subcategories is the language choice used in these codes. IF experiences describe

directing attention towards narrative elements such as storylines. On the other hand, AF codes do not focus on storylines: they describe how the visual aspect directs attention towards the music's affective components, completing, complementing, or clarifying them to the viewer (how this influences their affective reactions is discussed in greater detail below).

Participant reports that described being focused on how the features of the video (such as editing techniques, movement, colours, textures, etc.) aligned with or complemented the music's structural features (such as tempo, rhythm, melodies, harmonies, and lyrics) were categorized as Structural features. Structural features describe how attention was directed towards the synchronous relationship between the music and the video, and/or directing focus on human movements such as dance and performance gestures. Two subcategories emerged that reflect these concepts: *audio-visual synchrony* (AVS) and *movement and gesture* (MG), which were observed in 23.5% and 12% of cases respectively.

The perceived synchrony between the audio and visual modality guided and absorbed participants' attention during the MV experience. Codes categorized as AVS describe how the physical features of the video, such as camera angles, colours, and video editing techniques, aligned with or complemented the features of the music. These codes contrast those belonging to the Semantic features category since they highlight how the physical (i.e., structural) attributes of one modality direct attention to or highlight the coinciding attributes of the other.

Attention results complement previous research concerning the role of music in film and the cognitive processes highlighted in CAM (Cohen, 2013; 2001; Marshall & Cohen, 1988). Participants were aware of how the video influenced their attentional processes; in some cases, they asserted that their experience of the music was enhanced overall as a result of being more absorbed in the music, its narrative elements, and/or structural interactions between the audio and the video. Furthermore, our findings also shed light on how

Table 2. Descriptions of attention categories and subcategories, with examples.

Attention categories and subcategories

Category	Subcategory, frequency, and themes	Example (participant number)
Semantic features 65% of cases	*Interpretation-focus (IF)* 58.8% of cases • Focuses on how the meaning is represented with narrative elements such as storylines *Affect-focus (AF)* 17.6% of cases • Focuses on how the affective quality of the music is represented or clarified by the video	"I immerse myself more in the story of a song when watching a video." (P05) "I focus on how the video tells the story." (P23) "This video was interestingly edited and completed the mood of this song…" (P22) "I find having a visual often makes things much more explicit, even if there isn't a plot per se in the video… it can solidify parts of the song in a way that had been more nebulous in my head. Making feelings into images." (P25)
Structural features	*Audio-visual synchrony (AVS)* 23.5% of cases • Focuses on how well the properties of the music and video "fit", how they complement or contrast each other *Movement and gesture (MG)* 12% of cases • Focuses on musical performance gestures and/or choreography	"I focus on how the video [enhances] the instruments of the song." (P23) "Imagery that fits the tempo and lyrics of the song." (P28) "I tend to really watch how the [dancers'] movement in the video goes with the music [sic]." (P04) "It's more similar to a live viewing of it, where it's easier to feel the energy of it by not only hearing the music, but seeing it performed." (P21)

music perception (as opposed to visual perception) is influenced by audio-visual interactions, a finding previously reported by Boltz and colleagues (2009). Our results extend these findings by providing insight into how these interactions influence the perception of music with which the listener is already familiar (unless the MV experience is their first exposure to the song, which in most cases it was evident by their description that it was not). The influence of audio-visual interactions on affective outcomes is discussed in greater detail in the following section.

Reaction

Codes at the Reaction level offered new insights into how MV information influenced participants' affective outcomes. The extent to which MVs influenced affective outcomes on both valence and arousal dimensions varied among participants. Furthermore, it was commonly stated that whether an MV elicited any significant affective outcomes depended on factors related to their previous experiences with the music (or lack thereof) and the content of the video. These factors, which were first reported by Wilson et al. (2020), describe the musical, visual, and personal contingencies mediating affective responses to MVs and their underlying mechanisms (see Table 3).

Two Reaction categories, labelled Strong Affect and Weak Affect, were formed based on the extent to which MV experiences influenced participants' affective state, (see Table 4). Unlike codes belonging to other temporal levels, which could be simultaneously coded into more than one category or subcategory when appropriate, Strong and Weak Affect codes are mutually exclusive.

Table 3. Contingent factors of MV listening experiences (from Wilson et al., 2020; reprinted with permission).

Factors	Contingencies
Musical	Genre, emotional quality, artist, familiarity
Visual	Portrayal of emotion, presence of storyline or performance gestures, quality of cinematography
Personal	Appraisal of the music, current affective state, whether the MVs reflects their personal interpretation of the music

However, if the participant experienced outcomes categorized as Strong Affect and described which musical, visual, or personal factors were responsible for that outcome, simultaneous coding was used to account for those factors.

Strong Affect was observed in 47% of cases: these cases contain codes describing MV listening experiences which evoked salient affective responses. These cases highlight themes such as emotional contagion, feelings of connectedness, and strong sensations. Strong Affect includes examples of both positively and negatively valenced emotions evoked during or in direct response to the MV experience (for example, feeling sad in response to a MV of a sad song).

Weak Affect codes were less descriptive and reported less frequently: only 11.7% of cases assert that the MV experiences did not significantly influence their affective state. Although Weak Affect was only observed in four cases, half of them asserted that this lack of reaction was due to the *video* specifically and not the music. The Weak Affect categorization does not necessarily infer these participants are less reactive to music in general, nor does it rule

Table 4. Description of reaction categories and examples.

Reaction categories

Category	Description	Example (participant number)
Strong affect 47% of cases	MV elicits salient affective reactions. Feels more connected to the music and/ or artist. Emotional reactions mimic the emotional tone of the music (contagion).	"I think it just amplifies the mood of the song. If it's a sad song you feel more sad. If it's happy you feel happier." (P10) "I find [the video] makes an already powerful moment in the song even more intense… If it is a well made video, it typically just enhances the mood which the song already puts me in." (sic) (P16) "It can make whatever emotion I'm feeling become stronger." (P27) "… It made me feel like I knew her and felt all of her emotions inside me (that sounds so dramatic!) I cried at a lot of parts of it when watching it. Everything about it was SO in line with the things I'd already felt when listening to the music alone that it enhanced it so much beyond the sum of its parts." [sic] (P25)
Weak affect 11.7% of cases	MVs experiences do not elicit any significant emotional outcomes. MVs are described as less impactful than audio-only experiences.	"Videos don't tend to alter my mood. The music is the important part for me." (P12) "Sometimes a video will bore me, even though I like the song… [My emotional reaction is greater] when only listening." (P30)

out the possibility that the music could have still influenced their mood during the MV experience, even if the *video* did not contribute to these outcomes.

Although participants discussed their experiences with MVs they enjoyed, 14.7% of cases described how the MV negatively impacted their experience of the music. The factors or mechanisms that lead to this negative experience are also reflected in the contingencies outlined in Table 3, especially those found at the personal and visual level. For example, if the MV is appraised as being of inferior quality, the experience of the music suffered as a result (P16): "…some videos are hard to understand or poorly made and therefore take away from the experience of listening to the song." For other participants, it was not necessarily the quality of the video itself, but the interpretation of the music that caused them to negatively appraise or even avoid MVs for songs they enjoyed completely (P25):

> I honestly find that music videos often distract from the listening experience instead of enhancing it…because often how they [the artist] interpret their song in a video is much different than the feelings or visuals I may have had when I listened to it on my own, and therefore if I really love a song sometimes I'll consciously avoid the video (if there is one) because I don't want to know if the way I love the song is not how the artist feels the song is.

These findings suggest that some individuals may be more inclined to avoid MVs for songs that they have already connected with on a personal level in order to avoid compromising that connection. Future research should consider individual variables, such as personality traits and use of music to regulate mood and emotional health, and their relationship to certain MV experience outcomes such as the negative influence outlined above.

MV Experiences: Attention and Reaction

The results of the Attention and Reaction levels shed light on the attentional processes that occur during MV experiences, and the affective outcomes evoked as a result. The data provide novel insights into how MV experiences contribute to participants' experience of the music, when this contribution has a positive influence on their perception of the music, and when it does not. Participants describe how specific features, such as narrative content and structural features in both modalities guide attention and elicit affective responses. However, these results are currently limited. Since reactions vary depending on the MV in question, it cannot be assumed that because an individual has a strong affective response to one MV that they experience similar reactions to all MVs. Individual differences in musical expertise, personality, and emotional reactivity to music may provide insight into who experiences strong affective responses to MVs and who does not.

Situational variables can also provide more insight into the conditions influencing affective responses to MVs: who they are with and how much control they had over the experience might explain when and why some individuals' responses are more salient than others. For example, individuals have more positive responses to self-selected music (Krause et al., 2015; North et al., 2004), music they are familiar with (Schubert et al., 2014) and in situations where they have control over what music is being listened to (Krause & North, 2017). Furthermore, since participants were asked to watch a MV they were *familiar with* and that they *knew* they enjoyed, the data do not provide any insight into the types of responses individuals experience the first time they see the MV. While the contingent factors highlighted above provide some insight into the features responsible for evoking affective responses and the data suggest that MVs can, indeed, evoke strong affective responses, future research using quantitative measures or

experimental designs would be better suited for understanding the affective phenomena that occur during MV experiences, whether it is the first time the participant is being exposed to the MV or a subsequent exposure.

Retention

Retention level data describes the carry-over effects of MVs on subsequent, audio-only listening experiences (see Figure 5). Two duration categories were established depending on whether the participant indicated MVs have a long-term influence on how they perceived the music going forward or not. Participants who did not experience any long-term influence on how the music was perceived or stated their personal interpretation of the music's content had a stronger influence than the MV were categorized as Unaffected. Overall, 73.5% of cases reported MVs having a long-term influence on their experience of the music while 26.5% of cases were categorized as Unaffected (see Table 5). In addition, three categories emerged describing how MVs influenced subsequent listening experiences. These categories were labelled New Interpretation of Meaning (NIM), New Affect Perception (NAP), and Visual Mental Imagery (VMI). Descriptions and examples of these categories are found in Table 6.

MVs had the potential to significantly influence the perception of the song's meaning in subsequent listens.

These experiences are categorized as NIM and were observed in 56.8% of cases. Themes in this category describe how the characters and narrative elements such as storylines become associated with the music in subsequent listens, as well as how the MV content clarified the meaning behind the lyrics. For MVs containing culturally topical or socio-political messages, these messages became associated with the music in subsequent listens, therefore influencing the listener's interpretation of the song.

MV experiences could influence how the affective quality of the music was perceived. These outcomes, categorized as NAP, were reported in 17.6% of cases. Codes in this category describe how the MV changed how the affective quality of the music was perceived, making the song more emotionally impactful and potentially changing how the listener used the song for affect regulation purposes in subsequent listens. There was a significant overlap between cases reporting NIM and NAP effects, which suggests that the meaning of the music portrayed in the MV had an influence on their perception of the music's emotional quality as a result of it providing them with more context about the song.

The last Retention category, VMI, occurred the most frequently: it was observed in 76% of cases. This category reflects how imagery from the MV is recalled in subsequent listening experiences, even in cases where no long-term
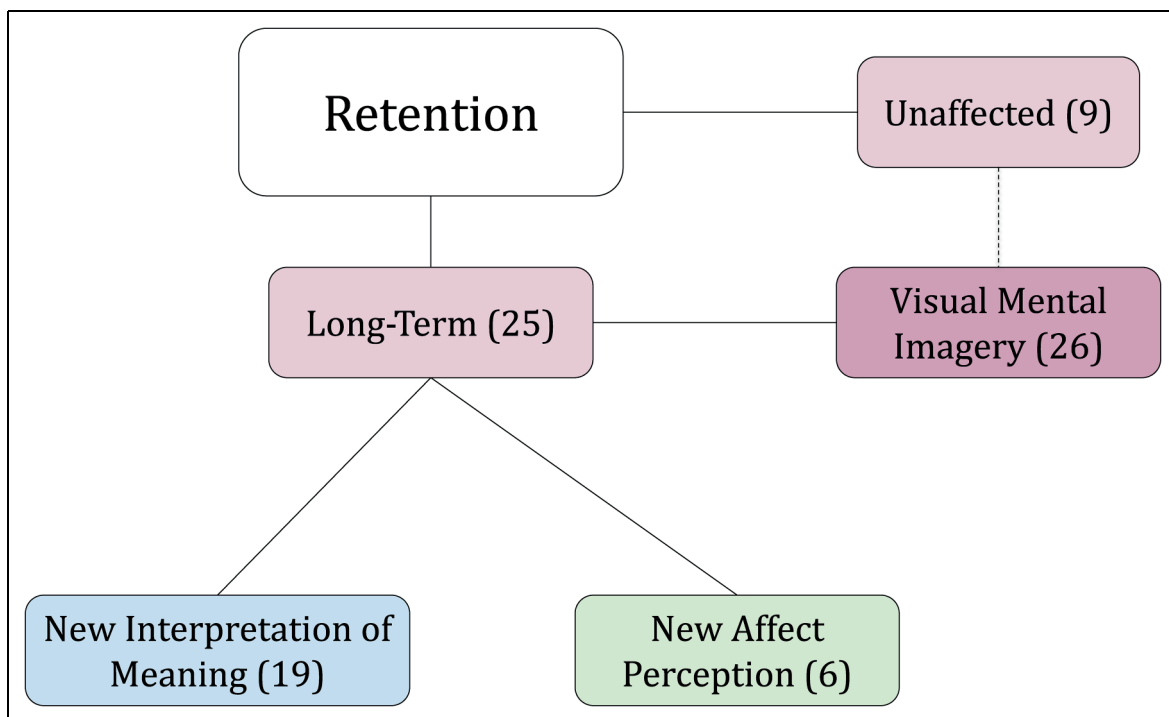
Figure 5. Retention outcomes are the effects which carry-over to subsequent listening experiences, as well as duration (Long-Term or Unaffected). Long term outcomes were divided into two categories. A third category, Visual Mental Imagery, was observed the most frequently, even in participants who stated the MV had no influence on their interpretation of the music or its affective quality. It is connected to the Unaffected category with a dotted line. The number of participants coded under each category are represented in brackets.

Table 5. Description of duration categories and examples.

Duration	Description	Example (participant number)
Long-term 73.5%	• Information from the MV influences subsequent listening episodes. • No longer retains previous interpretation if the MV contradicts it.	"As soon as I watch the video, I only imagine that once I hear the music again." (P02) "A good video will definitely be what I think of every time I hear [the] song permanently afterwards." (P15)
Unaffected 26.5% of cases	• MV does not have any influence on their perception of the music in subsequent listens, their personal interpretation is retained. • May experience a short-term change in perception but only during the MV experience.	"If I've already connected to a song, finding out the 'true meaning' intended by the artist doesnt change my interpretation" [sic]. (P12) "From memory, I can't think of a specific instance, but I feel like it probably has happened, more on a short-term basis [while watching]." (P33)

Table 6. Descriptions of retention categories and examples.

Category	Description/Themes	Example (participant number)
New Interpretation of Meaning (NIM) 56.8% of cases	• Interprets the meaning of the music based on the MV's content • Interprets the lyrics based on the storyline depicted in the MV • Associates the song with the social, cultural, or political themes depicted in the MV	"I always prefer my interpretation of a song but it can be hard to still think of it that way after seeing a video that shows it's meant to be interpreted differently." (P03) "The content of a music video has affected my interpretation of a song's meaning when there are actors who play a storyline based on what the lyrics are about." (P18) "[This song] was not a song I associated with [this social movement] at all, until I saw the music video, and now I think of it whenever I hear the song." (P07)
New Affect Perception (NAP) 17.6% of cases	• New perception of the music's emotional quality • Song more emotionally impactful compared to emotional responses prior to seeing the MV • May occur as a result of NIM, where the new interpretation creates a different affective response to the music in future listens.	"The imagery can [influence] how I perceive the mood of the song." (P32) "For example, sometimes a song will have a music video with a completely different tone/feel than the situations/context in which I would normally listen to the song (i.e., a happy song that I listen to when I'm in a good mood having a sad music video). So sometimes this makes me think more about the lyrics and hear things that I wouldn't have prior to seeing the music video." (P01) "I have thought a song was boring, with very little depth, then with a video's context it has completely changed the song forever for me." (P27)
Visual Mental Imagery (VMI) 76% of cases	• Song triggers mental imagery related to the content of the MV, such as narrative elements, thematic content, or characters • Imagery of performance gestures, dance choreography, and human movement.	"For example, in the video I just watched, I always think of [this very powerful scene] when the singer is screaming... I find it makes an already powerful moment in the song even more intense."(P16) "Especially in very sad or [emotional] songs I seem to always remember the video while listening to the song later."(P10) "These music videos often have iconic dance moves to them, so when just listening to music, I can see the dance in my mind when I listen to the music alone." (P21)

effects were reported concerning the perception of the music's meaning or affective quality. Themes for VMI codes include remembering images related to the content of the video, such as characters, storylines, topics, performance gestures, and dance choreography throughout the listening experience. Interestingly, 46% of all VMI

descriptions (35% of all cases in total) describe VMI of human gestures, such as the artist performing and scenes with dance and other choreographed movements (for an example, see Table 6).

The most interesting finding concerning MV-related VMI is that it could occur regardless of whether the content of the MV influenced their perception of the music's meaning or affective quality. Overall, two thirds (66.6%) of cases categorized as Unaffected reported VMI of MV content in subsequent listens. This may be a result of joint encoding: when the audio and visual are perceived as emotionally congruent, an integrated memory code is formed (Boltz, 2004). As a result, subsequent listens become a retrieval cue for MV imagery. Furthermore, these findings suggest that BRECVEMA mechanisms, particularly visual imagery, can change over time as new associations with the music are formed.

In addition, our findings highlight how MVs can make the music personally significant to the listener depending on whether the MV confirms or violates their expectations about the meaning of the music. Importantly, the MV did not need to confirm their expectations to be perceived as more meaningful (however, this may have had an influence on how much they empathize with the artist); however, it did seem important that the MV's content be in line with the listener's personal values. For example, MVs could have a negative long-term influence on the listener's perception of the song if it were perceived as representing antisocial or negative behaviours, such as glorifying violence or the sexual objectification of women. For example, as one participant explains (P15): "…he exploits her by making a nude model of her which I thought was kind of sleazy even though before that I enjoyed the song". Negative effects from MVs were observed in 18% of cases, and while some cases described disliking the MV because of it depicting imagery that was incompatible with their values, MVs also had a negative influence in cases where the MV showed an interpretation of the music that was not in line with their own interpretation of the song. For example, one participant explains how the content of the MV was perceived as being superficial and jarring, stating (P25): "I've been basically trying to erase the video from my brain so I can love the song the way I did before. This is not always possible." While the majority of participants (58%) provided examples of MVs that had a long-term positive influence on their perception of the music, this may have been a result of the study design, which asked participants to watch an MV they were already familiar with and enjoyed prior to completing the questionnaire. For whom and under what conditions the MV elicits positive or negative carry-over effects should be examined in future studies, and individual differences in emotional reactivity to music, personality traits, and music use for affect regulation purposes, should be measured. For example, individuals who are more emotionally stable and conscientious may be more likely to be unaffected by content they do not enjoy, even if the MV depicts an interpretation of the music that is contrary to their own personal interpretation.

Conclusions

The IARR framework provides new insight on the key characteristics of MV experiences.

- Intentions describe motivations for watching MVs. Internal goals include emotional or cognitive needs, such as to regulate affective states or to learn more about the meaning behind or production of the music, whereas External goals reflect experiences motivated by other people and the media.
- MV experiences absorb attention and influence affective reactions through their use of aesthetic imagery. While storylines and narrative content (Semantic features) were frequently mentioned, participants were also attending to how well the visuals and music complemented each other (Structural features), even in the absence of a narrative.
- MVs can evoke strong affective reactions and promote mechanisms such as emotional contagion. However, MVs can also distract from or hinder listener enjoyment of the music in the event they are perceived as being poorly made or violate the listener's expectations about the meaning of the music.
- The Retention level highlights how MV content is remembered and associated with the music in future listens, changing how the music's meaning and affective quality are perceived. In addition, participants reported that images from the MV were recalled, seemingly automatically, as visual mental imagery in subsequent listens, even if no other change in perception occurred.

This study found that MVs can enhance enjoyment of the music when the video features elements that give the song more depth, such as narrative components or imagery which make its emotional tone more salient. However, MVs can also hinder the experience of the music in the cases where the individual's personal interpretation of the music is incongruent with what is depicted in the video. The question of whether MVs enhance or hinder the perception and enjoyment of music, which was previously raised by Boltz (2013), does not have a simple answer: it depends on a range of personal, musical, and visual factors which could not be controlled in the present study.

The IARR framework complements previous research concerning the psychological function of music and the cognitive processes that occur when music is paired with visual information in a modern and popular listening context. It provides novel insight into how malleable and nuanced music listening outcomes can be, and the factors which contribute to this fluctuation. The study shows how new associations made possible by extra-musical sources such as MVs

influence the perception of musical meaning, as well as how affective outcomes to music may change as a result. Our results are also in line with Sun and Lull's study (1986). Despite these studies being conducted three decades apart from each other, the reasons for watching MVs and the perceptual effects they impose have remained consistent. We believe that future research needs to consider MVs and other visual presentations of music, especially considering the new methods of music listening currently available, to better understand how listening outcomes change over time.

While the IARR framework and its findings are novel, it is also a synthesis of existing models; it highlights the ways in which current models of multimedia and music perception complement each other in the context of MV experiences. For example, CAM provides a theoretical understanding of how music influences the interpretation of film, including how music directs attention to features in the video, and the interactions that occur as information stored in the viewer-listener's long-term memory interacts with working memory when processing the events in multimedia (Cohen, 2005). In the context of multimedia consumed as part of personal music listening experiences – including MV experiences – this interaction is especially nuanced, since the individual may have already established their own associations with the music, particularly in the case where the MV is for a favourite song or by a preferred artist. These findings are also in line with Cook's (1998) contest model of musical multimedia, which emphasizes how media is received by the viewer. The component media within the MV all contain their own sources of meaning; they are, as Cook suggests, "vying for the same terrain, each attempting to impose its own characteristics upon the other" (1998, p. 103). This contest is even evident at the Intention level, since it was frequently stated that understanding the meaning of the music was a significant factor motivating the individual to watch the MV in the first place. On the other hand, individuals who did not want to have their personal associations or perceptions of the music's meaning deconstructed stated they would avoid MVs for songs that were deeply personal to avoid conflicting meanings. Contest, however, is not the only model that applies to MV experiences, as many cases provided insight into how the music and the video complemented each other, and the impact this had on their reactions. According to one participant, the video (P16): "makes an already powerful moment in the song even more intense". In cases where the meaning is in line with the individual's associated meaning, this complementation can result in a salient affective reaction, as another participant describes (P25): "I cried at a lot of parts while watching it. Everything about it was SO in line with the things I'd already felt when listening to the music alone that it enhanced it so much beyond the sum of its parts [sic]." However, if the MV provides context about the meaning of the music that is not in line with the listener's initial personal interpretation or associations with the music, a negative affective response may occur. Our results suggest that when an MV is not in line with the listener's initial interpretation of the music's

meaning, the song can potentially become less impactful or even ruined for them in future listens as they try to "erase" the MV from their brain. However, as one participant asserts (P25): "This is not always possible." On the other hand, if the MV is not in accordance with their original interpretation but depicts content that the individual finds impactful or profound, the MV can have positive consequences on future listening experiences.

The increase in positive associations is in line with spreading activation theory, which posits that aesthetic pleasure from music listening occurs as a network of mental representations associated with the music become activated (Schubert et al., 2014). Mental representations are also responsible for the mechanisms behind music-evoked emotions (Juslin, 2013), and while research has studied intersections between BRECVEMA mechanisms and spreading activation (see Völker, 2021), more research should be done which considers how different theoretical models intersect or complement each other across diverse listening contexts.

While the results are novel and interesting, there are limitations to address. The online questionnaire design meant that the data were limited to the amount of detail provided by the participants: since the researcher was not present while the participant filled out the questionnaire, it was not possible to follow-up or ask the participant to elaborate on their descriptions. Some participants provided longer and more detailed descriptions than others. The results may have also been biased towards positive experiences with MVs given that they were asked to watch an MV they were familiar with and enjoyed prior to filling out the questionnaire. While some participants may have still provided data concerning their experiences with MVs they *did not* enjoy, this was not the case for all participants. Furthermore, we did not investigate the role of individual differences, such as musical engagement, preferences or background in the present analysis, and the small sample size was unsuitable for quantitative analysis. In addition, since some participants provided details about their experience with more than one MV (and not just the one they watched prior to doing the questionnaire), it was not possible to reliably ascertain whether there was an association with specific Intention factors leading to certain experiences or Retention outcomes.

The IARR framework is meant to provide a starting point for future research on the topic by providing insight into what variables or phenomena need to be considered and accounted for in study designs examining MV experiences and their effects. Furthermore, future studies should consider the role of individual differences in personality, emotional health and musical engagement behaviour in order to establish when and for whom MV experiences elicit what kinds of outcomes, and under what conditions. The use of more quantitative methodologies, such as surveys or experimental designs using control measures, can be used to explore relationships between categories at each temporal level and the potential relationships between experience patterns, MV

listening outcomes, and individual differences. In addition, a larger sample which includes participants from more diverse backgrounds is necessary.

While not every song on an album has an MV, the importance of visuals and their influence on listening outcomes is not limited to these types of experiences. For example, Spotify's Canvas feature allows artists to upload 3- to 8-s videos, including clips from MVs, that loop while a song is playing, and full MVs are available for Premium account holders in many countries, highlighting the industry's push to make visual content more available for music consumers. These platforms are also starting to include videos with the lyrics of the song, another media component which may influence the perception of the music. Music psychology research needs to consider these current trends and their influence on music listening behaviour, affective outcomes, and the perception of music in general. A continuous effort needs to be made in updating or expanding relevant theoretical models to reflect modern listening trends and their impact on music's psychological functions and overlaps between theoretical models need to be considered and explored.

Action Editor

Isabel Martínez, Universidad Nacional de La Plata, Facultad de Bellas Artes.

Peer Review

David Ireland, University of Leeds, School of Music.
Annabel Cohen, University of Prince Edward Island, Department of Psychology.

Contributorship

JDW researched literature, JDW and SS conceived and designed the study. JDW wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version.

Data Availability

Anonymized versions of the data may be obtained by request by contacting the corresponding author.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

ORCID iDs

Johanna N. Dasovich-Wilson https://orcid.org/0000-0003-1130-5507
Suvi Saarikallio https://orcid.org/0000-0002-4647-8048

Note

1. When MTV first debuted in the United States in 1981, MVs were aired 24 hours a day, seven days a week, until the channel was purchased by Viacom Inc. in 1985. MVs were still aired on MTV, its sister channels (such as MTV2), and affiliates (for example, MuchMusic in Canada). Although changes were made to the channel, it was still primarily MVs and other musical content (such as live performances) at the time of Sun and Lull's study (see Edmond, 2014 for an overview of MTV and its content during the 1980s and 1990s).

References

Baltazar, M., & Saarikallio, S. (2016). Toward a better understanding and conceptualization of affect self-regulation through music: A critical, integrative literature review. *Psychology of Music, 44*(6), 1500–1521. https://doi.org/10.1177/030573561 6663313

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2

Barton, K. C. (2015). Elicitation techniques: Getting people to talk about ideas they don't usually talk about. *Theory & Research in Social Education, 43*(2), 179–205. https://doi.org/10.1080/00933104.2015.1034392

Boltz, M. (2013). Music videos and visual influences on music perception and appreciation: Should you want your MTV. In S. Tan, A. J. Cohen, S. Lipscomb, & R. Kendall (Eds.), *The psychology of music in multimedia* (pp. 217–234). Oxford University Press.

Boltz, M., Ebendorf, B., & Field, B. (2009). Audiovisual interactions: The impact of visual information on music perception and memory. *Music Perception, 27*(1), 43–59. https://doi.org/10.1525/mp.2009.27.1.43

Boltz, M. G. (2004). The cognitive processing of film and musical soundtracks. *Memory & Cognition, 32*, 1194–1205. https://doi.org/10.3758/BF03196892

Cespedes-Guevara, J., & Eerola, T. (2018). Music communicates affects, not basic emotions A constructionist account of attribution of emotional meanings to music. *Frontiers in Psychology, 9*, 215. https://doi.org/10.3389/fpsyg.2018.00215

Chamorro-Premuzic, T., & Furnham, A. (2007). Personality and music: Can traits explain how people use music in everyday life? *British Journal of Psychology, 98*(2), 175–185. https://doi.org/10.1348/000712606X111177

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis (introducing qualitative methods series)*. Sage.

Charmaz, K. (2008). Grounded theory as an emergent method. In S. N. Heese-Biber & P. Leavy (Eds.), *Handbook of emergent methods* (pp. 155–172). The Guilford Press.

Chin, T–C, & Rickard, N. S. (2014). Emotion regulation strategy mediates both positive and negative relationships between music uses and well-being. *Psychology of Music, 42*(5), 692–713. https://doi.org/10.1177/0305735613489916

Cohen, A. J. (2001). Music as a source of emotion in film. In P. Juslin & J. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 249–272). Oxford University Press.

Cohen, A. J. (2005). The interpretation of film and video: Approaches from experimental psychology. *Selected Reports in Ethnomusicology, 12*, 15–36.

Cohen, A. J. (2013). Congruence-Association model of music and multimedia: Origin and evolution. In S. Tan, A. J. Cohen, S. Lipscomb, & R. Kendall (Eds.), *The psychology of music in multimedia* (pp. 17–47). Oxford University Press.

Cook, N. (1998). *Analysing musical multimedia.* Oxford University Press Inc.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music, 21*(2), 103–113. https://doi.org/10.1177/030573569302100201

Edmond, M. (2014). Here we go again: Music videos after YouTube. *Television & New Media, 15*(4), 305–320. https://doi.org/10.1177/1527476412465901

Goldberg, M. E., Chattopadhyay, A., Gorn, G. J., & Rosenblatt, J. (1993). Music, music videos, and wear out. *Psychology & Marketing, 10*(1), 1–13. https://doi.org/10.1002/mar.4220100102

Greasley, A. E., & Lamont, A. (2011). Exploring engagement with music in everyday life using experience sampling methodology. *Musicae Scientiae, 15*(1), 45–71. https://doi.org/10.1177/1029864910393417

Greb, F., Scholtz, W., & Steffens, J. (2018). Personal and situational influences on the functions of music listening. *Psychology of Music, 46*(6), 763–794. https://doi.org/10.1177/0305735617724883

Hargreaves, D. J., & North, A. C. (1999). The functions of music in everyday life: Redefining the social in music psychology. *Psychology of Music, 27*(1), 71–83. https://doi.org/10.1177/0305735699271007

International Federation of the Phonographic Industry (2019). *Music Consumer Insight Report.* IFPI. https://www.ifpi.org/ifpi-releases-music-listening-2019/.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews, 10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Juslin, P. N., Harmat, L., & Eerola, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music, 42*(4), 599–623. https://doi.org/10.1177/0305735613484548.

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion, 8*(5), 668–683. https://doi.org/10.1037/a0013505.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*(5). https://doi.org/10.1017/S0140525X08005293

Krause, A. E., & North, A. C. (2017). Pleasure, arousal, dominance, and judgments about music in every- day life. *Psychology of Music, 45*(3), 355–374. https://doi.org/10.1177/0305735616664214

Krause, A. E., North, A. C., & Hewitt, L. Y. (2015). Music-listening in everyday life: Devices and choice. *Psychology of Music, 43*(2), 155–170. https://doi.org/10.1177/0305735613496860

Laiho, S. (2004). The psychological functions of music in adolescence. *Nordic Journal of Music Therapy, 13*(1), 47–63. https://doi.org/10.1080/08098130409478097.

Liljestrom, S., Juslin, P. N., & Västfjäll, D. (2012). Experimental evidence of the roles of music choice, social context, and listener personality in emotional reactions to music. *Psychology of Music, 41*(5), 579–599. https://doi.org/10.1177/0305735612440615

Marshall, S. K., & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception: An Interdisciplinary Journal, 6*(1), 95–112. https://doi.org/10.2307/40285417

North, A. C., Hargreaves, D. J., & Hargreaves, J. J. (2004). Uses of music in everyday life. *Music Perception, 22*(1), 41–77. https://doi.org/10.1525/mp.2004.22.1.41

Randall, W. M., & Rickard, N. S. (2017). Personal music listening: A model of emotional outcomes through mobile experience sampling. *Music Perception, 34*(5), 501–514. https://doi.org/10.1525/mp.2017.34.5.501

Rentfrow, P. J. (2012). The role of music in everyday life: Current directions in the social psychology of music. *Social and Personality Psychology Compass, 6*(5), 402–416. https://doi.org/10.1111/j.1751-9004.2012.00434.x

Saarikallio, S. (2010). Music as emotional self-regulation throughout adulthood. *Psychology of Music, 39*(3), 307–327. https://doi.org/10.1177/0305735610374894

Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. *Psychology of Music, 35*(1), 88–109. https://doi.org/10.1177/0305735607068889

Saarikallio, S., Gold, C., & McFerran, K. (2015). Development and validation of the healthy unhealthy music scale. *Child and Adolescent Mental Health, 20*(4), 210–217. https://doi.org/10.1111/camh.12109

Saldaña, J. (2016). *The coding manual for qualitative researchers (No. 14).* Sage.

Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Frontiers in Psychology, 4*, 511. https://doi.org/10.3389/fpsyg.2013.00511

Schubert, E., Hargreaves, D. J., & North, A. C. (2014). A dynamically minimalist cognitive explanation of musical preference: Is familiarity everything? *Frontiers in Psychology, 5*, Article 38. https://doi.org/10.3389/fpsyg.2014.00038.

Sloboda, J. A., O'Neill, S. A., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the experience sampling method. *Musicae Scientiae, 5*(1), 9–32. https://doi.org/10.1177/102986490100500102

Sun, S.-W., & Lull, J. (1986). The adolescent audience for music videos and why they watch. *Journal of Communication, 36*, 115–125. https://doi.org/10.1111/j.1460-2466.1986.tb03043.x

Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 370–381. https://doi.org/10.1037/0278-7393.30.2.370

Tarrant, M., North, A. C., & Hargreaves, D. J. (2000). English And American adolescents' reasons for listening to music. *Psychology of Music, 28*(2), 166–173. https://doi.org/10.1177/0305735600282005

Timmermans, S., & Tavory, I. (2012). Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory, 30*(3), 167–186. https://doi.org/10.1177/0735275112457914

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition, 118*(2), 157–170. https://doi.org/10.1016/j.cognition.2010.11.010

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition, 101*(1), 80–113. https://doi.org/10.1016/j.cognition.2005.09.003

Völker, J. (2021). Measuring emotional music experience: Spreading activation and BRECVEMA mechanisms. *Psychology of Music.* https://doi.org/10.1177/03057356211010224

Vuoskoski, J., & Eerola, T. (2013). Extramusical information contributes to emotions induced by music. *Psychology of Music, 2*(43), 262–272. https://doi.org/10.1177/0305735613502373.

Wilson, J. N. (2018). *Young people's emotional engagement with music videos: An exploratory study.* [Master's thesis, University of Jyväskylä]. http://urn.fi/URN:NBN:fi:jyu-201808293953.

Wilson, J. N., Thompson, M., & Saarikallio, S. (2020). The function of music videos in everyday listening experiences. *Proceedings for Psychology and Music: Interdisciplinary Encounters*, October 24-26, 2019. Belgrade, Serbia. 135-144.

YouTube (2021, April 21). *Most Viewed Videos of All Time (Over 450M views).* YouTube. Accessed June 1, 2021. https://www.youtube.com/playlist?list=PLirAqAtl_h2r5g8xGajEwdXd3x1sZh8hC.

Appendix

1. Approximately how many music videos do you watch on average every week?
2. In what kind of environment, situation, or circumstances would you choose to watch a music video over listening to the music alone?
3. Do you ever re-watch music videos that you enjoy? If so, how frequently?
4. Which device do you usually watch music videos on? (You may select more than one answer for this question)
5. Laptop/computer
6. Tablet (e.g., iPad)
7. Smartphone
8. All of the above
9. Other (please specify)

The following questions are aimed at better understanding how music videos affect your experience with the music. You may refer to any music video in your responses; you are not limited to the video you watched before starting the questionnaire. Please be as descriptive and honest as you can.

1. Do you think the video enhances the music listening experience? Why or why not?
2. Describe your thought processes when watching music videos. Do you have similar thoughts when listening to the music alone? If not, how do they differ?
3. Has the content of a music video ever affected or changed your interpretation of a song's meaning? If yes, is it:
   Short term change (only while watching the video?)
   Long term change (every time you listen to the song)? or somewhere in between? Please describe.
4. Do certain scenes from the music video come to mind when you are listening to the music on its own? If yes, please give an example?
5. What kind of emotional outcomes or changes in mood do you experience when you watch music videos?
6. Is your emotional reaction to the music greater when you watch music videos? Or is your emotional reaction greater when you're only listening (not watching the video)?
7. Have music videos influenced your perception of or feelings towards the artist? If so, how?
8. Do you believe music videos have the ability to influence behaviour, whether in yourself or others? Please explain.
9. In your opinion, what makes a good music video?
10. Please include any other thoughts, opinions or feelings that you'd like to add that were not covered in the questionnaire. You may also include any feedback you have about the study here.
11. Please enter your age (in years)
12. What is your nationality?
13. Please indicate your gender

(Exit page). Thank you for participating in our study!

STUDY II

THE FUNCTION OF MUSIC VIDEOS IN EVERYDAY LISTENING EXPERIENCES

by

The Function of Music Videos in Everyday Listening Experiences

Johanna N. Wilson,[1] Marc Thompson,[2] and Suvi Saarikallio[3]

[1,2,3] *Department of Music, Arts and Culture Studies, University of Jyväskylä, Finland*

[1]johanna.n.wilson@jyu.fi, [2]marc.thompson@jyu.fi, [3]suvi.saarikallio@jyu.fi

Abstract

Streaming music videos (MVs) on social media platforms such as YouTube continues to be a popular method of music listening (see International Federation of the Phonographic Industry, 2018). Empirical research has shown that pairing music with visual stimuli has a significant influence on how the music is perceived and remembered in future listening episodes (Boltz, Ebendorf, & Field, 2009). This suggests MVs experiences may influence subsequent listening episodes, even when the video is no longer present. The aim of the study is to investigate MV watching experiences, particularly addressing the reasons for choosing this method of listening, the cognitive and emotional processes that occur during the experience, and the potential carry-over effects in subsequent listening episodes. Qualitative questionnaire data were collected and personality traits and the use of music for emotional health were measured using the TIPI and HUMS scales. Qualitative analysis was directed at identifying significant themes and causal relationships in the qualitative data. This analysis revealed that participants often experience changes in how they perceive the music's meaning after watching MVs and that future listening episodes would trigger visual mental imagery from the MV. The quantitative analysis revealed that participants with low scores in the trait emotional stability were more likely to experience negative long-term changes in how they perceived the music in subsequent listening episodes. Furthermore, participants with high scores in an unhealthy listening style were more likely to experience long-term effects from the MVs in general. The study identified several characteristics of MV listening experiences, providing new knowledge about this modern listening context. The results elaborate on our understanding of the personal meanings and affective impacts of everyday music listening.

Introduction

The way we listen to music is changing. Smartphones and other portable devices have allowed us to bring music with us virtually anywhere, and streaming services have replaced vinyl records and CDs due to their accessibility and massive music libraries. According to the Music Consumer Insight Report (see International Federation of the Phonographic Industry, 2018) YouTube is currently the most popular method of streaming music online. This suggests that individuals are more likely to consume music as multimedia such as music videos (MVs) as opposed to audio-only formats. Research has shown that music has an important role in young people's psychological and social development (Laiho, 2004; Miranda, 2013). The current study analyses the reasons, emotional outcomes and carry-over effects of this method of music listening in youth audiences and their relationship to certain traits such as personality and use of music for emotional help in order to establish a theoretical framework for understanding these experiences.

MVs differentiate themselves from other forms of music listening due to the fact they are multimedia. Empirical studies which examine the multi-modal component of music listening do exist, however, this previous research has been limited to specific elements such as performance gesture (Davidson, 1993; Vines, Krumhansl, Wanderley, Dalca, & Levitin, 2011) or as they pertain to film music (Boltz, 2004; Cohen, 2001, 2013; Marshall & Cohen, 1988). From a psychological perspective, MVs have mostly been studied in respect to how they influence behaviour in youth (Sun & Lull, 1986) and frequently focus on specific demographics (e.g., Bryant, 2008). Much has changed about MVs since they first appeared on MTV over thirty years ago, for example: they have helped shift the music industry towards a number of online streams as opposed to record sales, having evolved from a means to promote singles to a source of revenue for record companies (Ed-

mond, 2014). Furthermore, their availability on YouTube has allowed audiences to control *which* MVs they watch, wherever and whenever they want. The element of control has been linked to positive outcomes such as contentment and motivation, especially on platforms that allow for personal music collections and curation (Krause, North, & Hewitt, 2015).

Music listening can be used as a tool for regulating affect by maintaining positive mood states or changing negative ones. The GTSM model proposed by van Goethem and Sloboda (2011) highlights how music listening activities succeed in achieving regulation goals by using music listening as a tactic to engage in affect regulating strategies such as mental work or solace, which enable mechanisms such as evoking memories and mental imagery. Another important component to music listening outcomes are individual differences such as gender, age, and personality (Saarikallio, Nieminen, & Brattico, 2013; Vuoskoski & Eerola, 2011), listening style (Chamorro-Premuzic & Furnham, 2007), and use of music for achieving emotion regulation goals (Chin & Rickard, 2014; Saarikallio, Gold, & McFerran, 2015).

MVs are a unique method of listening that can further our current understanding of everyday music use for affect regulation. These listening contexts have been relatively underexplored, despite their popularity and modern role in shaping the music industry. The visual component of MV listening may be an integral source of information for young viewers, who use them as a method of social learning (Hansen & Hansen, 2000), as well for exploring their "private self" (Larson, 1995). This multimodal listening experience may have a significant impact on how the music is perceived, both during the watching episode and in subsequent audioonly listens. As a result, MVs have the potential to influence the way the listener uses the music for affect regulation purposes in the future.

The Current Study

The study aims to create a framework for analyzing MV listening experiences and their relationship to music listening for emotional health purposes. The study uses questionnaire data and scale measures (TIPI and HUMS) to achieve this. The study uses data from a larger qualitative analysis which aims to create a theoretical model for understanding the process of MV watching over four temporal stages, referred to as: *Intention, Attention, Reaction* and *Retention* (IARR) (Wilson, Thompson, & Saarikallio, in review). In this framework, *Intention* refers to the reasons an individual chose to watch the MV and what type of goals does aim to fulfil. *Attention* and *Reaction* make up the experience itself; *Attention* themes provide insight into which components of the audio and visual modalities were the object of focus and whether the MV distract them from enjoying the music, and *Reaction* themes provide insight concerning how and why MVs elicit an emotional reaction. The final category, *Retention*, describes the ways in which subsequent listening episodes are affected by the MV's visual information and the duration of these effects.

Select subcategories from the IARR model were chosen for the current study's convergent analysis on the grounds they provide reliable insight into the ways in which MVs complement or hinder the use of music to promote emotional health and well-being. These subcategories describe participants' listening goals, the emotional outcomes they experience, and the
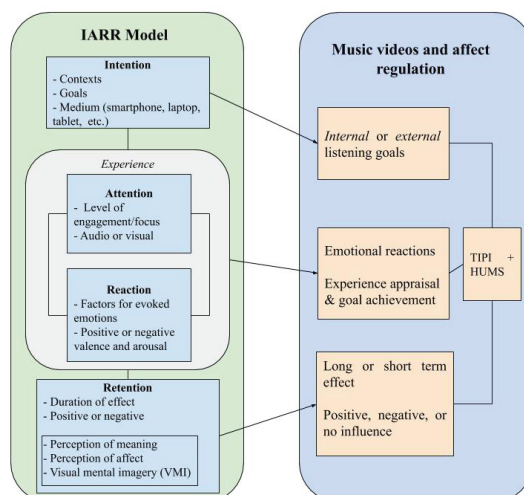

Figure 1. The IARR model (Wilson et al., in review) and the outline of the current study.

potential carry-over effects on future listening episodes. To establish a preliminary framework, these subcategories were transformed into nominal data; this allows for the triangulation of both qualitative and quantitative data sets in order to establish relationships between them. The IARR model and the current study's framework are visualized in Figure 1.

Method

Participants

Data were collected from 34 ($n = 34$) participants between the ages of 15 and 27 ($M = 22.4$). Participants were recruited via social media (Twitter and Facebook). In order to provide an incentive for their participation, all participants who provided a valid email address were entered into a raffle for an Amazon gift card. Participants were encouraged to invite friends to participate. For every friend that participated the original subject had an extra ballot entered for the raffle.

Design and Procedure

The study was accessible online via Qualtrics and took approximately 20–30 minutes to complete. Participants were first asked to complete the HUMS and TIPI measures before they completed the open-ended questionnaire component. Participants were asked to watch a MV they had already seen in order to limit recall bias and to include the song title and artist name or the MVs YouTube link. They were also informed that they may also refer to other MVs in their answers as well. In addition to the selected scale measures, four additional Likert-scale questions were added. Two of these questions were implemented in order to measure how much attention the subject attended to the MV, and whether they found themselves focused on the music or the visual component. The other two Likert questions pertained to how much the MV affected their mood and in what direction (positive or negative). The open-ended questionnaire component was designed in order to allow for detailed responses. Participants were encouraged to provide personal opinions and other details at the end of the study.

Qualitative Data and Analysis

The qualitative analysis was directed at uncovering relationships between patterns of experience and emotional health and well-being. An abductive method was chosen for qualitative analysis; this approach encourages the researcher to revisit the phenomenon in question after exploring and reframing the data from the perspective of existing theoretical frameworks (Timmermans & Tavory, 2012). This was deemed an appropriate method for analysis since individual concepts of MV watching can be understood in light of existing theories on the importance of music during youth and its role in psychological development, the use of music for affect regulation, media use, and gratification, and the cognitive encoding of audio-visual material.

Quantitative Data and Triangulation

The quantitative analysis used two-scale measures: the first measured personality (TIPI; Gosling, Rentfrow, & Swann, 2003) and participants' use of music for emotional health and well-being (HUMS; Saarikallio et al., 2015). Independent sample t-tests are conducted on the transformed nominal data from the qualitative analysis and individual variables of interest: personality (TIPI), specifically emotional stability, and healthy or unhealthy use of music for emotion regulation purposes (HUMS).

Results

Qualitative Findings

Reasons for engagement. Two main categories were established in the data: *internal* and *external* goals. *Internal* goals are motivated by the individual's psychological needs and are further subcategorized into *emotional* and *reflective* reasons. Emotional reasons include reports of the subject watching in order to achieve some affect related goal, such as relieving boredom or to change their negative emotional state into a positive one. Reflective reasons consist of watching the MV in order to better understand the meaning or interpretation of the music. In-

dividuals who watched the MV in order to learn how to play the music or to learn dance choreography were also included in this subcategory. *External* goals are influenced by social factors. These social factors include watching MVs that were shared with them by friends, or less personal social influences such as watching MVs that was hyped in the media or recommended to them by YouTube. *External* goals also include using MVs in order to relate to their favorite

media characters, a phenomenon Kristler, Rodgers, Power, Austin, and Hill (2010) referred to as "parasocial interactions". This form of social bonding occurs when individuals develop a perceived relationship with their favorite media characters, including musicians. Overall, goals motivated by *external* goals were reported in 24 cases (70.5%). *Internal* goals were reported in 23 cases (67.6%). Descriptions and frequencies are also reported in Table 1. Since some participants provided more data than others, it was possible for some individuals to provide codes in both categories.

Experience outcomes. Many participants asserted that whether their experience had any influence on their affective state was contingent on at least one of three key factors: the music, the video, and their personal interpretation of the music. These contingent factors are found in Table 2.

Table 1. Categories for intentions and goals of MV watching, descriptions and frequencies.

Intentions and goals		
Category & Frequency	Themes	Example codes
Internal goals 23 cases 67.6%	Emotional: To change or maintain the emotional state	When I am relaxing at night or having an anxiety attack. (P34) During my downtime when I'm procrastinating … (P36)
	Reflective: To create new associations with the music. To learn how to perform the music or MV choreography.	I love seeing the music videos to songs I love to see how the lyrics combine into visual art. (P24) … or when I am researching/ learning a song (P36)
External goals 24 cases 70.5%	MV received media hype or is shared by friends For parasocial engagement	Normally I'd only watch the video [if] someone recommended it or if I heard someone mention something special about it. (P10) I think it helps connect to the artist more and to understand what the song means to them. (P12)

Table 2. Contingent factors for emotions evoked by MVs.

Factors	Contingencies
Musical	Genre, emotional quality, artist, familiarity
Visual	Portrayal of emotion, presence of storyline or performance gestures, quality of the cinematography
Personal	Appraisal of the music (prior to watching MV), current affective state, whether the MVs reflects their personal interpretation of the music

Participants were categorized as having experienced *strong affect* if they disclosed intense emotional reactions in response to the MV. Individuals who did not experience any significant emotional outcomes or changes in mood were categorized as *unimpacted*. These two categories are mutually exclusive: participants could not provide evidence of *strong affect* and also be *unimpacted*. However, some participants are not categorized under either due to lack of data or the number of contingencies they report having an effect on their reaction. These categories and their frequencies are described in Table 3.

Table 3. Experience reaction categories, descriptions and frequencies.

Category & Frequency	Themes	Example codes
Strong affect 16 cases 47%	Experiences strong emotional outcomes during MV episode	I find it makes an already powerful moment in the song even more intense (P16)
	Reports that the video makes the emotional quality more salient.	… in general, I feel happier and more upbeat after I watch music videos (P19).
Unimpacted 5 cases 14.7%	No salient emotional outcomes reported.	[My emotional outcomes are] usually not significant. (P22)
	Reports that the MV does not elicit the same emotions as audio-only.	Videos don't tend to alter my mood. The music is an important part for me. (P12)

Table 4. Duration and valence of MV effects.

Duration	Valence & Frequency	Example codes
Long term	Positive 58%	*The [MV] is really powerful and had a long-lasting impression on me. Now, every time I hear the song, I am reminded of the video and the message [the artist] was trying to convey.* (P19)
	Negative 18%	*I've been basically trying to erase the video from my brain so I can love the song the way I did before…* (P25)
Unaffected	Neutral 24%	*Depends on how in depth the story is but mostly the song means what the song means to me despite what the video says.* (P35)

Duration of effects. Participants asserted that MVs had the potential to change the way they perceived or understood the meaning of the music in the future; however, these were not always perceived as positive changes. Participants who asserted that the MV had a negative effect on future listening episodes were distinguished from participants who experienced more positive or neutral outcomes in subsequent listens. Furthermore, there were eight participants (24% of cases) who asserted MVs had no significant impact on how they perceived the music in the future. As a result, three categories were created: long term positive, long term negative and unaffected. These categories are used in the triangulation component in order to examine relationships between the trait variables of interest and the duration and valence of MV outcomes. The criteria for these categories are outlined in Table 4.

Three particular subsequent outcomes were distinguished in the data. The first two, described as a change in *perception of affect*, change of *meaning and interpretation*, were used to establish whether the individual experienced salient long-term positive or negative salient outcomes. *Visual mental imagery* (VMI) was the most frequently reported subsequent outcome; more than three-quarters of the participants reported experiencing this effect, even those who said the MV had no long-term effects on their future listening episodes. All three categories are described in Table 5.

Quantitative Findings and Triangulation

Independent samples *t*-tests, Mann Whitney u-tests and one-way ANOVA were conducted in order to establish potential relationships between the trait variables of interest and categorical data. No statistical analyses were run on the contingent factors nor were any run on the descriptive categories for subsequent outcomes. These categories provide key insights into the type of long-term effects individuals experience, however, for the purpose of this analysis, the subsequent outcome categories were used to establish which participants experienced any long-term effects and whether that change was

Table 5. Descriptive categories for subsequent outcomes, descriptions, and frequencies.

Subsequent Outcomes		
Category	**Themes**	**Example codes**
Perception of affect	The emotional quality of the music is perceived differently in subsequent listens. This outcome often effects the subject's subsequent reasons for listening.	*It made me feel the song has heavier and more intense than before.* (P25)
Meaning and interpretation	Visual information from MV created new associations that changed the way the meaning of the music is understood	*Now, every time I hear the song, I am reminded of the video and the message that [the artist] was trying to portray in the video.* (P19)
Visual mental imagery (VMI)	The music triggers memories of visual scenes from the MV. This memory recall manifests as general thematic imagery or specific scenes triggered by specific moments in the music.	*[Now] every time I listen to the song and it hits a certain point that scene and the movement always replays in my head.* (P4)

perceived as positive or negative. For example, if a subject experienced a change in interpretation of the music's meaning in subsequent listening episodes, this was considered a long-term change. If the subject stated they disliked being reminded of the MV in their subsequent listening episodes, it was considered a long-term *negative* change. This was more suitable for the study's objective: to establish potential relationships between the duration and valence of MV

effects and individual traits related to the emotional use of music.

Independent samples t-tests reported a significant difference in unhealthy listening scores, where individuals who reported *externally* motivated goals for MV engagement had lower scores for unhealthy use ($M = 12.9$, $SD = 4.13$) compared to those who did not report these incentives for watching ($M = 19.5$, $SD = 3.27$): $t(32) = -4.48$, $p < .001$. Furthermore, individuals watching to achieve these goals had higher scores in emotional stability ($M = 4.88$, $SD = 1.35$) compared to those who did not ($M = 3.4$, $SD = 1.45$): $t(32) = 2.85$, $p = .008$. No significant relationships were observed for *internal* goals for watching and scale measures. All t-test data for *internal* and *external* goals are found in Table 6.

Table 6. T-test and U-test data for *Internal* and *External* goals.

	Internal		*External*	
	T	*P*	*t*	*p*
Healthy	94.0^	.235	0.09	.926
Unhealthy	-0.27	.79	-4.48	<.001**
Emotional Stability	111.5^	.59	2.85	.008**

* $p < .05$, ** $p < .017$ (Bonferroni correction), ^Mann-Whitney (all others t-test)

Experience outcomes were also significantly related to unhealthy listening scores, depending on if the subject experienced *strong affect* in response to the MV or were *unimpacted*. According to the independent sample t-test data, higher scores in unhealthy listening style were significantly related to *strong affect* ($M = 17.2$, $SD = 4.26$) compared to individuals who did not experience significant emotional responses ($M = 12.8$, $SD = 4.58$): $t(32) = 2.89$, $p = .007$. Unhealthy scores were also significantly lower for *unimpacted* individuals ($Md = 9.0$) compared to those who were uncertain or experienced strong emotions in response to the MV ($Md = 15.86$): $U = 12.5$, $p = .004$. Experience outcomes are reported in Table 7.

Table 7. T-test and U-test data for Experience Outcomes.

	Strong Affect		Unimpacted	
	t	p	t	p
Healthy	0.092	.93	-0.80	.926
Unhealthy	2.89	.007**	12.5^	.004**
Emotional Stability	-1.52	.14	47.5^	.23

* *p* < .05, ** *p* < .017 (Bonferroni correction),
^Mann-Whitney (all others t-test)

An independent samples *t*-test revealed that long-term outcomes, regardless of whether they were perceived as positive or negative, were significantly related to higher unhealthy use scores (*M* = 16.0, *SD* = 4.8) compared to individuals who did not experience long-term outcomes (*M* = 12.2, *SD* = 4.27): *t*(31) = 2.09, *p* = .044. T-tests also revealed that individuals who experienced long term outcomes were significantly related to lower scores in the trait emotional stability (*M* = 4.04, *SD* = 1.43) compared to those who did not, (*M* = 5.56, *SD* = 1.31),: *t*(31) = -2.77, *p* = .009. A one-way ANOVA revealed a significant effect of long-term outcome on emotional stability scores, *F*(2,30) = 5.56, *p* = .009. Tukey post-hoc comparisons revealed that those with no long-term outcomes (*unaffected*) (*M* = 5.81, *SD* = 1.36) had significantly higher emotional stability scores than both positive (*M* = 4.13, *SD* = 1.36) and negative (*M* = 3.67, *SD* = 1.13) long-term outcome groups. No significant effects were found for healthy or unhealthy scores.

Conclusion

The aim of the study was to establish a preliminary framework for understanding young people's experiences with MV media and its relationship to the use of music for emotional health. The study uses categorical outcomes outlined in the IARR framework (Wilson et al., in review) which provides new insight concerning the types of goals this method of music listening may accomplish, the emotional reactions they elicit, and the carry-over effects they impose on subsequent listening episodes.

The findings from this study suggest that MV watching is more than a form of music listening:

it is also a form of social media engagement. For some participants, MV watching was instigated not by an interest in the music or the artist or personal psychological goals, but external factors such as social media sharing, general media hype or feuds between artists. This finding also reflects an important paradigm shift in modern music listening, where an artist's success can be measured by a number of views their songs have on YouTube rather than albums sold. As a result, the line separating music listening from social media activities is becoming blurred. These *external* goals were significantly related to lower scores in unhealthy listening styles and higher scores in the trait of emotional stability, whereas no such relationships were observed for *internal* goals. The relationship between these measures and *external* goals may be indicative of the individual using MVs to feel connected to others, such as their peers. Participants who used MVs to achieve such goals frequently reported sharing their favorite MVs with their friends, sometimes watching them together. Previous research has suggested that music listening for the purpose of feeling socially connected to one's peers may have negative impacts on well-being when music replaces real friends (Chin & Rickard, 2014; Laiho, 2004) or when the individual suppresses their thoughts and feelings in order to feel connected with their peers (Chin & Rickard, 2014). Individuals who disclosed *external* reasons for watching may not necessarily be watching MVs for songs by favorite artists, instead of watching in order to be in on the "hype". This reflects a different kind of social connectedness through music that differentiates itself from previously identified social functions such as identity formation and peer-group identification (Laiho, 2004; Larson, 1995; North & Hargreaves, 1999).

The findings from the experience and subsequent outcome levels can provide a useful foundation for future research examining the links between emotional health and MV music listening during youth. The study finds evidence that stronger emotional outcomes (*strong affect*) in response to MVs were significantly related to higher scores in unhealthy listening styles, whereas individuals who experienced no

significant emotional outcomes (*unimpacted*) had significantly lower scores in unhealthy listening styles. Future research can benefit from using scales which specifically measure emotion regulation tendencies (i.e. Emotion Regulation Questionnaire; Gross & John, 2003) and use of music for mood regulation (i.e. Music for Mood Regulation Scale; Saarikallio & Erkkilä, 2007) in order to establish the relationship between outcomes from MV listening experiences and both musical and non-musical affect regulation strategies.

Future research should investigate whether traits such as empathy are a factor mediating these experiences, particularly when they are used for parasocial engagement. Whether these perceived relationships have a positive or negative influence on the use of music for emotional well-being is especially relevant during MV listening experiences based on the finding that participants frequently engage with this media in order to connect to their favourite artists.

The study identifies ways in which MVs modify future listening experiences by directly influencing mechanisms for music-evoked emotion and affect regulation. Visual mental imagery (VMI) has been identified as an underlying mechanism for music-evoked emotions (Juslin, 2013; Juslin, Harmat, & Eerola, 2014; Juslin & Västfjäll, 2008), however, there has been no research to date which analyses how visual representations of the music (such as MVs) influence this mechanism. VMI was the most frequently reported subsequent outcome, where images from the video were triggered automatically and often unintentionally. There also appears to be a link between this mechanism and scenes depicting human movement, such as performance gestures and choreographed movement. This finding may be particularly relevant for research on embodied cognition in music or neurology studies investigating mirror neuron systems (Overy & Molnar-Szakacs, 2009). Furthermore, the study highlights the ways in which multimodal listening experiences can change the way music is used to regulate affect, during the experience as well as in the future. For example, some participants suggested the MV allowed them to be more absorbed in the music than if

they were only listening to the audio since the visual component made the listening experience less passive and more engaging.

The study successfully highlights the cognitive and emotional processes involved during MV listening episodes, however, there are limitations to address. The online questionnaire component was beneficial in that participants were able to complete the study in an environment where they would usually engage with MVs, however, this also meant that there was no interviewer who could follow up or ask for more detail in the event a subject provided interesting details about their experience. Some participants provided more details than others, in some cases even providing details of more than one MV experience. For example, if a subject provided *internal* and *external* goals for MV watching and several potential subsequent outcomes, there was no way of establishing which outcomes were related to which initial reasons or goals the experience was meant to accomplish.

Overall, the current study is an important contribution to our knowledge of modern music listening, which today often occurs in the context of MV watching. Music psychological knowledge on music listening needs to be kept updated, taking these new contexts into account and we believe this study to be an important pioneering step in this endeavor.

References

Boltz, M. G. (2004). The cognitive processing of film and musical soundtracks. *Memory & Cognition*, *32*(7), 1194–1205.

Boltz, M., Ebendorf, B., & Field, B. (2009). Audiovisual interactions: The impact of visual information on music perception and memory. *Music Perception*, *1*(27), 43–59.

Bryant, Y. (2008). Relationships between exposure to rap music videos and attitudes toward relation-

ships among African American youth. *Journal of Black Psychology*, *34*(3), 356–380.

Chamorro-Premuzic, T., & Furnham, A. (2007). Personality and music: Can traits explain how people use music in everyday life? *British Journal of Psychology*, *98*(2), 175–185.

Chin, T., & Rickard, N. S. (2014). Emotion regulation strategy mediates both positive and negative relationships between music uses and well-being. *Psychology of Music*, *42*(5), 692–713.

Cohen, A. J. (2001). Music as a source of emotion in film. In P. Juslin & J. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 249–272). New York, NY: Oxford University Press.

Cohen, A. J. (2013). Congruence-association model of music and multimedia: Origin and evolution. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall (Eds.), *The psychology of musical multimedia* (pp. 17–46). Oxford, United Kingdom: Oxford University Press.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*(2), 103–113.

Edmond, M. (2014). Here we go again: Music videos after YouTube. *Television & New Media*, *15*(4), 305–320.

Gosling, S. D., Rentfrow, P. J., & Swann, Jr. W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*(2), 348–362.

Hansen, C. H., & Hansen, R. D. (2000). Music and music videos. In D. Zillmann & P. Vorderer (Eds.), *Media entertainment: The psychology of its appeal* (pp. 175–196). New York, NY: Routledge Press.

International Federation of the Phonographic Industry (2018). *Music consumer insight report 2018*. Retrieved from International Federation of the Phonographic Industry website: https://www.ifpi.org/downloads/Music-Consumer-Insight-Report-2018.pdf

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266.

Juslin, P. N., Harmat, L., & Eerola, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, *42*(4), 599–623.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575.

Kistler, M., Rodgers, K. B., Power, T., Austin, E. W., & Hill, L. G. (2010). Adolescents and music media: Toward an involvement mediational model of consumption and self-concept. *Journal of Research on Adolescence*, *20*(3), 616–630.

Krause, A. E., North, A. C., & Hewitt, L. Y. (2015). Music-listening in everyday life: Devices and choice. *Psychology of Music*, *43*(2), 155–170.

Laiho, S. (2004). The psychological functions of music in adolescence. *Nordic Journal of Music Therapy*, *13*(1), 47–63.

Larson, R. (1995). Secrets in the bedroom: Adolescents' use of media. *Journal of Youth and Adolescence*, *24*(5), 535–550.

Marshall, S. K., & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception: An Interdisciplinary Journal*, *6*(1), 95–112.

Miranda, D. (2013). The role of music in adolescent development: Much more than the same old song. *International Journal of Adolescence and Youth*, *18*(1), 5–22.

North, A. C., & Hargreaves, D. J. (1999). Music and adolescent identity. *Music Education Research*, *1*(1), 75–94.

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception: An Interdisciplinary Journal*, *26*(5), 489–504.

Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. *Psychology of Music*, *35*(1), 88–109.

Saarikallio, S., Gold, C., & McFerran, K. (2015). Development and validation of the healthy-unhealthy music scale. *Child an Adolescent Mental Health*, *20*(4), 210–217.

Saarikallio, S., Nieminen, S., & Brattico, E. (2013). Affective reactions to musical stimuli reflect emotional use of music in everyday life. *Musicae Scientiae*, *17*(1), 27–39.

Sun, S. W., & Lull, J. (1986). The adolescent audience for music videos and why they watch. *Journal of Communication*, *36*(1), 115–125.

Timmermans, S., & Tavory, I. (2012). Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory*, *30*(3), 167–186.

van Goethem, A., & Sloboda, J. (2011). The function of music for affect regulation. *Musicae Scientiae, 2*(15), 208–228.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*(2), 157–170.

Vuoskoski, J. K., & Eerola, T. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18–49.

Wilson, J. N., Thompson, M. R., & Saarikallio, S. (in review). *YouTube killed the radio star: Exploring the functions and effects of music video listening experiences.*

STUDY III

THE CHARACTERISTICS OF MUSIC VIDEO EXPERIENCES AND THEIR RELATIONSHIP TO FUTURE LISTENING OUTCOMES
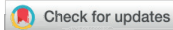
by

Johanna Dasovich-Wilson, Marc Thompson & Suvi Saarikallio

Original Empirical Investigations

The characteristics of music video experiences and their relationship to future listening outcomes

Johanna N Dasovich-Wilson (iD)**, Marc Thompson
and Suvi Saarikallio** (iD)

Abstract

Music videos are a popular method of consuming music; however, the characteristics of these experiences and their effects on music perception are poorly understood. An online survey ($N = 155$) was designed using theoretical insight from Dasovich-Wilson et al.'s (2022) Intention Attention Reaction and Retention (IARR) framework. The survey consisted of two parts: the first explored the key characteristics of music video experiences, and the second explored their effects on subsequent listening outcomes. Separate principal component analyses (PCAs) were performed on each part to differentiate between the experience itself (Experience components) and the effects on subsequent listens (Retention outcomes). Relationships between Experience components and Retention outcomes were explored using correlation and regression analyses. The results suggest that music video experiences characterized by performance gestures and narratives have the strongest influence on music perception. These findings shed light on how extramusical information from music videos influences mechanisms related to visual imagery and personal associations.

Keywords

audio-visual, perception, meaning, music videos, principal components analysis

Music videos have been a part of mainstream music culture for more than four decades. Originally intended as a marketing tool for record companies to promote new albums, this form of post-modern art became a cultural phenomenon popular among youth (Aufderheide, 1986); its nonlinear, non-narrative structure was seen as a reflection of the superficial and hedonic culture from which it emerged (Fiske, 1986). Today, music videos function not only as a

Department of Music, Arts and Culture Studies, University of Jyväskylä, Jyvaskyla, Finland

Corresponding author:
Johanna N Dasovich-Wilson, Department of Music, Arts and Culture Studies, University of Jyväskylä, P.O. Box 35,
FI-40014 Jyväskylä, Finland.
Email: johanna.n.wilson@jyu.fi

promotional tool but as a revenue stream in their own right. Short, self-contained, and open to interpretation, music videos are an ideal format during the internet age where new content is always in demand (Edmond, 2014). Whether available on music channels such as MTV or accessed on modern streaming platforms such as YouTube, they stand out as a popular method of consuming musical multimedia. However, the characteristics and effects of music video experiences have been largely overlooked in music psychology research.

A recent study by Dasovich-Wilson et al. (2022) proposed a framework for conceptualizing the psychological functions and effects of music video experiences. The purpose of this framework was to address this gap in music psychology research and provide a foundation for future research exploring the topic. In their study, qualitative data from 34 adolescents and young adults were collected and analyzed using an abductive grounded theory approach (see Charmaz, 2006). The resulting framework categorized music video experience-related phenomena across temporal levels: *intention*, *attention*, *reaction*, and *retention* (IARR). Thematic categories within each level provided insight into participants' reasons for watching, how the visual and musical content influence attention, the affective outcomes they evoke, and their influence on subsequent hearings. The categories in the *retention level* of the IARR reflect how perceptual phenomena that occur during the initial experience affect the perception of musical meaning, which can have an influence on subsequent listening outcomes. Dasovich-Wilson and colleagues concluded that extramusical information from music videos—including narrative content, visual effects, dance, and performance gestures—affects what the listener associates with the music. These associations can have a long-term influence on how the listener perceives a song's meaning and/or affective quality in subsequent listens. However, the IARR did not consider the relationships between conceptual categories, either within or across temporal levels.

The current study uses conceptual insight from the IARR framework to explore the characteristics of music video experiences and their effects on subsequent hearings using quantitative methods. The topic of music videos (and music multimedia in general) is especially relevant in today's music listening and multimedia climate, where music media is easily accessible from virtually anywhere at any time. To address this phenomenon, we highlight existing research on the functions of music and music videos and current theories of audio-visual perception.

The function of music in an information age

Music has been widely recognized for its ability to fulfill psychological goals which can be reduced to three dimensions: cognitive, affect-regulation, and social relatedness (Hargreaves & North, 1999; Rentfrow, 2012; T. Schäfer et al., 2013). These functions also apply to music video experiences. According to Sun and Lull (1986), youth watch music videos to better understand the meaning of their favorite songs as well as for emotion regulation, to pass the time, as part of social interactions and for social learning purposes—a finding which has also been supported in recent studies (Dasovich-Wilson et al., 2022; Wilson et al., 2020). Music video experiences, however, have evolved in recent years, in part due to concepts such as what is "trending" on social media platforms—a phenomenon that did not exist at the time of Sun and Lull's study. T. Schäfer et al. (2013) posited that the social function of music is less important compared to its cognitive and affect-regulating functions. They suggested this reflects Western ideals that emphasize the importance of individuality: "*self-acknowledgement and well-being appear to be more highly valued than social relationships and relatedness* (p. 7)." Schäfer and colleagues could not have predicted that social media and audio-visual platforms such as YouTube would become as commonplace as they have in recent years. These platforms have

become important methods of consuming, engaging with, and sharing music media. Schäfer and colleagues rightly pointed out the importance of self-acknowledgment and well-being. However, current trends in music engagement (including the platforms on which they occur) seem to more strongly suggest that music's social function has evolved.

Social media and streaming platforms like YouTube have contributed to this evolution. This has implications for music's cognitive and affect-regulation functions by virtue of the fact people may spend a lot of time engaging with these platforms: using them to cultivate their own sense of self, change their affective state, or both. However, the strongest contributor affecting these functions is the wealth of extramusical information they provide. For music videos, this comes in the form of visual content. The extent to which this information influences the perception of the music—and by extension, its function—is an important question that remains unanswered. Understanding how an individual processes and uses that information is a key step toward answering it, and the social component from which this information is accessed cannot be ignored.

Extramusical associations provide a cognitive explanation as to how listeners' thoughts and meanings about the outside world, such as personal memories and other associated concepts, contribute to their experience of music (Meyer, 1956). Many of the mechanisms of music-evoked emotion (BRECVEMA; see Juslin et al., 2014; Juslin & Västfjäll, 2008) are grounded in the experience and associations of the listener as well, including musical expectancy, episodic memory, and visual mental imagery. Associated meanings in the form of personal memories can trigger mechanisms such as episodic memory (Juslin & Västfjäll, 2008), resulting in emotional outcomes such as nostalgia (Barrett et al., 2010). Extramusical meaning can occur as a result of perceived shared experiences the listener has with the artist (K. Schäfer & Eerola, 2020), fostering parasocial relationships, especially among youth (Greenwood & Long, 2009; Kistler et al., 2010). These shared personal meanings can promote affect regulation through strategies such as *solace* (Saarikallio & Erkkilä, 2007), providing consolation for the listener during hard times (ter Bogt et al., 2017). However, for an individual whose musical engagement style indicates poorer emotional health and well-being (see Saarikallio et al., 2015), watching a music video for songs with deeply personal meanings attributed to them only to find out the "true" meaning of the song from the perspective of the artist can have potentially devastating consequences if those two meanings are in conflict with each other (Dasovich-Wilson et al., 2022).

Music's function depends on the individual, their relationship with music, and the context where the listening event takes place. Not all individuals place the same importance on music, and individuals who engage with music more often are more likely to use it to fulfill several functions simultaneously (Greasley & Lamont, 2011). Music videos can potentially change the type of associations attributed to a piece of music, which may affect its utility in fulfilling certain functions, for better or worse. Research that explores the cognitive mechanisms that make this possible will provide more insight into how the affective impact of music video content can manifest itself in subsequent listening experiences, even when the visual component is no longer present.

Audio-visual interactions and music perception

To understand how music videos can contribute to the experience of music, audio-visual interactions and their influence on cognition need to be considered. In film contexts, music can help establish mood, clarify the emotional meaning of a scene, and provide insight into the

motivations behind the characters' behavior (Cohen, 2001, 2013). According to Cohen's Congruence-Association Model (CAM; Cohen, 2001, 2013; Marshall & Cohen, 1988), the audience's attention is guided by shared structural congruencies between the audio and visual modalities, in line with the Gestalt principle of grouping by similarity. Visual elements that are temporally aligned with the music become the object of focus, and inferences about the emotional meaning of the scene are generated based on the viewer's past experiences. This top-down process of searching for information matching the emotional information of the music to the semantic elements of the film helps the audience establish an interpretation of film events. When the film's content is ambiguous, music has a polarizing effect: The emotional quality of the music is applied to the ambiguous content, which influences how the viewer understands the film's plot (Bullerjahn & Güldenring, 1994).

Cognitive schemas, which are cognitive networks created from past experiences that represent knowledge about concepts or other stimuli, their attributes, and the relationships between those attributes stored in long-term memory (Sakamoto & Love, 2004), may serve as an explanation for how long-term memory is activated in CAM (Shevy, 2008). A network of cognitive schemas stored in long-term memory is activated depending on the individual who is listening, the context where the listening event is taking place, and the features of the music. This "*spreading activation*" has been suggested as an underlying cognitive mechanism responsible for the role of familiarity in establishing musical preference (Schubert et al., 2014). In the case of musical multimedia, information about the meaning of the music, the character's actions, and perceived congruencies between the audio and the visual channels provide an opportunity for new associations to form. Spreading activation posits that as familiarity increases more associations are formed, making the music more enjoyable. However, spreading activation occurs at an unconscious level: The listener is not necessarily aware that they are making these connections (unless they are consciously trying to do so). For example, in a study by Strobin et al. (2015), the genre of the music that played during a film trailer influenced participants' expectations of the feelings that would be expressed in the film. Furthermore, they found that the genre of the music played in the trailer influenced expectations of the film's genre. The music, however, was not an important factor in determining whether they would go to see the film. This may be because the function of the music in this scenario is to fit with the content of the movie and to clarify the characteristics of the film (genre, events, and semantic quality), whereas music in music *video* contexts has a potentially more important psychological function that reflects the specific needs of the individual.

Unlike the influence of music in film trailer contexts, however, associations with visual material may be too intrusive to be subconscious in music video experience contexts. The most extreme example of this would be cases where hearing a song triggers mental imagery of the music video in subsequent listening episodes (Dasovich-Wilson et al., 2022). Joint encoding is a possible explanation for this phenomenon: when the music and video are perceived as mood-congruent (e.g., two individuals getting into a physical fight accompanied by angry music), information from both channels is integrated to form a single, audio-visual memory code (Boltz, 2004). This may explain why visual mental imagery from the video was a common occurrence in Dasovich-Wilson's study, even in the absence of any other perceptual changes.

A large body of research has investigated how extramusical information influences music listening experiences in both personal music listening and multimedia contexts. Despite this, music video experiences have not yet been adequately explored. This lack of previous research establishes the need and motivation for the current study.

Objectives

The current study explores the relationship between conceptual categories within the IARR framework using quantitative methods to establish the key components of music video experiences (i.e., *intention*, *attention*, and *reaction level* categories) and their influence on subsequent outcomes (i.e., *retention level* categories). The study investigates two research questions:

What are the key characteristics of music video experiences (Experience components) and their carry-over effects (Retention outcomes)?

Because the study was designed using Dasovich-Wilson et al.'s (2022) framework as a starting point, Experience components should reflect themes such as engaging with narrative content, seeing the artist perform, for emotion regulation purposes, and as part of social media engagement. We expect Retention outcomes to reflect how visual and personal associations with the music change after having seen the music video.

How are Experience components related to Retention outcomes?

This second question is exploratory; however, we expect the strongest predictors for Retention outcomes to be experiences that emphasize visual content such as narratives outlining the meaning of the music and/or lyrics, and performance gestures (musical playing or dance).

Methodology

Participants

Participants were recruited via the authors' university email lists, Twitter, and Survey Circle (www.surveycircle.com). The first 25 participants were given an Amazon voucher valued at €5. No other compensation was provided; however, all participants who completed the survey and provided a valid email address were entered in a raffle to win one of two Amazon gift cards (value of €25). Participants who did not complete the entire survey were removed for analysis, as well as three participants who filled out the same answer for each question. The final sample consisted of 155 participants between the ages of 15 and 44 ($M = 27.3$, $SD = 5.8$); 66% of the sample were female and 62.3% were currently living in a European Union (EU) country. The most frequently observed nationality was German (12.25% of cases), followed by British (11.6%), Finnish (11.6%), and French (9.67%). Most participants identified as music-loving non-musicians (44.5%; a full breakdown of participants' nationalities and musical backgrounds can be found in the Supplementary Data file).

Questionnaire

Participants provided questionnaire data concerning their experiences with music videos using a 5-point Likert scale (1 = *disagree strongly*, 5 = *agree strongly*). The items were designed using the qualitative categories from Dasovich-Wilson et al.'s (2022) IARR framework as a starting point. The items reflecting these qualitative categories were organized in two parts. The first part contained items describing the main categories of the intention, attention, and reaction levels of the IARR framework. Intention items describe the cognitive and emotional goals that motivated music video engagement, as well as social influences such as peers, as well as social

media platform use. The attention level reflects the structural and semantic content of the music video to gain insight into the *type* of content that would absorb attention. The reaction level reflects the type of affective outcomes that occur during and immediately following the music video experience, such as a boost in energy, as well as emotional processes that occur as a result of mechanisms such as empathy and contagion. Items from the intention level subcategories were labeled as cognitive (C), emotional (E), and social (S), and items derived from the attention and reaction levels were labeled A and R, respectively (see Table 1). Additional information concerning subcategories (e.g., *interpretation-focused* or *affect-focused* at the attention level) was not considered necessary for the purpose of this study.

The second set of items explored the potential carry-over effects music videos would impose on their subsequent music-listening experiences. These items were based on the three retention-level categories in the IARR: (1) *visual mental imagery*, which reflects how content from the music video was recalled as in subsequent listens; (2) *new interpretation of meaning*, which reflects a change in understanding of the music's meaning as a result of the video's semantic information, and (3) *new affect perception*, which reflects a change in affective reactions, whether it be a different emotional response to the song or different appraisal. Retention-level items were also labeled according to the IARR thematic categories they were derived from (see Table 2).

To provide situational context to the experiences, nominal and categorical data were collected concerning the locations where music video experiences occur, whether they prefer to watch music videos alone or with others, and the devices they used were also collected in the questionnaire. This information is available in the Supplementary Data.

Procedure

The study was hosted online via the webropol survey platform. Participants indicated their consent to participate on the opening page, which also detailed the study's purpose and instructions. Because the study did not put participants at any risk and participation was entirely voluntary, ethics approval was not required from the researchers' host institution according to the guidelines of the Finnish National Board on Research Integrity.[1] Once participants completed the survey, they were given the option to provide their email to participate in the prize draw.

Results

Preliminary analysis

In total, two principal component analyses (PCAs) were performed. Components with an eigenvalue above 1.0 were retained and sampling adequacy was evaluated using the Kaiser-Meyer-Olkin (KMO < .60; Kaiser, 1974). The first PCA was conducted to reduce the dimensionality of items at the Intention and Experience level (referred to as Experience components for the rest of the article), whereas the second PCA reduced the dimensionality of items at the Retention level (R1). Promax rotation was used to increase the interpretability of the loadings because correlations between components are expected. Once components were established, correlations were performed to investigate the relationship between Experience components and Retention outcomes. Linear regressions were performed to establish which Experience components were the strongest predictors for each Retention outcome (R2).

Table 1. Principal Component Structure for Experience Items.

	Item	*M (SD)*	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Int7	I watch music videos that are like short films and have storylines (C)	3.51 (1.15)	**0.51**		**0.51**				0.43
Int6	I watch music videos to see the artist's interpretation of the music (C)	3.53 (1.18)	**0.50**	**0.52**	**0.61**				
Int5	I watch music videos to gain more insight about the meaning of the music and/or lyrics (C)	3.46 (1.17)			**0.81**				
Int8	I watch music videos to reflect on the content of the music/lyrics (C)	3.35 (1.22)		0.42	**0.69**				
Int14	I watch music videos to see the artist and how they perform the music (C)	3.79 (1.14)		**0.67**	0.46				
Int9	I watch music videos to experience the music differently than if I were only listening (C)	3.63 (1.19)	0.48	**0.64**	0.45				
Int13	I watch music videos to watch and/or learn dance moves and choreography (C)	2.53 (1.32)						**0.62**	
Int1	I watch music videos when I am bored have and nothing to do (E)	3.43 (1.18)					**0.62**	0.47	
Int2	I watch music videos when I need a boost of energy and motivation (E)	3.40 (1.26)		**0.58**	0.42		**0.65**		
Int3	I watch music videos to distract myself when I am in a bad mood or negative emotional state (E)	3.28 (1.21)					**0.8**		
Int4	I watch music videos to help me relax and unwind (E)	3.50 (1.22)	0.40		0.48		**0.69**		
Int11	I watch music videos that are hyped in the media and have a reputation for something (S)	3.23 (1.19)						**0.72**	
Int10	I watch music videos as part of social activities with friends (S)	3.05 (1.33)						**0.60**	
Int12	I watch music videos when YouTube recommends them to me or auto-plays them (S)	3.22 (1.12)							**0.72**
Exp7	I enjoy music videos that tell a story (A)	3.94 (1.06)	**0.77**						
Exp8	I enjoy music videos that show an alternative interpretation of the music and is different from my own interpretation (A)	3.52 (0.97)	**0.71**	0.40					
Exp11	I enjoy music videos that contain special effects and engaging cinematography (A)	3.65 (1.07)	**0.71**			0.41		0.43	

(Continued)

Table 1. (Continued)

Item	M (SD)	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Exp10 I enjoy music videos that are ambiguous and open-ended (A)	3.35 (1.05)	**0.66**					0.40	
Exp12 I enjoy music videos that are visually interesting and conceptual art (A)	3.97 (0.96)	**0.59**	**0.55**				0.41	
Exp15 I enjoy music videos that are recordings of live performances (A)	3.30 (1.26)		**0.70**					
Exp13 I enjoy music videos that highlight the music's features in creative ways (A)	3.95 (0.99)	**0.54**	**0.68**					
Exp6 I enjoy music videos that show an interpretation of the music that is similar to my own personal interpretation (A)	3.44 (0.99)	0.45			**0.73**			
Exp9 I enjoy music videos that highlight the meaning of the lyrics and clarify their true meaning (A)	3.66 (1.02)			0.42	**0.67**			
Exp14 I enjoy music videos that contain dance, choreography (A)	3.49 (1.19)	**0.52**	0.45		0.43		**0.59**	0.43
Exp16 I enjoy music videos that contain clips of the artist performing but are not live recordings (A)	3.48 (1.05)		**0.53**		**0.57**			
Exp1 The emotional quality of the music video defines my emotional reaction (R)	3.37 (1.07)				**0.65**			
Exp2 Music videos can motivate me or energize me, and give me a boost of energy (R)	3.57 (1.01)		**0.72**	0.44	0.45	0.40		0.41
Exp3 Music videos can elicit strong emotional reactions in me (R)	3.72 (1.05)	**0.50**	**0.6**	**0.57**				
Exp4 Music videos can help me understand the song on a deeper emotional level (R)	3.58 (1.07)	**0.54**	**0.5**	**0.51**				
Exp5 Seeing the music video can make me relate to the artist more as a person (R)	3.46 (1.04)	**0.52**	**0.51**	**0.55**	**0.54**			
Eigenvalue		8.82	1.87	1.74	1.48	1.34	1.16	1.14
Variance Explained		29.40%	6.23%	5.81%	4.94%	4.46%	3.87%	3.80%

PC: principal component. Loadings above or equal to .50 are in bold.

"Int" items were derived from the main thematic categories from the Intention level of Dasovich-Wilson and colleagues' (2022) IARR framework. "Exp" items were derived from Attention and Reaction level categories, which pertain to the experience itself. Labels in parentheses: (A) items refer to specific content features of music videos that grab attention, including structural and semantic features (C) items reflect music's cognitive function, (E) items reflect emotional uses of music videos, (R) items reflect affective reactions from music videos, (S) refers to social functions.

Table 2. Principal Component Structure for Retention Outcome Items.

	Item *When I hear a song I have seen the music video for. . .*	*M (SD)*	PC1	PC2	PC3
R7	I picture the choreography from the music video in my mind (VI)	3.29 (1.15)	**0.62**		
R8	I picture scenes of the artist performing in my mind (VI)	3.35 (1.11)	**0.78**		
R9	I picture specific scenes or imagery from the video in my mind (VI)	3.63 (0.99)	**0.77**	**0.57**	-0.45
R6	I think of the storylines and/or characters from the music video (VI/NIM)	3.51 (0.99)	**0.66**	0.45	
R3	It is hard not to think of the music video (NIM)	3.14 (1.14)	**0.71**		
R2	The music video can have a stronger influence than my own personal memories or associations with the song (NIM)	2.98 (1.15)	0.49	**0.59**	
R10	I think of the topic(s) of the music video (NIM/NAP)	3.39 (1.04)	**0.65**	**0.56**	
R1	The song can be more meaningful and/or personally significant to me (NAP)	3.46 (1.03)		**0.78**	
R4	The mood of the song can seem different (e.g., a sad song sounds less sad) (NAP)	3.35 (0.99)		**0.54**	
R5	How I use the song to influence my mood can change (e.g., I used to listen to the song when I was happy, now I listen to it when I am sad) (NAP)	2.89 (1.02)		**0.74**	
R11	I almost never think back to the music video (UA)	2.48 (1.05)			**0.85**
LT	Music videos have a long-term influence on how I perceive and/or understand the meaning of the music[a]	3.35 (1.06)	**0.60**	**0.51**	
NLT	Music videos have no influence on my future listening experiences[a]	2.59 (1.12)			**0.87**
	Eigenvalue		4.66	1.37	1.00
	Variance explained		35.85%	10.55%	7.71%

PC: principal component; VI: visual mental imagery; NIM: new interpretation of meaning; NAP: new affect perception; UA: unaffected; LT: long term; NLT: non-long term.
Labels in parentheses refer to the retention-level categories from Dasovich-Wilson's et al. (2022) original study.
Loadings above or equal to .50 are highlighted in bold.
[a]LT and NLT items were presented as written.

Principal component analysis

Experience level. A PCA was performed to investigate the key characteristics of music video experiences. The KMO measure verified excellent sampling adequacy (*KMO* = 0.88). Seven Experience components were established (see Table 1) which account for 58.51% of the variance. These were labeled: *(a) Narrative Content*, *(b) Visual Performance*, *(c) Meaning of Lyrics*, *(d) Personal Interpretation*, *(e) Emotional Use*, *(f) Social Use*, and *(g) YouTube Recommender.*

The first four components are labeled according to the particular features of music videos that absorb attention and/or motivate listeners to watch them. The strongest items for Narrative Content describe watching music videos that are story-driven (*I enjoy music videos that tell a story; I watch music videos that are like short films and have storylines*), and depict a different interpretation of the song's meaning than their own (*I enjoy music videos that show an alternative interpretation of the music that is different from my own interpretation*). The Visual Performance component consists mostly of items that emphasize the importance of visual content, especially seeing the artist perform (*I enjoy music videos of live performances; I watch music videos to see the artist and how they perform the music*). However, the item with the strongest loading for this component describes how the experience increases state arousal (*music videos can motivate me or energize me and give me a boost of energy*). The items with the strongest loading for the Meaning of Lyrics component reflect the intention to learn about the meaning of the music and to reflect on that meaningful content (*I watch music videos to gain more insight about the meaning of the music and/or lyrics; I watch music videos to reflect on the content of the music/lyrics*). Furthermore, this component has strong loadings on items describing certain affective reactions, including a deeper understanding of the music's emotional tone and an increase in state arousal. The reaction item with the strongest loading reflects how music videos can cause listeners to feel more connected to the artist on a personal level (*seeing the music video can make me relate to the artist more as a person*). The fourth component, Personal Interpretation, also reflects experiences where attention is directed toward how the meaning of the music (or lyrics) is represented but emphasizes the importance of the listener's personal interpretation of the music (*I enjoy music videos that show an interpretation of the music that is similar to my own personal interpretation*).

The last three components more prominently feature items describing the reasons (intentions) for engaging with music videos. Emotional Use items reflect music's affect-regulating function and contain items reflecting the use of music videos with the intent to distract themselves from a negative affective state, to relax, to feel more motivated, or to relieve boredom. The Social Use and YouTube Recommender components reflect the social functions of music listening and the role of social media platforms (including but not limited to YouTube) and other media outlets in influencing music video consumption. The Social Use items with the strongest loadings reflect the influence of both peers and broader media hype (*I watch music videos that are hyped in the media and have a reputation for something*), and videos with choreography (*I watch music videos to watch and/or learn dance moves and choreography; I enjoy music videos that contain dance and choreography)*. The YouTube Recommender component primarily reflects how YouTube recommendations and its auto-play function influence music video engagement (*I watch music videos when YouTube recommends them to me or auto-plays them*).

Retention level. A second PCA was performed to investigate the characteristic Retention outcomes.

The KMO measure revealed excellent sampling adequacy (*KMO* = 0.85). Three Retention outcome components were established, and account for 54.11% of the variance (see Table 2). These components were labeled: *(a) Visual Mental Imagery*, *(b) Personal Significance Increase*, and *(c) Unaffected*, and reflect the qualitative characteristics of music video effects and their duration.

The Visual Mental Imagery and Personal Significance Increase Retention outcomes highlight how music videos influence the type of associations made with the music, a song's personal significance to the listener, and its influence on affective states. Both reflect the carry-over effects imposed on future listening experiences. The Visual Mental Imagery outcome sheds light on how visual information is recalled during listening, whereas Personal Significance Increase reflects how this information affects the song's influence on their mood or emotional state. The item, "*When I hear a song that I have seen the music video for, it is hard not to think of the music video,*" was a key item for Visual Mental Imagery; this suggests that the mental imagery of the video may be triggered automatically, making it difficult for the individual to disassociate the song with the video in the future. Personal Significance Increase not only reflects how music video content can influence the perception of the music's emotional character (*the mood of the song can seem different . . .*) but also how this content makes the song more significant to the listener (*the song can be more meaningful and/or personally significant to me*). The Unaffected component reflects experiences where the music video has no salient influence on how a song is perceived or what is associated with it in subsequent listens. Because this study is concerned with music video experiences and effects, the rest of the analyses and discussions will primarily focus on Visual Mental Imagery and Personal Significance Increase (all correlations and linear regression analyses conducted for the Unaffected outcome are reported in the tables).

Correlation and linear regression results

Correlations and linear regression analyses were performed to establish which Experience components are the strongest predictors for each Retention outcome (R2). All correlations are reported in Table 3 and all linear regressions are in Table 4. Bonferroni correction was applied to all regression results ($p = .007$).

Both Visual Mental Imagery and Personal Significance Increase Retention outcomes had a moderately strong, positive correlation with the Experience components: Narrative Content, Visual Performance, Meaning of Lyrics, Personal Interpretation, and Emotional Use. The Experience components with the strongest positive correlations with Visual Mental Imagery were Visual Performance, $r(153) = .53$, $p < .001$, and Personal Interpretation, $r(153) = .45$, $p < .001$. A moderate positive correlation was also observed between both Visual Mental Imagery and Personal Significance Increase, $r(153) = .57, p < .001$.

Visual Mental Imagery. Linear regression analysis revealed the extent to which each Experience components could predict the Retention outcome Visual Mental Imagery. Interestingly, all seven components were predictors of this outcome; however, there were four Experience components that significantly and positively predicted Visual Mental Imagery and explained more than 10% of the variance. Visual Performance explained 27% of the variance, $R^2 = .27$, $F(154) = 59.25$, $p < .001$, 95% confidence interval (CI) [0.39, 0.66]; Personal Interpretation explained 20% of the variance, $R^2 = .20$, $F(154) = 40.16$, $p < .001$, 95% CI [0.31, 0.60]; Narrative Content explained 15% of the variance, $R^2 = .15$, $F(154) = 28.49$, $p < .001$; and Meaning of Lyrics explained 14% of the variance, $R^2 = .14$, $F(154) = 25.53$, $p < .001$, 95% CI [0.23, 0.526].

Personal Significance Increase. The linear regression analysis revealed significant positive relationships between Personal Significance Increase and Experience components. The strongest predictor for Personal Significance Increase was also Visual Performance, which explained 23% of the variance, $R^2 = .23$, $F(154) = 47.28$, $p < .001$, 95% CI [0.35, 0.63]. Meaning of Lyrics explained 21% of the variance, $R^2 = .21$, $F(154) = 42.13$, $p < .001$, 95% CI [0.32, 0.61];

Table 3. Correlations between EC and RO.

Variables	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Narrative content[a] (EC)	–									
2. Visual performance (EC)	.506***	–								
3. Meaning of lyrics (EC)	.444***	.494***	–							
4. Personal interpretation[a] (EC)	.307***	.383***	.177*	–						
5. Emotional Use[a] (EC)	.245**	.382***	.433***	.204*	–					
6. Social use (EC)	.368***	.265***	.118	.250**	.253**	–				
7. YouTube recommender[a] (EC)	.192*	.132	−.018	.268***	.020	.218***	–			
8. Visual mental imagery (RO)	.395***	.528***	.378***	.447***	.280***	.287***	.242**	–		
9. Personal significance increased (RO)	.463***	.486***	.465***	.361***	.470***	.228***	.187*	.572***	–	
10. Unaffected[a] (RO)	−.149	−.214**	−.188*	−.124	−.121	−.050	−0.015	−.442***	−.286***	–

EC: Experience components; RO: Retention outcomes.
[a]Spearman's rho.
***$p < .001$; **$p < .01$; *$p < .05$

Table 4. Retention Outcomes Predicted by Experience Components.

Retention outcome	Experience component (predictor)	β	*SE*	β/*SE*	*p*
Visual Mental Imagery	Narrative content[a]	.396	.074	5.338	**<.001****
	Visual performance[a]	.528	.069	7.697	**<.001****
	Meaning of lyrics	.378	.075	5.053	**<.001****
	Personal interpretation[a]	.456	.072	6.337	**<.001****
	Emotional use	.265	.078	3.402	**<.001****
	Social use	.287	.077	3.713	**<.001****
	Youtube recommender	.244	.078	3.111	.002*
Personal Significance Increase	Narrative content[a]	.456	.072	6.329	**<.001****
	Visual performance[a]	.486	.071	6.876	**<.001****
	Meaning of lyrics[a]	.465	.072	6.491	**<.001****
	Personal interpretation[a]	.393	.074	5.280	**<.001****
	Emotional use[a]	.439	.073	6.038	**<.001****
	Social use	.228	.079	2.902	.004*
	Youtube recommender	.220	.079	2.794	.006*
Unaffected	Narrative content	−.138	.080	−1.726	.086
	Visual performance	−.236	.079	−3.004	.003*
	Meaning of lyrics	−.241	.078	−3.067	.003*
	Personal interpretation	−.126	.080	−1.569	.119
	Emotional use	−.108	.080	−1.347	.180
	Social use	−.008	.081	−0.104	.917
	Youtube recommender	−.046	.081	−0.576	.566

SE: standard error.
β is the unstandardized parameter estimate; β/*SE* is the *t*-value associated with the parameter estimate
[a]Are principal components with an $R^2 \geqslant .15$.
**p < .001; *p < .007 (Bonferroni correction). The significance of bold values is ** *p*<.001.

Narrative Content explained 20% of the variance, $R^2 = .20$, $F(154) = 40.06$, $p < .001$); Emotional Use explained 19% of the variance, $R^2 = .19$, $F(154) = 36.45$, $p < .001$; and Personal Interpretation, which explained 15% of the variance, $R^2 = .15$, $F(154) = 27.88$, $p < .001$). Social Use and YouTube Recommender components explained only 5% and 4% of the variance, respectively. These findings were in line with our hypothesis that experiences that emphasize a focus on the music's meaning or narrative and affect-regulation functions would be the strongest predictors of Retention outcomes.

Discussion

The aim of the study was to explore the key characteristics of music video experiences (Experience components) and their effects on future listening experiences (Retention outcomes). The study also explored the relationship between Experience components and two Retention outcomes: Visual Mental Imagery and Personal Significance Increase.

At the Experience level, seven principal components emerged. These components reflect the different characteristics of music video experiences based on the content they contain, the reasons for watching, and the affective reactions they evoke. Furthermore, these components highlight the ways in which music videos can satisfy the different cognitive, emotional, and social functions attributed to music. Narrative Content, Visual Performance, Personal

Interpretation, and Meaning of Lyrics were characterized by the content they contain, and not necessarily a specific function-based intention such as to change one's emotional state. The Personal Interpretation component is particularly unique in this respect. Whereas the other three content-based components feature at least one intention-related question (e.g., *I watch music videos to gain more insight about the meaning of the music and/or lyrics*), Personal Interpretation does not: It is characterized by the music video validating an interpretation of the music that is similar to that of the individual. On the other hand, Meaning of Lyrics is characterized by the intent to know what the song is about on a deeper level.

Although the first four components are characterized primarily by their content, the last three reflect why (or how) music video experiences come about. Emotional Use and Social Use reflect specific functions that music video experiences can fulfill. On the other hand, the YouTube Recommender component reflects the influence of the platform itself, which generates recommendations based on the user's previous activity on the site (see Zhou et al., 2010) and was characterized mostly by a single item (*I watch music videos when YouTube recommends them to me or auto-plays them*). This reflects YouTube's popularity as a streaming platform (77% of participants indicated that they sometimes or often use YouTube to stream music, see Supplementary Data). Interestingly, the other two items loading on this component describe enjoying music videos that have engaging choreography and storylines, and getting a boost of energy from the experience (notably, the loadings of these items are comparatively weaker). Given that many of YouTube's most watched videos have dance moves associated with them which often contribute to their going viral, this was not surprising. For example, the first music video on the platform with over two billion views was PSY's *Gangnam Style* (Benjamin, 2014) and the dance from this video was a viral meme at the time. In addition, the YouTube Recommender component reflects music video experiences from using YouTube for music streaming. While the algorithms involved in recommending and auto-playing videos are beyond the scope of this study, this component was included and labeled after the platform itself because it demonstrates a method of listening that has, to date, been under-explored in empirical research.

At the Retention level, two principal components were established that reflect the potential long-term effects of music video content on music perception: Visual Mental Imagery, which reflects how the visual imagery mechanism is affected, triggering internal imagery of the video in subsequent listens; and Personal Significance Increase, which reflects how the personal significance of a song is positively influenced by music video content, influencing the perception of and subsequent responding to the song's emotional quality in the future. This finding reflects the overlap between changes in the perception of the music's affective quality and changes in the interpretation of meaning previously reported by Dasovich-Wilson et al. (2022). A moderate correlation was observed between Visual Mental Imagery and Personal Significance Increase outcomes, suggesting they may co-occur depending on the music video and the individual's relationship with the song. The last Retention component, Unaffected, represents the absence of effects in future listens.

The second aim of the study was to explore potential relationships between Experience and Retention components. The results are in line with our initial expectation that Experience components emphasizing the visual experience and meaning of the music would be the strongest predictors of Retention outcomes. While Visual Mental Imagery was predicted by all Experience components, the strongest predictor was Visual Performance. The cognitive mechanisms involved in mental imagery (which include visual, auditory, and motor imagery) offer a potential explanation for this relationship. Mental imagery is activated by the same neural processes involved during perception of the same modality (in this case, visual) as well as

memory, emotion, and motor control mechanisms (Kosslyn et al., 2001). Visual Performance experiences involve attending to physical performance gestures such as a musician playing or dance choreography, and strong affective reactions including a boost in energy. The relationship with mental imagery may be a result of a joint memory code being formed during the experience being made salient due to the emotion and motor mechanisms activated. Visual Mental Imagery of human gestures may reflect the activation of mirror neuron mechanisms: neurons that activate during both observation and execution of physical action (Overy & Molnar-Szakacs, 2009). Although beyond the scope of the present study, this finding may be of interest for research on embodied cognition or neurological studies of music and multimedia. This is particularly relevant for studies exploring the relationship between music and memory retrieval.

Other Experience components which were strong predictors of Visual Mental Imagery were Personal Interpretation, followed by Narrative Content. Narrative Content experiences reflect an alternative interpretation of the music—or one that is ambiguous and open-ended— whereas Personal Interpretation experiences reflect the preference for the video depicting an interpretation similar to the meaning the listener has already attributed to the song. In both cases, the video's depiction of the music's meaning potentially provides a visual association that is recalled in future listens. However, the relationship with Personal Interpretation complements previous research concerning how cognitive networks are activated during music listening. When the music video is congruent with the listener's personal associations or understanding of the music's meaning, information from the music video may be more easily stored in memory. This may strengthen existing cognitive networks and create new ones as the music becomes more familiar (Schubert et al., 2014). This finding may also reflect how CAM functions in music video contexts; the individual is not only directing attention toward the music and videos shared structural features (see Cohen, 2013), but the similarities between the content of the video and their previous associations. It may be easier to store the visual information from the video in long-term memory because it confirms their expectations (i.e., previous associations) about the music, thus reinforcing that association. On the contrary, the reverse can also be true: if the video challenges one's expectations about the meaning of the music, or if this violation is a result of ironic contrast between the music and the video, having one's expectations violated may result in a more memorable experience. This complements existing research which has found that ironic contrast can enhance recognition of visual content compared to affectively congruent audio-visual pairings. This suggests that emotion-specific perception drives this process of memory formation in these contexts, as opposed to changes in felt emotion (Damjanovic & Kawalec, 2022).

All Experience components predicted Personal Significance Increase to some extent; however, the strongest predictors were Visual Performance and Meaning of Lyrics. These findings suggest music videos can influence aesthetic judgment and increase absorption, which is an important moderator influencing the strength of emotional responses to music (Sandstrom & Russo, 2013). The Visual Performance component is closely tied to music's affect regulating function, specifically boosting energy levels. However, there is also a cognitive component to this experience, because it helps the individual gain a deeper understanding of the artist's interpretation of the music by watching them perform. This is in line with previous research which has found that being able to see the artist perform influences audience perceptions of a song's emotionality (Vines et al., 2006). The current study extends these findings by applying them in an everyday personal music-listening context and highlights the potential long-term effects that observing an artist perform can have on how one understands and relates with the music. The relationship between Meaning of Lyrics and Personal Significance Increase sheds light on how underscoring

the meaning of the music in the video reduces any ambiguity about what the lyrics are referring to, thereby giving a song more depth. This relationship also highlights how Cohen's (2001, 2013) CAM works in music video contexts; the structural features of the video draw attention to the lyrics, allowing for new associations to form that potentially have more depth and significance compared to whatever prior meaning was associated with the lyrics (if any).

While the study provides interesting insight into music videos and their effects on future listening experiences, there are limitations to consider. The study does not account for whether some experience types occur simultaneously, and if some occur together more often than others. For example, moderately strong, positive correlations were observed between the first three Experience components: Narrative Content, Visual Performance, and Meaning of Lyrics. A single music video may contain content that is both narrative *and* performative and may draw attention to the lyrics in creative ways. However, whether one or more components characterize the experience depends on the individual and their reason for watching. Future research should consider how individual differences contribute to and/or influence music video experiences and outcomes. Individuals who are highly engaged with music and consider it an important part of their lives may identify multiple characteristics of their experience. This would be in line with previous research which has found that individuals who are more engaged with music overall tend to use it to fulfill multiple functions simultaneously (Greasley & Lamont, 2011).

The intent behind the music video experience is an important factor that can potentially influence subsequent outcomes. In the current study, there were two components: Emotional Use and Social Use, which were characterized by the listener's intentions. Both components may act as experience mediators and/or moderators. For example, Emotional Use experiences have a more deliberate function, therefore the video they select should be cohesive with their preferred affect-regulation strategy (Saarikallio & Erkkilä, 2007; van Goethem & Sloboda, 2011). If watching a music video to receive a boost of energy, the individual may lean toward a music video that features energetic performance gestures (i.e., Visual Performance). On the contrary, Social Use experiences do not necessarily have a deliberate goal beyond entertainment. Whether Emotional Use and Social Use experiences mediate the relationships between other Experience components and Retention outcomes should be explored further. Future research should consider individual differences in music listening and social media behavior, as well as individual traits (e.g., gender, musical training, personality, empathy, life satisfaction, etc.) to better understand when, how, and for whom music video experiences yield certain outcomes over others. Future research should consider using experience sampling methods (ESM) to gain more insight into how these experiences influence outcomes (see Randall & Rickard, 2017). More sophisticated methods of statistical modeling are necessary to gain better insight into the relationships between individual differences, experiences, and future listening outcomes. Another limitation of the current study is that participants were not listening to music or engaging with music videos prior to filling in the survey, therefore, the data are subject to recollective recall bias. This could be addressed in future research by either using an ESM, or a controlled experiment using a combination of experimenter and participant-selected music videos to address the question of how attention is directed during music video experiences, and how this is affected by the music.

Conclusion

Music videos are an easily accessible and popular method of music engagement. The current study explored the key characteristics of these experiences and the main effects they impose on future listening experiences. The results posit that music videos can have a salient influence on

mechanisms such as mental imagery, as well as the individual's personal associations with the song. These findings synthesize previous research on the function of music in film and research on everyday music-listening experiences by providing insight into how visual information about a song can affect personal music-listening outcomes. Future research should explore these relationships further and account for individual differences to better understand when, how, and for whom these experiences yield the strongest effects, and whether these effects are perceived as positive or negative.

Contributorship

J.N.D.-W. researched the literature, J.N.D.-W., M.T., and S.S. conceived and designed the study, and J.N.D.-W. wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version.

Funding

ORCID iDs

Johanna N Dasovich-Wilson https://orcid.org/0000-0003-1130-5507
Suvi Saarikallio https://orcid.org/0000-0002-4647-8048

Supplemental material

Supplemental material for this article is available online.

Note

1. https://tenk.fi/en/ethical-review/ethical-review-finland.

References

Aufderheide, P. (1986). Music videos: The look of the sound. *Journal of Communication*, *36*(1), 57–78. https://doi.org/10.1111/j.1460-2466.1986.tb03039.x

Barrett, F. S., Grimm, K. J., Robins, R. W., Wildschut, T., Sedikides, C., & Janata, P. (2010). Music-evoked nostalgia: Affect, memory, and personality. *Emotion*, *10*(3), 390–403. https://doi.org/10.1037/a0019006

Benjamin, J. (2014, July 30). PSY's "Gangnam style" hits 2 billion YouTube views. *Billboard*. https://www.billboard.com/pro/psy-gangnam-style-2-billion-views-youtube-kpop/

Boltz, M. G. (2004). The cognitive processing of film and musical soundtracks. *Memory & Cognition*, *32*(7), 1194–1205. https://doi.org/10.3758/bf03196892

Bullerjahn, C., & Güldenring, M. (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology*, *13*(1–2), 99–118. https://doi.org/10.1037/h0094100

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis (introducing qualitative methods series)*. Sage.

Cohen, A. J. (2001). Music as a source of emotion in film. In P. Juslin & J. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 249–272). Oxford University Press.

Cohen, A. J. (2013). Congruence-association model of music and multimedia: Origin and evolution. In S. Tan, A. J. Cohen, S. Lipscomb & R. Kendall (Eds.), *The psychology of music in multimedia* (pp. 17–47). Oxford University Press.

Damjanovic, L., & Kawalec, A. (2022). The role of music-induced emotions on recognition memory of filmed events. *Psychology of Music*, *50*(4), 1136–1151. https://doi.org/10.1177/03057356211033344

Dasovich-Wilson, J. N., Thompson, M., & Saarikallio, S. (2022). Exploring music video experiences and their influence on music perception. *Music and Science*, *5*, 1–18. https://doi.org/10.1177/20592043221117651

Edmond, M. (2014). Here we go again: Music videos after YouTube. *Television and New Media*, *15*(4), 305–320. https://doi.org/10.1177/1527476412465901

Fiske, J. (1986). MTV: Post structural post modern. *Journal of Communication Inquiry*, *10*(1), 74–79. https://doi.org/10.1177/019685998601000110

Greasley, A. E., & Lamont, A. (2011). Exploring engagement with music in everyday life using experience sampling methodology. *Musicae Scientiae*, *15*(1), 45–71. https://doi.org/10.1177/1029864910393417

Greenwood, D. N., & Long, C. R. (2009). Mood specific media use and emotion regulation: Patterns and individual differences. *Personality and Individual Differences*, *46*(5–6), 616–621. https://doi.org/10.1016/j.paid.2009.01.002

Hargreaves, D. J., & North, A. C. (1999). The functions of music in everyday life: Redefining the social in music psychology. *Psychology of Music*, *27*(1), 71–83. https://doi.org/10.1177/0305735699271007

Juslin, P. N., Harmat, L., & Eerola, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, *42*(4), 599–623. https://doi.org/10.1177/0305735613484548

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. https://doi.org/10.1017/S0140525X08005293

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kistler, M., Rodgers, K. B., Power, T., Austin, E. W., & Hill, L. G. (2010). Adolescents and music media: Toward an involvement-mediational model of consumption and self-concept. *Journal of Research on Adolescence*, *20*(3), 616–630. https://doi.org/10.1111/j.1532-7795.2010.00651.x

Kosslyn, S., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature*, *2*, 635–642.

Marshall, S. K., & Cohen, A. J. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception: An Interdisciplinary Journal*, *6*(1), 95–112. https://doi.org/10.2307/40285417

Meyer, L. (1956). *Emotion and meaning in music*. University of Chicago Press.

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception*, *26*(5), 489–504. https://doi.org/10.1525/mp.2009.26.5.489

Randall, W. M., & Rickard, N. S. (2017). Personal music listening of emotional. *Music Perception*, *34*(5), 501–514. https://doi.org/10.1525/mp.2017.34.5.501

Rentfrow, P. J. (2012). The role of music in everyday life: Current directions in the social psychology of music. *Social and Personality Psychology Compass*, *6*(5), 402–416. https://doi.org/10.1111/j.1751-9004.2012.00434.x

Saarikallio, S., & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. *Psychology of Music*, *35*(1), 88–109. https://doi.org/10.1177/0305735607068889

Saarikallio, S., Gold, C., & McFerran, K. (2015). Development and validation of the Healthy-Unhealthy Music Scale. *Child and Adolescent Mental Health*, *20*(4), 210–217. https://doi.org/10.1111/camh.12109

Sakamoto, Y., & Love, B. C. (2004). Schematic influences on category learning and recognition memory. *Journal of Experimental Psychology: General*, *133*(4), 534–553. https://doi.org/10.1037/0096-3445.133.4.534

Sandstrom, G. M., & Russo, F. A. (2013). Absorption in music: Development of a scale to identify individuals with strong emotional responses to music. *Psychology of Music*, *41*(2), 216–228. https://doi.org/10.1177/0305735611422508

Schäfer, K., & Eerola, T. (2020). How listening to music and engagement with other media provide a sense of belonging: An exploratory study of social surrogacy. *Psychology of Music*, *48*(2), 232–251. https://doi.org/10.1177/0305735618795036

Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Frontiers in Psychology*, *4*, Article 511. https://doi.org/10.3389/fpsyg.2013.00511

Schubert, E., Hargreaves, D. J., & North, A. C. (2014). A dynamically minimalist cognitive explanation of musical preference: Is familiarity everything? *Frontiers in Psychology*, *5*, Article 38. https://doi.org/10.3389/fpsyg.2014.00038

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

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. *Academy of Marketing Studies Journal*, *19*, R244–R260.

Sun, S. -W., & Lull, J. (1986). The adolescent audience for music videos and why they watch. *Journal of Communication*, *36*(1), 115–125. https://doi.org/10.1111/j.1460-2466.1986.tb03043.x

ter Bogt, T. F. M., Vieno, A., Doornwaard, S. M., Pastore, M., & van den Eijnden, R. J. J. M. (2017). You're not alone: Music as a source of consolation among adolescents and young adults. *Psychology of Music*, *45*(2), 155–171. https://doi.org/10.1177/0305735616650029

van Goethem, A., & Sloboda, J. (2011). The functions of music for affect regulation. *Musicae Scientiae*, *15*(2), 208–228. https://doi.org/10.1177/1029864911401174

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80–113. https://doi.org/10.1016/j.cognition.2005.09.003

Wilson, J. N., Thompson, M., & Saarikallio, S. (2020). The function of music videos in everyday listening experiences. *Conference proceedings for Psychology and Music: Interdisciplinary Encounters* (pp. 135–144), October 24–26, 2019. Belgrade, Serbia. https://psychologyandmusicconference2019.wordpress.com/proceedings-pam-ie-2019-conference/

Zhou, R., Khemmarat, S., & Gao, L. (2010, November 1–30). The impact of YouTube recommendation system on video views. In *Proceedings of the for the 10th ACM SIGCOMM Conference on Internet Measurement* (pp. 404–410). Association for Computing Machinery. https://doi.org/10.1145/1879141.1879193

STUDY IV

THE INFLUENCE OF MUSIC VIDEO CONTENT ON LISTENING OUTCOMES: AN EXPERIENCE SAMPLING METHOD STUDY

by

Johanna Dasovich-Wilson, William Randall, & Suvi Saarikallio

The influence of music video content on listening outcomes:

an Experience Sampling Method study

Johanna Dasovich-Wilson[1], William Randall[2], & Suvi Saarikallio[3]

[1]jonawils@jyu.fi, [2]will.m.randall@jyu.fi, [3]suvi.saarikallio@jyu.fi

Abstract

Music videos (MVs) are a popular form of music engagement which allows music listeners to experience their favourite songs in an audio-visual format. MVs have been made more accessible since the rise of platforms like YouTube, giving individuals the control to select MVs over other forms of music engagement. Previous research has shown that MVs can influence the content of visual imagery experienced during music listening. In addition, visual imagery has been associated with modulating arousal levels for musically trained individuals, suggesting a link between musical training and the effect of visual imagery on music induced emotions. This paper features the preliminary results from a pilot study which explored how exposure to MV content influenced enjoyment and emotional reactions to music in everyday listening experiences using the experience sampling method (ESM). 152 observations were collected from 15 (N =15) participants. Linear mixed models (LMMs) were performed, incorporating the participants as random intercepts. No significant effects were observed for enjoyment, however, results suggest that Image Recall, which reflects the amount of MV content recalled as visual imagery during music listening, influences emotional reactions in everyday listening experiences. Image Recall also moderates the effect of changes in both the amount of personal significance placed on the music by the individual, and their interpretation of the music's meaning on emotional reactions. Interestingly, the influence of interpretation on emotional reactions became weaker as the amount of imagery recalled from the MV increased. These results suggest that not only does MV content have a salient influence on the visual imagery mechanism, this influence moderates the strength of effects of other perceptual changes on emotional reactions. These findings support the notion that that visual content, which is readily and freely accessible online, can influence emotional responses to music in everyday life. These findings are discussed in respect to how music educators can use MV and MV-related media to promote better emotional expression in students' musical performances, as well as teach them how to recognize multiple emotions in music and promote cognitive abilities related to mental imagery.

Keywords: Music video, mental imagery, music perception, experience sampling method, emotional outcomes

Introduction

According to the International Federation of the Phonographic Industry (IFPI, 2022), the two most popular methods of music engagement are audio-streaming and video streaming. Music consumers spend over 20 hours a week engaging with music, using six or more different methods of engagement (i.e., radio, audio-streaming, livestreams, etc). This has been facilitated due to the use of smartphones, especially among youth: 94% of 16- to 24-year-olds use their smartphones for music (IFPI, 2018).

Audio-visual recordings of musical performances are an invaluable resource for music teachers and students. Being able to see *how* sounds are produced, whether it be with an instrument or vocally, can highlight the underlying structure of a musical passage (Vines et al., 2011), provide insight into the performer's expressive intent (Davidson, 1993), and influence perception of tone duration (Schutz & Lipscomb, 2007). This information is also useful for ensemble performers by allowing them to better coordinate and synchronise with each other, especially when transitioning between phrases (Boltz, 2013; Schober & Spiro, 2014). The benefits of audio-visual methods of music consumption and their influence on music perception apply to everyday music listening experiences. This has been observed with music videos (MVs), which can affect the perception of musical meaning by influencing what the listener associates with the music and promoting visual imagery in subsequent listens (Dasovich-Wilson et al., 2024; 2022). MVs give listeners access to information about the meaning of the music and often contain scenes of the artist(s) performing. The analyses presented in this paper feature the preliminary results from a pilot study designed with the intent to explore how exposure to MV content can influence on emotional outcomes and enjoyment on subsequent listens to the audio in isolation. It explores how MVs influence music listening outcomes by affecting the visual imagery mechanisms, personal significance, and the interpretation of musical meaning using the experience sampling methodology (ESM). These findings may prove useful for music educators teaching emotional expressivity and emotion recognition in music performances and who wish to implement a culturally relevant and youth-friendly teaching aid in the form of MVs.

Audio-visual Perception

MVs and other musical multimedia which present both audio and visual information are an engaging method of consuming music which requires more attention than audio-only

music listening. Although the content of both the music and video can vary widely in MV contexts, previous research on audio-visual perception and theoretical models outlining the cognitive processes involved provide insight into how MVs may influence the perception of music in subsequent listens. Cohen's Congruence-Association model (CAM; Cohen, 2013; 2001; Marshall & Cohen, 1988) highlights the bottom-up and top-down processes that occur in film (which can refer to both cinematic movies and other multimedia contexts, such as music videos and advertisements). These processes allow individuals to attend to and understand the film's events as they unfold. During bottom-up processing, the structural components of the film and music are identified; this includes the physical features of the film (for example: locations, objects, characters, and dialogue), as well as the structural elements of the music (for example: pitch, tempo, rhythm). These components are grouped together based on their structural similarity. The emotional meaning of the music is transferred to the audio and visual information which attention is directed towards. This transfer relies on top-down inference processes, which allow the individual to make sense of the events unfolding in the film by matching the emotional information from the music with the associated meanings stored in their long-term memory (LTM). The individual's past associations with music are retrieved from LTM and superimposed on the events of the film, allowing them to create a working narrative. These associations are elicited by the structural information of the music and the individual's perception of valence and arousal (Cohen, 2013).

Research has shown that visual information can influence the perception and remembering of the music, as well as enhance the perception of certain musical characteristics. For example, performance gestures influence music perception by providing a clearer concept of a musical passage's phrase structure (Vines et al., 2011). Musicians' movements can infer when a phrase begins and ends, portray musical affect, and foreshadow changes in emotional content (Vines et al., 2006). Furthermore, performance gestures in MVs may be recalled as mental imagery in subsequent listens, and potentially influence how the emotionality of the music is perceived (Dasovich-Wilson et al., 2024; 2022). Narrative information, such as that found in MVs with storylines, can also influence how music is perceived and remembered. Visual information can cause melodies to sound louder, faster, more rhythmic and more active compared to when a melody is presented on its own, regardless of affective quality or whether it is presented as a film clip or montage of still images (Boltz et al., 2009). Boltz and colleagues suggest that the mere presence of visual information can cause music to sound more interesting and/or meaningful, providing

an interpretation of the music's meaning and its underlying message. Even simple narrative descriptions about the music, without visual accompaniment, can influence how music is perceived and influence the content of visual imagery experienced during listening (Vuoskoski & Eerola, 2015).

According to Juslin & Västfjall (2008), *visual imagery* "refers to a process whereby an emotion is induced in a listener because he or she conjures up visual images while listening to the music (p.566)." Visual imagery (also referred to as *mental imagery*) that occurs during music listening may represent a metaphorical representation of the music's structural features and characteristics (Juslin et al., 2014). Adolescents use this mental imagery during listening to explore their personal identity and create a temporary sense of self (Larson, 1995). Emotion regulation through imagery has also been associated with musical training (Küssner & Eerola, 2019). Visual imagery has been associated with the perception of emotions expressed by music and musical training. For example, research has shown students with musical training perform better at tasks involving visual imagery, which is the cognitive ability mentally to create, maintain, and transform images (Commodari & Sole, 2019). This ability can facilitate musical performance: using anticipatory, motor or visual imagery can help students control performance aspects related to timing, intensity, articulation and intonation (Kellar, 2012).

Aims

The study explores the influence of music video content on the *enjoyment of* and *emotional reactions to* music in everyday listening experiences. Three variables describing the mechanisms by which content from an MV affects music perception are explored as potential predictors influencing these outcomes: Interpretation, Personal Significance, and Image Recall. The findings from this study intend to provide insight into how mechanisms underlying everyday music experiences are influenced by visual content, which in turn can affect how the emotional content of the music is perceived and enjoyed. While the study does not consider individual differences in music training as a possible factor influencing listening outcomes, these findings are expected to provide novel insight into how the recall of MV content as mental imagery influences the effects of other mechanisms on music enjoyment and music evoked emotional responses.

Hypotheses

Hypothesis 1 (H1): Personal Significance is expected to have a stronger effect on enjoyment as Image Recall Increases.

Hypothesis 2 (H2): Interpretation and Personal Significance are both expected to have a stronger effect on emotional reaction as Image Recall Increases.

Method

Procedure

The study was completed entirely on the MuPsych application for Android OS. Music experience sampling reports (ESRs) were presented to participants while they were listening to music on their phone. Participants were required to complete a minimum of 20 ESRs over the course of a two-week period, after which they were entered into a draw for one of two €25 Amazon gift cards.

ESRs were presented to participants (via notification) in three sections: the first at the moment of music commencement; another after five minutes of listening, then a final section after approximately thirty minutes of listening. Section one consisted of standard MuPsych items, with the first screen (titled 'How do you feel?') presenting two 7-point sliders, assessing initial *Valence* (labelled at points 1, 4, 7: Very negative–Neutral–Very positive) and initial *Arousal* (labelled at points 1, 7: Very low energy–Very high energy). The dimensions of valence and arousal have been demonstrated to be efficient and reliable measures of music- induced emotion, explaining a high proportion of variance (Vuoskoski & Eerola, 2011).

Section two was presented if participants were still listening to music five minutes later. The first screen assessed *Valence*, *Arousal* and *Intensity* for a second time, with the same 7-point slider format and labels as the initial measures. A screen related to the currently playing music was presented, measuring (all on 7-point slider scales): subjectively perceived *Music Valence* (labelled Very negative–Neutral–Very positive) and *Music Arousal* (Very low–Very high). Following this, participants were asked if they had seen the music video for the song that was playing at the start of section two of the music ESR (responses: Yes, I'm very familiar with it; Yes, I'm somewhat familiar with it; Yes, but I'm not very familiar with it; No I haven't; I'm not sure; No, but I've heard it in a movie, ad, or TV show). If they responded with 'No I haven't' or 'I'm not sure', they were presented with the same question, but in reference to the song playing at the end of section one of the music ESR. If they had

seen either of these MV, the next screen asked: "How has the music video influenced your current listening experience of <Track Name>?", with five questions on 7-point slider scales: "How has your enjoyment of the music changed?" (*Enjoyment*; Enjoyed less - No change - Enjoyed more); "How has your emotional reaction to the music changed?" (*Emotional Reaction*; Not at all - Very much); "Has your interpretation of the lyrics/ music changed?" (*Image Recall*; Not at all - Very much); 'Has the music video made the song more personally significant?" (*Personal Significance*; Not at all - Very much). (*Interpretation*; Not at all - Very much); "Did you remember images/ visual content from the video?" Responses to these MV-ESRs were used in the analysis.

If the participant had not seen the MV of either of the two songs, they were presented with an alternate screen relating to the music, consisting of three 7-point slider scales: "Do any visuals come to mind when you hear this song?" (Not at all - Very much); "What type of memory do you associate with this music?" (Very negative - neutral - Very positive); "How personally significant or connected to this song are you?" (Not at all - Very much). This screen was added in order to keep the number of screens consistent regardless of whether they had seen the video or not, in order to disincentivise participants from indicating that they had not seen the MV - even if they had - in order to complete the number of ESRs required in order to be entered in the prize draw: this data is not analysed in the present study.

Participants

Data were collected from 15 participants (60% female, 20% preferred not to say) between the ages of 18 and 37 ($M = 21.7$; $SD = 4.56$). All participants who completed the demographic survey reported having at least a high school education. The most represented nationality in the sample was Finnish (N=4), followed by Canadian (N=2) and American (N=2). Two participants did not provide data concerning their nationality or education background.

Analysis

The dataset analysed in this paper consists of 152 MV-ESRs across all 15 participants. Two Linear Mixed Models (LMMs) were performed incorporating participants as random intercepts. This approach considers the initial differences between participants and allows for individualized intercepts for each participant. Emotional Reaction and

Enjoyment were entered as dependent variables, whereas Image Recall, Personal Significance and Interpretation scores were entered as predictors. The interactions of Image Recall scores with Personal Significance and Interpretation scores were also investigated in order to explore whether Image Recall acts as a moderator influencing the effect of Personal Significance and Interpretation on each of the two outcomes. In case of a significant moderation, simple effect analysis was performed in order to detect the slope for each predictor (Personal Significance and Interpretation) and its significance level at different levels of the moderator (Image Recall).

Results

H1

Results from the LMM analysis showed no significant influence of any of the three predictor variables on enjoyment. No interactions were observed between Image Recall and either Personal Significance or Interpretation changes occurring as a result of MV content.

H2

Results from the LMM analysis showed a significant effect of Interpretation on Emotional Reaction (b=0.94, SE=0.25, p< .001), indicating that greater changes in interpretation of the music as a result of MV content was associated with greater changes in Emotional Reaction. These effects were significantly moderated by Image Recall (b= –0.11, SE=0.05, p=.028); in that, as Image Recall increased, Interpretation had a weaker effect on emotional reaction (see Figure 1). In more detail, the effect of Interpretation on Emotional Reaction was significant at all levels of Image Recall (p< .001); however, the effect was stronger at low levels of Image Recall (i.e., Mean – 1SD; b = 0.68, SE =0.14, p< .001) and weaker at high levels of Image Recall (i.e., Mean +1SD; b = 0.31, SE =0.08, p< .001).
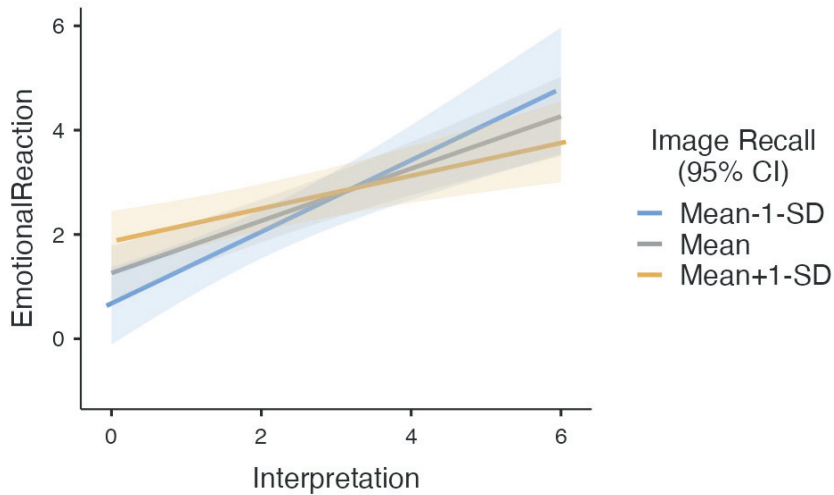
Figure 1. Effects plots for the interaction between Image Recall and Interpretation on Emotional Reaction.

No significant effect was observed for Personal Significance on Emotional Reaction. However, an interaction effect was observed between Image Recall and Personal Significance (b = 0.14, SE = 0.05, p = .006). This suggests that Image Recall has a positive influence on the relationship between Personal Significance and emotional reaction (see Figure 2). This effect was significant for higher levels of Personal Significance (i.e., Mean + 1SD, b = 0.33, SE = 0.09, p<.001).
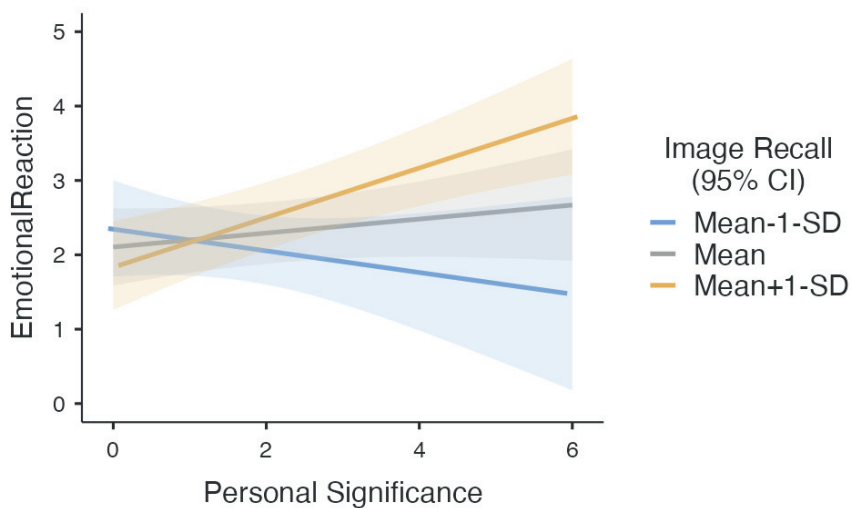

Figure 2. Effects plot for the interaction between Image Recall and Personal Significance on Emotional Reaction.

Discussions

The study confirmed the hypothesis that Image Recall, which reflects the remembering

of MV content as visual imagery during subsequent listens, has a significant effect on Emotional Reaction. However, this influence was not observed for music enjoyment. Image Recall moderated the influence of both Interpretation and Personal Significance on Emotional Reaction, however these effects varied depending on the amount of imagery experienced.

Interestingly, the interactions between Image Recall and Interpretation on Emotional Reactions suggests that changes in interpretation of the music's meaning from MV exposure has a weaker influence on music evoked emotions as the amount MV content recalled as visual imagery increases. These findings support the notion that emotional reactivity is related to the interactions between the music and the images conjured by the listener (Commodari & Sole, 2020; Juslin & Västfjall, 2008; Juslin et al., 2014). Furthermore, these findings suggests that the cognitive processes at work in CAM (Marshall & Cohen, 1988; Cohen, 2013; 2001) are also at work in subsequent listening experiences. In other words, MV content becomes part of the listener's working narrative in subsequent listens in the form of visual imagery from which emotional inferences can be drawn. Therefore, as the amount of imagery increases, the semantic interpretation of MV content is less influential on the emotional response than the imagery itself.

CAM processes may also explain the interaction between Image Recall and Personal Significance on Emotional Reactions. In these cases, the listener's interpretation of the music may not necessarily change but is instead reinforced by the content of the MV. If MV content is line with the individual's existing interpretation of the music, a stronger interaction between visual imagery and the significance of the music on emotional outcomes may occur as a result of those feelings about the music being validated by the video. This would be in line with the findings of Dasovich-Wilson and colleagues (2024) who found that increases in personal significance and the experience of visual imagery in subsequent listens is predicted by MV experiences characterised by a desire to understand the meaning of the music, especially if that meaning was in line with the listener's interpretation. Furthermore, even if the listener has no strong personal connection to the music prior to seeing the MV, content that adds depth to the music or is in line with the listener's moral values has been associated with greater appraisal and stronger connections to the music in subsequent listens (Dasovich-Wilson et al., 2022).

While the results from the study are novel, there are limitations to address. First, musical background was not considered in these analyses. Research on visual imagery conducted on both musically trained and untrained individuals has found that, while both groups experience visual imagery during listening experiences, musically trained

individuals use visual imagery to modulate arousal (Küssner & Eerola, 2019). However, previous research on mental imagery in musically trained and untrained populations has not considered the influence of outside sources like MVs on this mechanism. Future research should account for musical training in particular in order to investigate this influence; since students with musical training have better cognitive abilities in respect to mental imagery (Commodari & Sole, 2020), it is possible that musical background is related to better control over *which* images are recalled from MVs in subsequent listens depending on their reasons for listening. Furthermore, this these analyses are from data collected during the piloting phase of a much larger study, which will also collect additional data about participants' engagement with MVs and musical multimedia in general, as well as survey measures for individual differences such as musicianship, personality, empathy, and use of music (eg., MUSEBAQ: Chin & Rickards, 2014; HUMS: Saarikallio et al., 2015).

Conclusion

MVs are a popular method of consuming music whose extramusical information, in the form of performance gestures and narrative content, can potentially influence the mechanisms underlying emotional reactions in subsequent listens, especially visual imagery. Visual imagery abilities are also associated with musical training and has useful applications in music performance. Therefore, MVs may be useful for music teachers who wish to improve emotional expressivity in students' musical performances which is also a socio-culturally relevant and popular method of consuming music among students.

References

Boltz, M. (2013). Music videos and visual influences on music perception and appreciation: Should you want your MTV. In S. Tan, A. J. Cohen, S. Lipscomb, & R. Kendall (Eds.), *The psychology of music in multimedia* (pp. 217–234). Oxford University Press.

Boltz, M. G., Ebendorf, B., & Field, B. (2009). Audiovisual interactions: The impact of visual information on music perception and memory. *Music Perception*, *27*(1), 43–59. https://doi.org/10.1525/mp.2009.27.1.43

Chin, T–C, & Rickard, N. S. (2014). Emotion regulation strategy mediates both positive and negative relationships between music uses and well-being. Psychology of Music, 42(5), 692713. https://doi.org/10.1177/0305735613489916

Cohen, A. J. (2013). Congruence-Association model of music and multimedia: Origin and evolution. In S. Tan, A. J. Cohen, S. Lipscomb, & R. Kendall (Eds.), *The psychology of*

music in multimedia (pp. 17–47). Oxford University Press.

Commodari, E., & Sole, J. (2020). Music education in junior high school: Perception of emotions conveyed by music and mental imagery in students who attend the standard or musical curriculum. *Psychology of Music*, *48*(6), 824–835. https://doi.org/10.1177/0305735619832962

Dasovich-Wilson, J. N., Thompson, M., & Saarikallio, S. (2024). The characteristics of music video experiences and their relationship to future listening outcomes. *Psychology of Music, 0*(0). https://doi.org/10.1177/03057356231220943

Dasovich-Wilson, J. N., Thompson, M., & Saarikallio, S. (2022). Exploring music video Experiences and Their Influence on Music Perception. *Music and Science*, *5*. https://doi.org/10.1177/20592043221117651

International Federation of the Phonographic Industry (2022). *Engaging with Music.* IFPI. https://www.ifpi.org/wp-content/uploads/2022/11/Engaging-with-Music-2022_full-report- 1.pdf

International Federation of the Phonographic Industry (2018). *Music consumer insight report*. IFPI. https://www.ifpi.org/ifpi-releases-2018-music-consumer-insight-report/

Juslin, P. N., Harmat, L., & Eerola, T. (2014). What makes music emotionally significant? Exploring the underlying mechanisms. *Psychology of Music*, *42*(4), 599–623. https://doi.org/10.1177/0305735613484548

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5). https://doi.org/10.1017/S0140525X08005293

Keller, P. E. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, *1252*, 206–213

Küssner, M. B., & Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 90–99. https://doi.org/10.1037/pmu0000238

Larson, R. (1995). Secrets in the bedroom: Adolescents' private use of media. *Journal of Youth and Adolescence* (24).

Marshall, S. K., & Cohen, A. J. (1988). Effects of Musical Soundtracks on Attitudes toward Animated Geometric Figures. *Music Perception*, *6*(1), 95–112. https://doi.org/10.2307/4028541
Schober, M. F., & Spiro, N. (2014). Jazz improvisers' shared understanding: A case study. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00808

Saarikallio, S., Gold, C., & McFerran, K. (2015). Development and validation of the Healthy-Unhealthy Music Scale. *Child and Adolescent Mental Health, 20(4)*, 210–217. https://doi.org/10.1111/ camh.12109

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*(6), 888–897. https://doi.org/10.1068/p5635

Vines, B., Krumhansl, C., Wanderley, M., & Levitin, D. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80–113. https://doi.org/10.1016/j.cognition.2005.09.003

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to

my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*(2), 157–170. https://doi.org/10.1016/j.cognition.2010.11.010

Vuoskoski, J. K., & Eerola, T. (2011). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, *15*(2), 159–173. https://doi.org/10.1177/1029864911403367

Vuoskoski, J. K., & Eerola, T. (2015). Extramusical information contributes to emotions induced by music. *Psychology of Music*, *43*(2), 262–274. https://doi.org/10.1177/0305735613502373