Author(s): Dasovich-Wilson, Johanna N.; Thompson, Marc; Saarikallio, Suvi

Title: Exploring Music Video Experiences and Their Influence on Music Perception

Year: 2022

Version: Published version

Please cite the original version:

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

Research Article

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 audio-visual 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
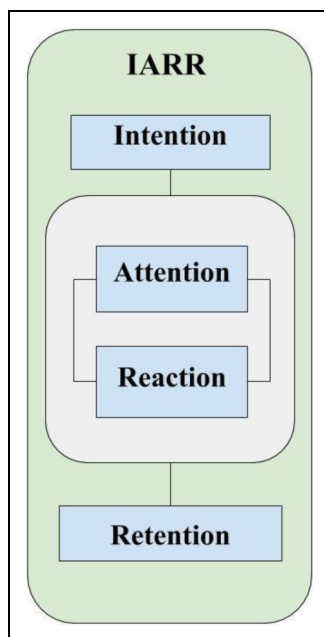
Figure 1. The four temporal levels of MV listening experiences: Intention, Attention, Reaction, Retention (IARR). Intention refers to all antecedent factors leading to MV engagement. The experience of watching MVs is divided into two levels (Attention and Reaction). Retention refers to all carry-over 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
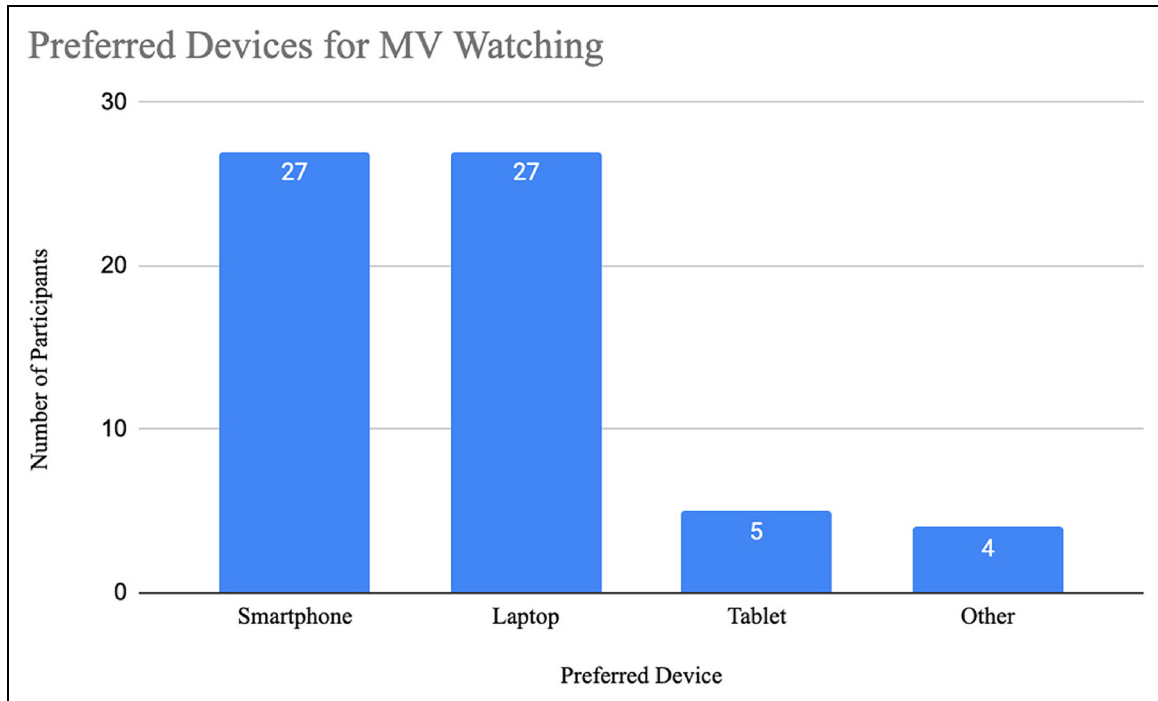
Figure 2. Preferred devices for watching MVs. Participants were allowed to select one of the three options provided, or specify the device if selecting "Other". Other devices were non-portable, such as watching on television, or via YouTube on a gaming console.
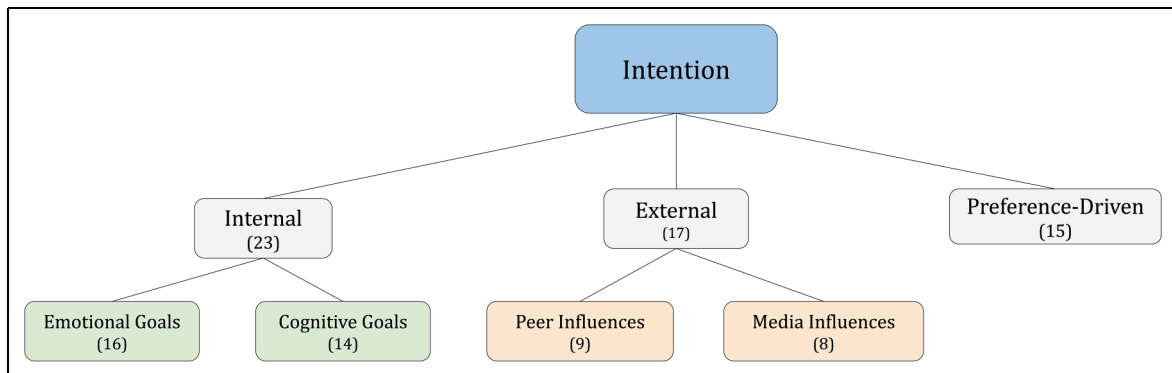
Figure 3. Intentions for watching MVs. The number of participants coded under each category and sub-category are represented in brackets.

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.
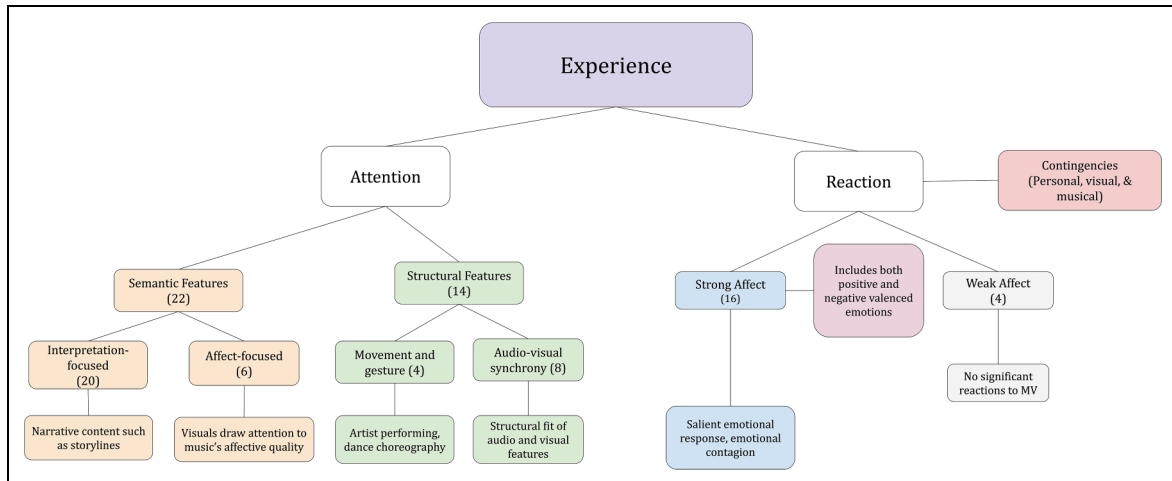
Figure 4. Experience is divided into two levels: Attention and Reaction. The number of participants coded under each category and subcategory are represented in brackets.

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.
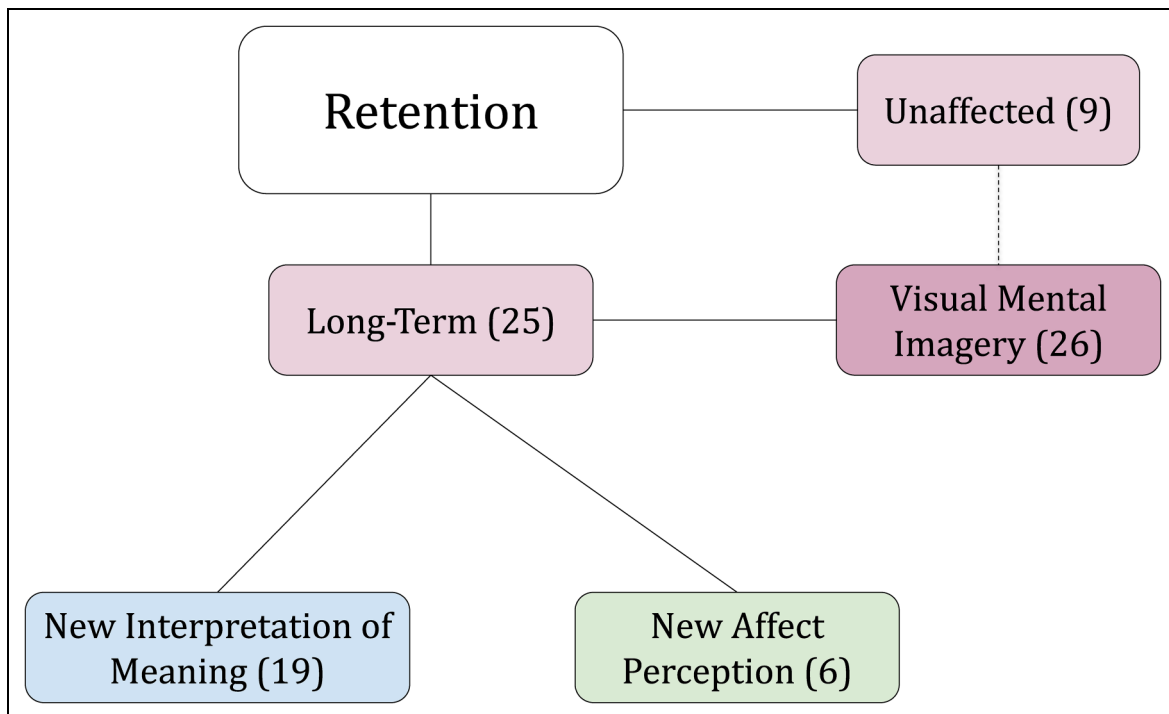
Retention

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 anti-social 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.

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/10.1177/0305735616663313

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/10.1177/0305735693 02100201

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!