

The Impact of Visual Stimuli on Music Perception

Haverford College Psychology Senior Thesis Project

Advised by Professor Boltz

Brittany Ebendorf

Abstract

While many studies have explored music's effect on the perception of visuals, the consideration of the reverse is more limited. The purpose of this experiment was to provide a further and more systematic investigation into the effects of visuals on the affective and acoustical perception of music. Videos and montages pre-rated as positive or negative were paired with unfamiliar and affectively ambiguous tunes to determine if visual affect and format influenced the perceptual processing of music through the analyses of within subject evaluative tasks and ratings for each stimulus condition – four audiovisual and one no visual control. Results indicate that the mood and format of visual accompaniment does affect musical perception on acoustical and emotional dimensions. These findings are both theoretically significant and provide valuable practical implications for such fields as education, marketing, and the music industry.

In today's world, music is everywhere—from the background music at the grocery store to movie soundtracks or even the song being hummed by someone within earshot. This constant exposure to music has a number of important effects upon listeners at both an explicit and implicit level. Experimental evidence suggests that humans are not only sensitive to music, but also associate musical stimuli with themselves. Music has been seen as expressing emotion, faith, beauty, human character, objects and events, motion, and both political and social conditions (e.g. Bharucha, 1994; Krumhansl, 1990). Furthermore, these emotions and personal connections are often experienced with great intensity. Having a meaningful relationship with music does not require one to be a connoisseur of it; mere exposure is often enough.

The power and inescapability of auditory stimuli is surpassed only by visual stimuli, and much research has examined the relationship between these two sensory modalities. Film is a primary example of this relationship. Since the ending of the silent film era, films are seen as incomplete without some form of auditory influence through speech, music, or sound effects embedded in the film itself. While there are many examples of visual information beyond film, film is the most marketable and widespread form of visual stimuli. Auditory stimuli can also come from many sources and influence almost all situations, but music is the main focus of

laymen and researchers alike. Unlike film, music is frequently enjoyed without any added sensory information. Music's ability to stand alone or be paired with other stimuli leads to the question of how the cognitive processing of music is altered, if at all, by the inclusion of visual information. Is the impact of visual stimuli on musical compositions similar to the more established effects of music on film? And, in what ways and under what circumstances does this relationship manifest itself?

The interplay between auditory and visual modalities has created an interesting area of psychological research in recent years, contributing to researchers' knowledge of cognitive processing activities, impression formation, and behavioral responses. Much is understood about the impact of music on visuals such as film, but far less is known about the reverse relationship. It is important to ascertain whether findings in the former area can be directly applied to the latter, or if different processes are at play. The question of how auditory stimuli are influenced by visual information and the theoretical and practical implications of this relationship will be explored throughout this paper.

Musical Aspects

One of the leading psychologists to examine the relationship between visual and auditory domains, Annabel Cohen, suggests that music can be viewed as having two aspects: an affective component and an acoustical or structural component. While these two components are often examined separately, they invariably work together to produce musical effects. People tend to be more conscious of the affective components of music than the structural ones, yet systematic responses to the more implicit acoustical dimensions have also been noted. Cohen (2001) terms the tendency of film audiences to extract music's emotional information while failing to attend to its acoustical aspects 'inattentional deafness,' based on her understanding of Mack and Rock's

(1998) ‘inattentional blindness.’ Since awareness is dependent on attention and one’s attentional capacity is limited, individuals tend to use their resources in the most efficient manner. Thus, the affective aspects of music would be seen as having more real life value and deserving of one’s limited resources than the acoustical. To support this idea, Cohen (2000) cites the Adaptive Resonance Theory proposed by Grossberg (1995) in which information in short-term memory becomes conscious only if it is consistent with predictions about reality. Reality functions as a standard, developed from information in long-term memory about similar events and experiences. Therefore, an individual’s limited mental capacities may cause him or her to acknowledge only the more easily processed affective components of music, though the structural aspects also have an impact. Composers and professional musicians may be the exception to this rule of emotional dominance, as they consciously attend to both musical aspects. Thus within any section of the population, the affective and structural qualities of music, considered alone and as an interactive relationship, play an important role in both the psychological and practical explorations of auditory and visual materials.

The Emotional Aspect of Music

A piece of music has an overall emotional effect as well as specific emotional characteristics. In a study on the experience of music in everyday life, Sloboda and O’Neil (2001) found that music made participants ‘feel better,’ or more positive, alert, and focused. Anecdotally, people are frequently brought to tears or raised to states of exhilaration through listening to music, showing that musical stimuli have the potential for great emotional impact. The affective aspects of music vary greatly in how they are developed, experienced, and acknowledged, but remain a consistent and at least partially conscious influence.

An iconic relationship often develops due to a resemblance between the emotional ‘tone’ of an event or agent, and a certain musical structure (e.g. Dowling & Harwood, 1986). For example, loud, fast-paced music is comparable to loud, fast-paced events. Both evoke high-energy emotions such as excitement. The listener does not need to be consciously aware of the emotional agent nor the musical structure to experience the reaction. Sloboda and Juslin (2001) note how these iconic emotional sources are part of music’s extrinsic effect. The overall patterns of tension and release are similar to those experienced in non-musical contexts. Similarly, music’s intrinsic emotional components play on personal expectations. Each individual note within the piece adds to the overall effect, creating expectancies. These are both learned and innate perceptual processes, such as the results of priming or the ‘gestalt laws’ of perception, respectively, that together provide a predicted musical trajectory. The emotional intensity of a piece of music is highly dependent on the confirmation or violation of these expectancies and how the listener engages with the piece (Meyer, 1956). Music’s extrinsic and intrinsic aspects work together to convey or enhance emotional meaning.

Although the effect of a given piece may not be universal among listeners, commonality is frequently found. Personal experiences, associations and expectancies, as well as contextual factors, come into play. Sloboda and O’Neil (2001) found that music has different emotional functions depending on the context of the exposure. A person’s emotional response to music is a complex combination of the reaction to the material itself, including related associations, and to the social context in which the music is embedded. For instance, a love song such as *Home Again* by Carole King could evoke strong feelings of joy and satisfaction at a wedding, or sadness and longing at a funeral. In this case, the music’s impact results from its structural characteristics and lyrical meaning as well as the context – such as other listeners’ emotions, the

general tone of the event and one's personal mood. Cohen (2001) considered contextual impact on the emotional perceptions of musical stimuli by using various films or pairings of pictures. These studies do not rely on personally charged situations such as weddings or funerals; they use devices like film clips to show the impact of contextual aspects such as subjects, setting and time. It is consistently seen that even constructed visual scenes, such as films, provide contextual information that systematically impacts perception. These findings will be explored in detail later. Additionally, the types of emotions that music can represent are limited. While some emotions such as happiness, sadness and excitement are easily encompassed, a complex emotional concept such as gloating, which is grounded in a cognitive context, is more challenging (Sloboda & O'Neil, 2001). For gloating to be perceived or experienced, there must be a cognitive comparison between one's own situation and that of another, less successful, person. Hence while music can have a powerful emotional impact, it does not necessarily represent all emotional variants.

Systematic measures have been used to study emotions on an empirical level within a musical context. Hevner (1935a, 1935b, 1936, 1937) developed the first of these dimensional approaches to emotions identified from musical stimuli. This model consists of eight adjective clusters of affectively analogous words placed next to similar clusters and across from their presumed opposites. In Hevner's initial studies, descriptive words from the adjective-wheel were used as dependent-measure options for subjects to describe piano pieces that varied along structural lines (e.g. mode, melody and tempo). While Hevner's model is still useful, the more recent Circumplex Model developed by Russell (1980) has become the most widely used dimensional emotional scale (Sloboda & Juslin, 2001). Using three different scaling techniques across multiple experiments, Russell created a valid and reliable system for categorizing

emotions by their valence and arousal level. Emotional adjectives are placed within a four-quadrant space created by a horizontal axis of negative-to-positive valence and a vertical axis of high-to-low arousal (see Figure 1). Although the adjectives used and the relationships between them are similar to Hevner’s earlier model, Russell’s model validates clear bi-polar opposites while using valence and arousal to understand the relationship between different emotions. In this model, although “sad” and “angry” are both negative emotions, they have opposite levels of arousal. Similarly, “frustrated” and “excited” are both highly charged emotions, although opposite in valence. These dimensional emotional scales are frequently used to investigate subjects’ emotional reactions to music by utilizing these adjectives and others in Likert-type scales or descriptive choice tasks. Open-ended questions about the emotional attributes of a musical piece are also typically used.



Figure 1: Russell’s (1980) Circumplex model presents a multidimensional scaling for 28 affective words

The ability to identify and measure the emotional aspects of music is key to further understanding how they form, arise, and create responses. But Gabrielsson and Lindstrom (2001) make the important distinction between a piece of music being perceived as “happy” and a person feeling happy while listening to it. A song may be correlated with a certain emotion without instilling that emotion within an individual. As previously mentioned, the emotional impact may depend on the listeners themselves and the context of the presentation, but the identification of emotional aspects helps to illuminate the cognitive processing of emotions. The perceived, not necessarily experienced, emotion is greatly influenced by the structural aspects of music.

The Structural Aspects of Music

The structure of a musical composition is a key factor in determining its emotional meaning, and is central to understanding the cognitive processing of musical stimuli. Structure consists of acoustical dimensions such as tempo, rhythm, pitch, melody, harmony, mode and loudness. These musical components are implicit knowledge—though a listener may not be consciously aware of them, he or she nevertheless responds to them in systematic ways. Musical pieces are rarely spontaneous or random; instead they are methodically created utilizing these structural components to express the artist’s intentions. When conducting empirical studies of musical structure, researchers rely on three common methodologies for musically experienced and inexperienced subjects alike: free phenomenological descriptions, choice among provided descriptive terms, and the appropriateness of certain terms to the musical stimuli (Gabrielsson & Lindstrom (2001). While these methodologies appeared in the 1930s, recent studies rely on similar yet increasingly complex techniques. To date, a sizable body of research supports the relationship between structural components and emotional qualities of music. The division of

structural components into time, pitch and texture by Bruner (1990) provides a useful way of thinking about the individual and interactive aspects of music's acoustical characteristics. Bruner points out that musical time and pitch appear in almost all of the empirical studies, strongly validating the effects of those characteristics (e.g. Henkin, 1955, 1957; Nielzen & Cesarec, 1982), whereas research on the third factor, texture, is more sparse.

Musical time mainly refers to rhythm, or patterns of durational change over time, as well as tempo—the speed or rate of the music. Gabrielsson & Lindstrom (2001) observe that aspects such as note density or the density of melodic or harmonic changes may influence the listener's perception of tempo. Tempo is also frequently described as the most influential structural factor in emotional expression (Gudlack, 1935; Hevner, 1937; Juslin, 1997; Rigg, 1964; Scherer & Oshinsky, 1977). Similarly, a fast or slow tempo was found to be the most important predictor of happiness and sadness, respectively (Juslin, 1997 Experiment 2). Based on past research, Bruner (1990) concluded that fast tempos are associated with exhilaration, joy, and happiness, whereas slow tempos elicit tranquil, sentimental, or sad reactions (Gundlach 1935; Hevner, 1937; Rigg 1940a; Scherer & Oshinsky 1977; Swanwick 1937; Watson 1942; Wedin 1972). These emotions were evoked at exposure times of less than a second in studies conducted by Peretz and colleagues in 1998.

While these findings are highly consistent, the emotions related to various temporal speeds are also largely dependent on the actual context of the music and the presence of other musical factors. For example, Rigg (1940) found an exception to the trend in a song expressing lamentation and sorrowful longing. In this case, the piece of music was still perceived as sad even at the highest tempo. This study serves as a reminder of the strong impact of context and the interplay between various structural components of music.

Rhythm, the other main component of musical time, also has a number of consistent associations. Gabrielsson & Lindstrom (2001) and Bruner (1990) found across numerous studies that smooth rhythms, featuring a regularly occurring beat over time, trigger emotions related to happiness, dignity, playfulness and serenity. Irregular or rough rhythms, characterized by more unpredictable patterns, evoke feelings of uneasiness and anger. Additionally, varied rhythms or music with many rhythmic changes are often viewed as joyful, whereas pieces with consistent or firm rhythms are judged to be more serious or sacred and associated with sadness and dignity. While findings of tempo and rhythm are often presented separately, as they are here, acoustical characteristics still influence one another. For example, tempo can be thought of as a rhythm's speed, and emotional qualities associated with tempo or rhythm alone may also have an interactive effect. Although both of these factors are influential, Lindstrom (1997) found that only harmonic functions had a greater impact than rhythm on the music's emotional expression as judged by the subjects. These greatly influential harmonic functions, and other pitch-related acoustical qualities, pertain to the second major structural category to be discussed.

While musical pitch is mainly defined by its overall "highness" or "lowness," technically speaking, pitch is the frequency of a sound wave. High pitch is linked to a wide variety of emotions including happiness, serenity, excitement, surprise, anger, and fear. Conversely, lower pitched music suggests sadness, boredom, dignity, as well as excitement and vigor. These findings are somewhat contradictory in terms of valence and activity, but when the pitch variation or range (the magnitude of difference between the highest and lowest notes within a musical tune) is considered, the trends are more systematic. Large pitch variance is associated with happiness, activity and surprise, whereas small variations are tied to boredom, anger, and fear (Gabrielsson & Lindstrom, 2001; Bruner, 1990). Therefore, a lower-pitched piece with a

large range might be interpreted as exciting, and a similar low-pitched tune with a smaller range viewed as sad.

The affective quality of a tune is greatly impacted by other pitch-related variables such as melody, harmony, tonality, and mode. Melodies include factors that are both vertical, referring to the ratio relationships among simultaneously played notes, and horizontal, referring primarily to musical scales that contribute to tonal attributes. Within this context, harmonies are most frequently classified as consonant or dissonant due to the simultaneously played notes (i.e. chords) following lawful or unlawful ratios of pitch, respectively. A number of studies cited by Bruner (1990) and Gabrielsson & Lindstrom (2001) classify consonant harmonies, which are more pleasing to the ear, as happy, playful, or serene, as opposed to dissonant or complex harmonies that are more sad, agitating, or angry (Hevner, 1936; Watson, 1942; Wedlin, 1972). While consonant and dissonant harmonies can be conceptualized as vertical due to simultaneously played notes, melodic attributes occurring horizontally or over time are characterized as the “tonality” of a tune.

Musical compositions can be either tonal or atonal. The former refers to music comprised of notes taken from the underlying scale in which the melody is composed, while the latter contains notes that fall outside the underlying musical scale. When examining the impact of tonality on emotion, Thompson & Robitaille (1992) found that composers use tonal sounds to create joyful, dull and peaceful melodies, and atonal sounds to communicate anger. In Western music, the underlying scale upon which the melody is based is usually described as either major or minor. In 1990, Kastner and Crowder confirmed that by the age of 7, people associate major modes with happiness and joy, and minor modes with sadness. This study provides insight on how mode is linked to perception, and even more importantly, shows that these structural

associations exist from a very young age. Therefore, in a sense, these associations are inherent in the music. The study also found that interpretations of mode are heavily influenced by other contextual factors such as the pitch and loudness of chords.

Texture, the last structural category, is much less understood than the previous two. Bruner (1990) includes timbre, orchestration, volume and its dynamics in this category. Factors related to volume have generated the most interest in the past. Juslin (1997, Experiment 2) found that loudness is the most important predictor of anger. Additionally, loud music is associated with intensity, tension and anger, whereas soft music is perceived as more tender, sad, and fearful. Furthermore, large variations in volume, or amplitude, create more fear than small variations, which are associated with happiness. Lastly, rapid volume changes are experienced as more playful or pleading, whereas less frequent changes are perceived as more sad or peaceful (Juslin 1997).

Numerous reliable trends emerge from the extensive research on music's structural components. Songs that are perceived as happy tend to be fast in tempo, fluid or rhythmic, higher pitched, more tonal, of a consonant harmony, and without major volume changes. Sad songs are associated with the opposite structural characteristics, most notably slow tempi, lower pitches, and greater atonality. While these findings are consistent, context and the interplay between various factors must be considered. Emotional responses elicited by the structural attributes of musical stimuli vary not only by valence (positive or negative), but also in their level of arousal (high or low). The ability of some music to stimulate angry, negative emotions, and other music to evoke serene, positive emotions, makes a dimensional emotional model, such as Russell's (1980) classification of adjectives along affective and activity axes, both applicable and helpful. Despite a slight variance among past studies examining the relationship between

structure and affect, these trends suggest that music's numerous structural elements convey emotional meaning to virtually all listeners from a very young age.

These findings also support Scott's (1990) alternative view of music as language-like, with individual words or musical phrases providing meaning to the overall conversation. Like language, music communicates through the emotions it evokes as well as less consciously via structural elements. Cohen (2001) draws a parallel between the role of music in film and that of prosody in speech perception. The knowledge and use of music's unconscious structural components in relation to its more conscious emotional aspects have also been experimentally explored. For example, musical pieces composed to express specific emotions were characterized by listeners as pertaining to that emotion without any knowledge of the composer's goal. Thus without instructing the composers to use particular musical techniques or directing the listener's attention, this experiment by Thompson & Robitaille (1992) shows that music can be perceived systematically. Additionally, Libscomb and Kendall (1994) found that when subjects were asked to pair various video clips and soundtrack excerpts, they consistently placed them in their original pairings. The correspondence was not necessarily identical, yet both the creators and perceivers saw the affective and structural components as complementary. Therefore, music's affective and structural aspects are influential both when evaluated alone or with a visual pair.

Music and Film

Music has long been an area of interest for social scientists, but it is not always examined in isolation. Indeed, a sizeable body of work describes how music and film interrelate. Before exploring this relationship, a basic understanding of film should be established.

According to Cohen (2000), “In the language of cognitive psychology, cinema is a multisensory stimuli that, millisecond by millisecond, impinges on sensory receptors and excited networks of neuronal activity in the mind of every film spectator.” (p. 360). Film is a visual narrative using people, places, events and their interaction to convey a message, influence emotions, and provide entertainment. Over the years, films have increasingly allowed us to step outside our everyday experience and live vicariously through the lives of characters on the screen. Beyond a spectator’s enjoyment, films also create new schematic knowledge of situations and utilize existing knowledge to convey a message. To the extent that film mirrors real life, schematic information helps people understand the sequence, meaning, and implications of actions taking place on the screen. Silent films were the first to emerge, though live music was quickly added to mask the projector’s loud noise. As technological capabilities advanced, so did the use of audio stimuli – music, speech, and sound effects – to enhance the film. Musical stimuli mainly consist of background music to foreshadow, accompany, or provide ironic contrast within a scene. Foreshadowing occurs when suggestive music is used in an otherwise ambiguous scene to create expectations of what is about occur. Numerous examples of this can be found in horror movies in which unsettling or disturbing music is paired with an otherwise mundane scene, such as cleaning up the kitchen, to create a sense of impending doom. The expectations created by musical foreshadowing can then be confirmed or violated in subsequent scenes to influence the viewer in different ways. Sometimes music doesn’t provide new information to a scene, but rather enhances the visual meaning through accompaniment. Romantic, soothing music will often accompany a love scene, whereas the climactic scene of a great challenge, such as reaching the top of a mountain, will be accompanied by triumphant and dramatic music. Conversely, ironic contrast is created when the affect of accompanying music is

incongruent with that of the visual information, such as a violent montage paired with cheery, upbeat music. This technique can be used for many reasons, including comic relief or social commentary, to produce atypical responses to visual material.

Regardless of how music or any other form of auditory stimuli is used, in today's world no film is complete without engaging both the auditory and optical senses. Cohen (2000) argues that the musical components bring a film to life, noting, "Without music, images are prosaic, mundane, and even lifeless." (p. 261). The well-researched and defined impact of musical stimuli on film will be examined first, followed by the less explored influence of visual stimuli on music. The vast majority of existing research examines the impact of auditory stimuli on visual material without considering the reverse. Thus, it is unclear to what extent or through what processes visual stimuli may influence musical perception. The limited research suggests that a similar relationship exists, but further consideration of the available data as well as additional experiments are necessary to confirm this.

Music's Impact on the Perception of Visual Information

Cohen (2000) claims that music adds considerable value to film, as it engages faculties other than those required for processing and remembering visuals. Drawing on her own and other researchers' work, Cohen has identified eight main functions of music in a film or multimedia context: to cover extraneous noise, create consistency between shots, heighten the sense of reality, add an aesthetic effect, guide attention by structural or associationist congruence, provide further meaning especially within ambiguous scenes, generate mood induction, and draw on associations in memory to connect thematic events (Cohen, 2001). The last four of these functions speak to the cognitive impact of music on film and will be further explored by examining the role music plays in comprehension (including attention and processing) and

emotionality. Music's impact on film and visuals will also be considered in a marketing context, mainly through its ability to activate and strengthen associations. Lastly, Cohen's multimedia model will be examined to address the various interactive influences of music and film.

In addition to speech, film incorporates atmospheric auditory stimuli such as background music, and sound effects that mimic reality. It may seem paradoxical that both unrealistic sounds (the former) and realistic ones (the latter) are incorporated within the visuals. Yet both types of auditory stimuli act as preattentive steps that guide the viewer throughout the film. As music guides attention and influences both interpretation and memory, it can be thought of as a schema. Boltz (2001) was the first to show the added schematic value of music within films. Schemas provide an interpretive framework that draws on past personal experiences in similar situations and the understanding of social norms. By activating and engaging schematic information, music helps the viewer understand the sequence of events, the motivation behind them and what might come next. Using three ambiguous Hitchcock film clips paired with positive, negative, or no musical accompaniment, Boltz (2001) showed that subjects' performance on character ratings and object recall tasks varied significantly, and in a mood-congruent fashion, due to musical stimuli. For instance, participants exposed to clips with negatively-valenced music rated the characters' actions, motivations, and personalities more negatively than the control group that viewed the same clip without any music. Therefore the comprehension of the film was directly affected by the music, suggesting that it played a vital role in how the viewer processed and made sense of the visuals. Prior to this, the schematic value of music was not directly considered, though music's ability to provide subtle attentional, processing and comprehension cues, especially in ambiguous scenes, had been explored.

Earlier work focused on the role of association and structural congruence in aiding the interpretation of film. It is important to note that these concepts and the idea of music acting as a schema are not mutually exclusive and actually work together to help a viewer make sense of the stimuli. The associative value of music allows for a direct transfer of its meaning to the film's content, setting, or plot. Structural congruence, which often evokes associative principles, directs attention to corresponding visual information – i.e., uplifting music enhances uplifting visual information (Cohen, 2005). In 1975, Vinovich provided early support for the impact of musical mood on video comprehension. Different interpretations of the same ambiguous video led Vinovich to conclude that subjects used the emotionality of music to create a predictable cognitive interpretation of the scene. Similarly, Cohen (1993) found that a video of ambiguous male-female interactions was viewed as either romantic or aggressive depending on the musical accompaniment. Musical mood, however, had little impact on an unambiguous fight scene, suggesting that music is most influential when viewers need help making sense of the visual cues. These studies, and numerous others, provide strong support for music's role in directing attention, processing, and comprehending films, particularly when visual information is ambiguous or incomplete.

On a physiological level, Manfred Clynes (1975, 1977, 1980; Clynes and Nettheim, 1982), a music and neurophysiology scholar, found that the brain processes corresponding to emotional responses are activated by appropriately structured music. In support, Thayer and Levenson (1983) demonstrated that music has a direct impact on viewers' internal arousal levels. As subjects were exposed to films of industrial accidents paired with manipulated musical scores or no music, their physiological responses varied according to musical pairing – stressful horror music increased physiological activity, while relaxing documentary music decreased

electrodermal responses. Additionally, a number of studies conducted by Alpert and Alpert (e.g. 1990, 2005) show that musical effects in commercials are transferred to the viewer, causing changes in mood that closely resemble the music.

Music not only influences film by affecting the viewer's emotions, but it also influences the evaluation of the film and characters in a mood-congruent fashion. Marshall and Cohen (1988) found that personality ratings and the interpretation of geometric-shaped characters' actions in a film were dependent on the musical soundtrack. The stimuli used in this experiment included a large triangular character that was pre-rated to be aggressive and agitating relative to smaller geometric figures in the scene. The negative evaluations of the large triangle increased with the addition of louder and more active music, whereas the inclusion of softer and weaker music produced a less unfavorable evaluation of the large triangle. Interestingly, the inclusion of music that was independently rated as positive did not cause the geometric figures to be seen as positive per se; instead, the intensity of the negative evaluations was lessened. While these findings suggest that musical stimuli have an emotional impact on character ratings, they do not show a direct transference of musical meaning onto visual information. It may be difficult to translate human emotions to inanimate objects through music, yet better-suited pairings produce larger evaluative impacts. For example, clips of ambiguous interactions among wolves were viewed as friendly or combative depending on the musical accompaniment (Bolivar, Cohen, & Fentress, 1994). Thus, music's effect is two-fold—influencing viewers' emotions and their interpretation of the film to varying degrees.

The joint representation of musical and visual information has been used for marketing tactics based on the principles of classical or Pavlovian conditioning. This associative learning theory suggests that after repeated exposure to a neutral stimulus paired with an affectively

charged stimulus, the once neutral stimulus evaluated independently will elicit a response characteristic of the other stimulus' affect (e.g., an audiovisual pairing in which the song is neutral and the visual is positively charged leads to positive evaluations of the music even in the absence of the visual information). Gorn's 1982 study employed this principle to condition product choice through pairings with appealing or unappealing musical stimuli. Gorn reported that the same product, a pen, paired with appealing music during initial exposures led to a more positive response to it later, even without the pleasant music. However, the validity of these findings was brought into question when Kellaris & Cox (1989) were unable to replicate Gorn's findings across three similar studies.

The Elaboration Likelihood Model (ELM), offered by Petty & Wegner (1999), can help psychologists and marketers understand how music impacts mental processing. The ELM describes a central and a peripheral route to persuasion at opposite ends of an elaboration continuum. The subject's motivation and ability to engage in evaluating the presented stimuli influences the degree of elaboration. For example, an individual who greatly values higher education viewing a video about a prospective college would most likely engage in central route processing. This route is used when the person is highly motivated to utilize all applicable and available cognitive resources as a means of attending and evaluating the stimulus information. However, low motivation or low ability triggers the peripheral processing route that relies more on heuristic cues or stereotypes. Park & Young (1986) found that in advertising, music is most effective when peripheral processes are engaged, and relatively ineffective with central processing. Similarly, Bruner (1990) noted that music has less impact on decisions that require high cognitive engagement, such as decisions about purchasing a car or insurance. But when

consumers are experiencing low cognitive involvement or high affective reactions, such as with beer or cosmetic products, music may be more influential.

While the ELM is somewhat helpful, the Congruence-Associationist Model presented by Cohen (e.g. 2000, 2001, 2005), provides one of the most valid and comprehensive models for understanding how music and film interrelate. Cohen's model assumes that comprehending the visual narrative is the primary goal of film-watching. This goal is achieved by utilizing limited short-term and long-term memory in conjunction with sensory information. Auditory stimuli in films thus help to guide interpretation of visual cues through associations created by inter-domain structural congruence (Cohen, 2005).

Cohen's model addresses the significant domains of film – music, speech, sound effects, film images and written text – through five parallel channels of stimulus processing. Each channel includes four levels of processing, and both simultaneous bottom-up and top-down processes that meet at the level of conscious attention and short-term memory. Drawing on data across all domains, bottom-up processes utilize structural and associationist information, whereas top-down processes activate long-term memory from preattended information. The first two phases interpret the sensory system's response to and processing of the stimuli in a domain-specific fashion, respectively. In the second phase, structural characteristics and associative meanings can lead to cross-domain interactions if there is sufficient overlap. In such cases, a visual domain conveying a similar message to an auditory domain is attended to more and remembered better. For example, if background music that is tonal, slow in tempo and fluid in its rhythm accompanies a peaceful ocean scene, the viewer will pay more attention to the visual cues. A working narrative is created by combining the cross-modal information from the bottom-up processing and top-down processing based on past experiences. This unified

representation of all domains results in a conscious understanding of the film. Congruency of stimuli and the viewer's emotional state influence the attentional processes, utilization of short-term and long-term memory and understanding of the visual narrative. Cohen's framework gives key insight into the empirically supported influence of musical stimuli on film's cognitive processing. Yet, no model to date considers the impact of visual stimuli on the cognitive processing of music. This reverse effect will be addressed next.

Visual Stimuli's Affect on Musical Perception

In contrast to studies investigating the effects of music on film, there is a relative lack of research on the influence of visual stimuli on musical perception—perhaps, in part, due to visual dominance. The subject is still of great interest, however. Its further exploration could address a number of issues, including: Can visual stimuli influence a listener's cognitive or affective response to music? How are the affective qualities of music altered, if at all, by visual stimuli? Does visual information act as a distracter, enhancer, or both in the emotional and cognitive processing of music? Are people's musical tastes affected by visuals?

One of the most consistent findings in this area is that including visual information with musical stimuli heightens the emotional interpretation and positive evaluation of the auditory information. Two different studies have shown that visual stimuli improve a subject's emotional response and understanding of the message. First, Adams (1994) measured the responses of musicians and non-musicians to symphony performances. Subjects' emotional responses after exposure to the performance were gauged from one of three conditions – music only, music plus video, or video only. Regardless of the subject's musical background, the first two conditions provoked higher emotional responses. Furthermore, the musicians regarded the music plus video condition as significantly more emotional than the other two conditions. This shows that the

inclusion of visual information does in fact enhance music's emotional meaning, but why might this be so? Davidson (1993) suggests that the visual cues associated with musical styles enhance the experience. Davidson instructed his skilled musicians to alter the expressive qualities of a piece by using one of three styles – deadpan, projected, or exaggerated. The projected style attends to various aspects of the music equally, whereas the deadpan and exaggerated styles constitute two extremes. The former style stresses note accuracy while minimizing other acoustical characteristics, while the latter dramatically overstates these characteristics. Results showed that interpretation of the musical styles was dependent on the inclusion of visual information. The visual only presentation group was better able to predict which expressive style the musicians were utilizing than the audiovisual or music only subjects. The inclusion of visual stimuli – taped performances – added emotional value and understanding.

Across several performance evaluation studies, the inclusion of visual stimuli improves the perceived musical ability and quality. Musical experts' evaluations of the same violin and viola performances were rated more positively when presented in an audiovisual vs. audio only format (Gillespie, 1997). Using five scales – amplitude of vibrato, speed, evenness, pitch, stability, and overall sound – a panel of experts rated audiovisual performances of inexperienced players individually and again one month later using only an audio tape of the performance. The procedure was repeated six months later with performances by experienced players. While the trends varied between the two types of performers, Gillespie consistently found improved ratings when visual accompaniment was included despite the fact that the actual musical performance was the same at both evaluation times.

Yarbrough & Hendel (1993) found similar results in the evaluation of high school choral teachers. Elementary and high school students used a 100-point scale to evaluate teaching

performance based on one of three presentation formats. Subjects were exposed to either a real-life rehearsal through an audiovisual clip, visual only clip, audio only clip, or simply a written “teaching script” (i.e. lesson plan) that varied on several dimensions. The audiovisual presentation and script-only conditions received higher ratings than either the visual or audio presentation alone. Neither stimulus produced as positive an evaluation as when presented in tandem (or in the script-only case, imagined as happening together). Lastly, a study of a disabled-persons’ choir by students and adults using the Florida Music Educators Association choral adjudication guidelines produced similar results (Cassidy & Sims, 1991). Besides the evaluator’s age, the between-subject factors included the performers’ disability label and the type of performance presentation (either audiovisual or audio only). Not only did the inclusion of visual information enhance musical evaluation, but the more information provided, such as a disability label, the more favorably the musical performance was rated. When considered as a whole, this body of research suggests that the inclusion of visual stimuli greatly enhances both the evaluative and affective dimensions of music. Visual stimuli do in fact affect certain forms of musical perceptions and, in some circumstances, can enhance musical ratings.

Studies that consider the interplay of visual and auditory stimuli in music videos provide further insight. Unlike studies utilizing musical performances, which presumably have high levels of consistency between the two stimuli, music videos draw on visuals with differing affective and structural qualities, congruence, and elaboration needs. A sociological study conducted by Hall and colleagues (1986), found that young people prefer music videos that emphasize the storyline, and visual imagery that aids in interpreting the lyrics—providing further support for visual enhancement when congruency is present. Furthermore, Blanchard-Field and colleagues (1986) found that subjects characterized rock videos as more ambiguously structured

than television dramas, which are fairly thematic. The more abstract rock videos may require greater cognitive involvement when there is a lack of thematic or congruent information from the audiovisual stimulus. Drawing on the ELM discussed in the marketing section, this suggests that visual stimuli, as found in music videos, can enhance the likelihood of central route processing if the viewer is motivated to understand the message. Moreover, Goldberg and colleagues (1993) found that musical wear out (the amount of time it takes to lose interest) is forestalled in music videos. Music videos that lacked closure, or whose narratives were left open to individual interpretation, were more successful at preventing wear-out than music without any visual stimuli. While the idea of “keeping it simple” (Roman & Maas, 1976) may be effective if there is minimal exposure, the positive cognitive elaboration associated with lack of closure in the video was posited as the mediator in the cases of high exposure explored by Goldberg and colleagues. Once again, this study suggests that visual accompaniment’s ability to encourage central processing increases the impact of musical stimuli. The ability of music videos to increase cognitive elaboration implies that appropriately constructed music videos can lead to heightened understanding and evaluations, both desirable effects for the music industry. These studies also provide general support for certain aspects of the ELM and congruency models to this area of study.

Geringer and colleagues (1996, 1997) ran the two most comprehensive studies to date on the cognitive and affective impact of visual stimuli on music. They addressed whether the listeners’ musical experience or response – affectively and cognitively – would be changed by the inclusion of video imagery. In the first study, subjects were randomly assigned to a music plus video or music only condition. All subjects were exposed to two musical pieces – Bach’s *Toccato and Fugue in D minor* and Dukas’ *The Sorcerer’s Apprentice* – and all video excerpts

were taken from the Walt Disney movie *Fantasia*. For the music plus video conditions, Bach's piece was accompanied by an abstract or non-programmatic clip of *Fantasia*, and a narrative story or programmatic clip was paired with Dukas' music. Subjects in the music only condition listened to the same two tunes, without any excerpts from *Fantasia*. After a three-minute stimulus exposure to each piece, participants were asked, in two open-ended questions, to detail their emotional experience of the music, film or both, and which portions of the music caught their attention. Likert-type scales captured the affective responses in terms of likeability, degree of involvement in the stimuli, degree of emotionality experienced, and degree of feeling versus thinking involved. Lastly, a cognitive task evaluated accuracy of the subjects' perceptions of musical characteristics – tempo, melody, harmony, meter, texture, instrumentation and dynamics – in a multiple-choice format, following five short musical excerpts of both tunes.

A main effect of presentation was found only for the Dukas piece; subjects in the audiovisual group that used a narrative movie clip performed higher on cognitive measures and gave more positive ratings than subjects in the music only condition. No main effect was noted for the Bach piece paired with an abstract video nor were any reliable differences found between the audiovisual and music only conditions. These findings suggest that visual information alone may not enhance cognitive processing of musical stimuli but, instead, the content of the visual information may determine the performance levels. Thus the type of video may be a mediating factor in the impact of visual stimuli on music perception. In addition, the answers to the two open ended questions suggest that the audiovisual groups felt a stronger emotional connection to the musical stimuli than the audio only conditions. In these tasks, the audiovisual conditions relied more heavily on descriptions of the visual stimuli and emotional meaning derived from them, with little mention of any musical information, whereas the music only conditions were

more attentive to musical characteristics (the only stimuli they were given). These findings could be interpreted to mean that visual stimuli distract from the music, due to visual information bias in audiovisual subjects' free responses. Yet this conclusion is not supported; the subjects' preference for describing visual information does not mean that they did not attend to the musical information (the subjects in the Dukas audiovisual condition performed highest on the cognitive tasks). It's more likely that visual information is easier for subjects to describe in words and more salient due to visual dominance.

In a subsequent study, Geringer and colleagues (1997) added a video of the Vienna Philharmonic Orchestra to the visual stimulus material, as well as the narrative sections of *Fantasia* used in the earlier study. Utilizing two movements of Beethoven's Symphony No. 6 in F major as the musical stimuli, this study had four conditions – one orchestra video clip, one with an animated video clip, while the remaining two groups had no visual stimuli. The procedure and dependent measures used in this experiment were virtually identical to the study run one year earlier (Geringer et al., 1996).

The lowest performance on the various cognitive tasks was observed in the animation treatment groups. For the first movement, the performance video condition performed highest on the cognitive tasks. And, although there was no significant difference in affective ratings between groups, both the animation and the first movement conditions consistently received higher affective ratings. The results of the second study further reject the notion that visual information paired with music always acts as a cognitive distracter. The consistently higher affective rating of the more melodic first movement across all groups shows that the emotional responses are tied more to visual than musical cues. Lastly, ELM can explain the consistently lower performance of the animation groups on cognitive tasks, despite high scores

on likeability measures. When in a pleasurable state, such as that produced by the animated visual stimuli, subjects tend to be less scrutinizing and rely more on heuristics, which mitigates the need for high cognitive elaboration and peripheral processing. This supports the 1996 study's supposition that including visual stimuli enhances musical perception, and further illustrates that different types of visual stimuli work toward different goals. For example, performance videos enhance cognitive awareness, whereas animated videos increase the positive emotions associated with musical stimuli. Although the findings from these studies (1996 and 1997) corroborate the previously established trends, many questions remain. Future studies should address issues such as the potential familiarity of subjects with pieces like *Fantasia*, and ensure that adequate information is given about the cognitive tasks. The relatively small effect sizes of most findings also suggest the need for further empirical work.

Given the research on visual material's impact on music, it is clear that while many of the same processes are affected when looking at the interplay with music as the central focus, further research needs to address these issues directly and in a methodologically sound manner. And even though most studies have suggested that visual stimuli have the potential to enhance the perceptual and cognitive processing of musical material, a better understanding of the mechanisms at play would help maximize this relationship for applications such as music education, music videos and marketing in general.

Current Study

Cohen's (2005) congruence-associationist model considers music's influence on the cognitive processing of film – including its effects on the emotional impact of a scene (e.g.

Cohen, 1993), elaborative inferences about the characters' motivations and actions (e.g. Marshall & Cohen, 1988), and how memorable a scene may be (e.g. Boltz et al., 1991).

Due to the primary assumption of visual dominance over auditory stimuli, the congruence-associationist framework focuses on the impact of music on film perception, but does little to address the reverse effect. The existing research addressing this issue is sparse, but suggests that the reverse effect may in fact exist. For example, under specific circumstances, visual information appears to enhance the emotional aspects of musical stimuli (e.g. Adams, 1994), increase cognitive elaboration (e.g. Goldberg et al., 1993), and aid performance on perceptual and memory tasks (e.g. Geringer et al., 1996, 1997). Various studies have shown that visual stimuli systematically influence the cognitive processing of musical information, and that the visual stimulus best suited for a tune is somewhat dependent on the specific goals of that pairing (such as education or entertainment). To date, the research of Geringer (1996, 1997) provides the most direct analysis of this issue by examining how different types of visual stimuli (animated narrative, animated abstract, musical performance) influence performance on subsequent perceptual and cognitive tasks. While this literature provides greater understanding of visual information's influence on music, methodological concerns exist. The use of vague dependent measures, potentially familiar musical and visual stimuli, and affectively charged stimuli render general conclusions difficult. . For example, an individual's previous exposure to the stimuli could interfere with results, due to pre-formed associations with the stimuli affecting cognitive and emotional perceptions. Additionally, the stimuli vary on affective dimensions, which makes a systematic understanding of the relationship challenging. A more rigorous test of the impact of visual stimuli on music would rely upon relatively neutral or ambiguous musical stimuli paired with film clips pre-judged as positive or negative. This in turn would allow one to

examine whether different emotions are associated with the same piece of music when accompanied by differing film clips. If so, then the music should receive different ratings on affective and acoustical dimensions found to convey emotional meaning. For example, a given melody may be perceived as having a smoother, more regular rhythm and faster tempo when accompanied by a positive visual scene, and a more irregular rhythm and slower tempo when accompanied by a negative visual scene.

If film does in fact influence music, then the role of visual characteristics should be questioned. Specifically, does the structural format of the accompanying visual information also influence music perception? Visual stimuli can differ in format along several dimensions – moving vs. still images, narrative vs. abstract, animation vs. video, realistic or related visual accompaniment vs. unrealistic or unrelated visuals. Some of these dimensions have been explored in past studies (e.g. Blanchard-Fields et al., 1986; Geringer, 1996, 1997), but more research would substantiate findings. One pertinent question is whether a thematic and affectively charged video has a different impact than a positive or negative montage made of still images. A positive video and a positive montage may influence cognitive and affective perception differently. For example, the montage may be inherently choppy than the video's moving images. If this rougher visual montage is transferred onto the music, then the music may be perceived as less rhythmic and thus more negative than a similarly charged video. The music education field, as well as the entertainment and advertising industry, would benefit from the discovery of systematic formatting differences.

In sum, despite the past emphasis on music's impact on visual stimuli, it is also likely that visual information influences music such that a film's mood and structure lead to different perceptions of the same song. This idea is implicitly understood in the music industry, where

artists use videos to help market their songs. Musical and audio stimuli are of primary concern in the film industry, and rightly so considering the impact of various audio techniques. But are visual stimuli like music videos central to the music entertainment industry? Taking into account the growing trend of hit songs with accompanying music videos, the answer is yes. But it's also possible that visual information distracts from musical perception. This idea cannot yet be confirmed or denied. But through rigorous methodology that isolates possible variables and systematically examines structural and affective perception, the present research aims to extend and enhance the existing body of work.

These issues are addressed in the present experiment by presenting subjects with five ambiguous tunes paired with visual scenes that vary in format (video, montage, or no visuals) and affective valence (positive or negative). After the presentation of each pair, subjects are asked to evaluate various affective and acoustical characteristics of the melody, as well as certain dimensions of the visual scenes. The primary hypothesis is that ambiguous songs paired with positive visual stimuli will receive more positive affective ratings and be perceived as displaying more structural components associated with positive affect (e.g. faster tempo, greater tonality, more regular rhythm), relative to ambiguous songs paired with no visual stimuli (control group). The reverse effect is predicted for songs paired with negative visual material. A secondary hypothesis is that visual format (montage vs. video) may also influence perceptual evaluations of the stimuli. For example, a positive montage pairing may be evaluated as less positive (e.g., more choppy) than a positive video pairing.

Methods

Design and Participants

The design was a (5 x 2 x 3) x 10 mixed factorial. All participants listened to five mood-ambiguous songs paired with video scenes that varied in their overall affect (positive or negative) and format (visual montage, coherent video story, or no visual accompaniment). The single between-subjects factor was Set, which counterbalanced each song's presentation order and pairing with each of the visual stimulus options (including the no visual control) across the ten Sets.

Forty-six participants received course credit for an introductory psychology course at Haverford College for completing the experiment. Each had normal (or corrected) vision and hearing. Of the participants, thirty-four were female and twelve were male. One male chose to leave in the middle of the experiment due to his reaction to the stimuli. This incomplete data set along with four others were not included in the final analyses due to rating complications. Subjects were randomly assigned to one of the ten Sets to ensure that there were at least four complete sets of usable data for each condition.

Stimulus Materials

Musical Stimuli

Five, ninety-second affectively neutral and structurally ambiguous musical excerpts obtained from various Internet databases (e.g., <http://music.download.com/>; <http://www.music-free-download.net/>) were used as experimental stimuli. The final musical stimuli were selected using a pre-test of six subjects, in which each provided three descriptive adjectives and Likert-type familiarity, affect, and arousal ratings for twenty-eight, thirty-second musical excerpts. The five tunes chosen from this pre-test displayed unfamiliarity, affective ambiguity, and moderate

arousal (See Appendix A for a complete list of final tunes and artists). Furthermore, all musical stimuli were instrumental (without lyrics), of the neoclassical genre, and originally written in a major mode with a regular rhythm, high degree of tonality, and a moderate tempo with an average beat duration of 0.7 beats per second, considered to be affectively neutral according to past literature (e.g., Gabrielsson & Lindstrom, 2001).

Visual Stimuli

Montages: For the ninety-second montage visual accompaniment stimuli, two positive and two negative versions were created from intermixed picture sources. Images were taken from the International Affective Picture System (IAPS) (Center for the Study of Emotion and Attention, 2001) and an independent database compiled from Film and Musical Soundtrack research at Haverford College (Boltz, 2004). The IAPS images were selected from a CD containing over 700 still images pre-rated for affect and arousal using 9-point Likert-type scales, with 5 as a neutral score (Lang et al., 2001). The IAPS images that were selected as experimental stimuli were moderate in arousal (average scores between 3.5 and 6.5) and strongly positive or negative in their affect (greater than 5.5 and less than 4.5 for positive and negative respectively). Photos selected from the Haverford database had been pre-rated using 11-point Likert-type arousal and valence scales (from -5 to 5) and followed similar selection criteria as the IAPS photos. All images had been rated between -2.5 and 2.5 on the arousal scale; positive photos were above 2 in affect whereas negative images received ratings of -2 and below. Within each of the four montages, the affect of images is always the same, all positive or negative, and level of arousal is matched both within and between conditions. Negative images included photos of natural disasters and human suffering, while positive images featured flowers, appetizing food, and happy people. Sibelius, a musical notation program, was used to determine

the average beat duration of each tune in order to best synchronize the auditory and visual information. The montage speed was set at 0.7 seconds per picture for all tunes. Each montage featured approximately 130 pictures, with a total of over 500 pictures from the two databases combined.

Videos: Four unfamiliar video clips were created from existing DVDs, two negative and two positive, all ninety seconds long to match the duration of the music. In a pre-test, seven subjects rated fifteen, thirty-second video clips in terms of familiarity, affect, and arousal. These fifteen clips were selected from sources such as *Baraka* and *Chronos* by Ron Ficke; *Koyaanisqatsi* and *Powaqqatsi* by Godfrey Reggio; and *Darr: A Violent Love Story* by Yash Chopra. The final four video clips were selected based on their unfamiliarity, desired affective charge, and moderate levels of arousal. Three of the four final clips were taken from *Baraka* and the remaining clip was from the Bollywood film *Darr*. Nature scenes, such as flowing waterfalls and tranquil bodies of water, were the central focus of positive video A. Positive video B was a romantic scene between a man and woman from an obscure Bollywood film. Version A of the negative videos showed a chicken-processing factory with chicks being branded; version B featured homeless people in an urban setting (See Appendix B for complete video clip descriptions).

Musical and Visual Pairings

All stimuli were the same length, so that each of the five ambiguous tunes could be paired with each visual stimulus. A total of ten different sets were created. Each set included all five tunes – one paired with a negative montage (A or B), one with a positive montage (A or B), one with a negative video (A or B), one with a positive video (A or B), and one with no visuals (control group). Across the ten sets, melodies were paired with different visual stimuli so that all

possible combinations occurred in a randomized presentation order (see Appendix C for a complete set list).

Apparatus

iMovie and iTunes software, manufactured by Apple, was used for video editing and synchronizing videos with their accompanying soundtracks. This software was also used to keep amplitude constant across all five tunes. To determine the presentation speed of images in the montages, Sibelius (manufactured by Avis Technologies, Inc.) was used to ascertain the average beat duration. During the actual experiment, paired stimuli were presented at a constant volume on a set of four Dell (Dimension 4500) computers containing a DVD player and individual headphones to prevent any interfering noise.

Procedure

Participants were randomly assigned to one of the ten stimulus sets and tested in groups of two to four participants. Prior to the experiment, subjects were asked to sign a consent form and told that they would be making perceptual ratings on a series of visual and auditory stimuli. Participants were then given unlimited time to read and ask questions about the written instructions that described the experiment's procedure and the set of ratings they would be making. After the presentation of each song, subjects were informed they would be prompted to pause the DVD using the space bar, after which they would have an unlimited amount of time to complete stimuli ratings sets A-D. In rating set A, participants evaluated the stimuli (music alone or paired with visual information) with measures using 7-point Likert-type scales of likeability, affect, and arousal (see Appendix D). Next, in rating set B, subjects were asked to select any of the twenty-eight randomly arranged adjectives that described the stimuli. These particular adjectives were selected from past studies (Hevner, 1937; Russell, 1980) to ensure that

each adjective could be characterized in terms of affect (positive or negative) and arousal (high or low); there were seven adjectives within each of these four groupings. The complete list of adjectives and their respective quadrants is found in Appendix E. Upon completion of the adjective choice task, subjects were instructed to complete rating set C, comprised of six 7-point Likert-type scales designed to assess cognitive perceptions of the musical structure in terms of tempo, rhythm, tonality, volume, and harmony. The vocabulary used in these structural bi-polar adjectives was geared toward non-music experts; more complicated musical characteristics, such as tonality, were described in written instructions at the onset of the experiment (see Appendix F for the complete musical structure scales). Lastly, participants completed rating set D, which applied only to the four stimulus presentations that contained visual information. This rating set included 7-point Likert-type scales that measured the perception of the visual stimulus alone (theme, likeability, and affect), as well as in conjunction with its musical pairing (degree of audiovisual fit) (see Appendix G for the visual scales). These latter measures were included to provide a stimulus check and ensure that both the visual and musical material would be attended to. After completing all applicable ratings, subjects re-started the DVD by pressing the space bar to move on to the next stimulus. They repeated this process until all five stimuli were individually rated. The experiment lasted approximately thirty-five minutes. At the end, subjects were thanked and reminded to not discuss the experiment with potential participants. Subjects were debriefed via e-mail once the study was concluded.

Results

The set of behavioral measures assessed in this experiment can be categorized into three subgroups: stimulus measures that required participants to evaluate the audiovisual (or audio alone) presentations in a holistic fashion; musical measures that directed attention to the songs'

acoustical or structural dimensions; and visual measures that assessed visual attributes independently and in relation to the musical pairing. Each is discussed in turn.

Stimulus Dependent Measures

The first two ratings sets (Sets A and B) asked participants to complete Likert-type scales and an adjective choice task for each stimulus. For the four audiovisual conditions, the measures refer to the musical and visual information as a cohesive unit, while in the no visual or control condition they refer to the musical stimuli alone, since the subjects listened to the music while watching a blank computer screen.

Adjective selection: The adjective-choice task consisted of a list of twenty-eight adjectives from which the subjects were asked to select those that described the stimulus. The adjectives varied on two dimensions – mood (positive or negative) and arousal (active or passive) – creating four adjective quadrants. Seven words from each quadrant were presented in a random order for the subjects to choose from while rating set B. This classification system and the adjectives used are based on the work of Hevner (1935a, 1935b, 1936, 1937) and Russell's Circumplex Model (1980) discussed earlier. The primary issue of interest was not the individual adjectives chosen, but which adjectives in terms of affect and arousal were associated with each stimulus. Therefore, results are shown as percentages of the mean total of adjectives selected out of a possible seven for each quadrant. An overall 5 (experimental condition) x 4 (quadrant) repeated measures factorial was conducted to obtain the mean percentages for each of these 20 cells, and the MS_{error} value required to conduct Tukey's HSD post-hoc comparisons (HSD=.82). This was followed by a second ANOVA that relied on a 2 (mood) x 2 (format) x 4 (quadrant) repeated measures factorial (omitting the control group) in order to assess any potential main effects or interactions among these variables due to mood and format.

This latter ANOVA revealed two significant interactions. First, Table 1 depicts the adjective choice frequencies for the significant interaction between mood and adjective quadrant [F(3,120)=220.9, p<.0001]. The exposure to positive visual stimuli resulted in a greater percentage of adjectives selected from positive quadrants, and a higher frequency of negative adjectives in the negative visual information conditions. This is one indication that, as suggested by the visual stimuli pre-test and database image ratings, the visual mood manipulations were successful in eliciting their intended affective responses.

Table 1
The Percentage of Adjectives Selected for Positive and Negative Audiovisual Pairs as a Function of the Four Affect/Arousal Quadrants

	Positive Visual Information	Negative Visual Information	<i>Adjective Quadrant Mean</i>
Positive/Active	44.4%	2.4%	23.4%
Positive/Passive	39.1%	4.0%	21.6%
Negative/Active	3.4%	49.4%	26.4%
Negative/Passive	2.6%	51.4%	27.0%
<i>Affect Mean</i>	22.4%	26.99%	

All values are significantly different from the mean, p<.05.

Second, a significant three-way interaction between mood, format, and quadrant [F(3,120)=20.51, p<.0001], shown in Table 2, suggests that the arousal dimension of the chosen adjectives depends on format (i.e. video or montage). Specifically, subjects used more positive/active words to describe positive montages than positive videos, although both formats yielded more responses than either the control no visual condition or the two negative formats. For the positive/passive quadrant, adjective choices indicated that montages are perceived as more active than videos, since a significantly lower percentage of the passive/positive adjectives were used to describe the positive montages than the positive videos. Furthermore, a significantly higher proportion of positive/passive adjectives was used to describe the positive

video condition as compared to the other audiovisual groups – positive montage and both negative conditions. In both of the negative quadrants, negative montages received significantly more adjectives from both arousal dimensions compared to the other four conditions. While the positive visual conditions did not differ significantly from the control in their frequencies of negative/active adjectives, both negative visual formats had significantly greater frequencies than the control. In the negative/passive adjective quadrant, however, all visual conditions differed significantly from the no visual condition along affective lines with the negative conditions receiving higher selection frequencies than the positive. Important format differences emerged within the negative quadrants similar to what was seen within the positive quadrants – montages were more frequently characterized by active adjectives and less by negative adjectives. Specifically, the negative/active adjectives were used more often to describe negative montages than videos, whereas negative videos showed higher selection percentages within the negative/passive quadrant than the montages (although this trend did not reach significance).

Table 2
The Percentage of Adjectives Selected for Positive and Negative Audiovisual Pairs as a Function of the Four Affect/Arousal Quadrants and Two Visual Format Conditions

	Positive Video	Negative Video	Positive Montage	Negative Montage	No Visual (Control)
Positive/Active	31.7%*	1.0%	51.1%*	3.9%	12.6%
Positive/Passive	49.9%	7.3%*	28.6%*	0.7%*	53.7%
Negative/Active	4.6%	34.1%*	2.4%	64.9%*	3.3%
Negative/Passive	3.9%*	48.1%*	1.4%*	54.7%*	16.0%

Asterisk (*) denotes significant difference from the control, $p < .05$.

Ratings of Stimulus Likeability, Affect, and Activity: In addition to the adjective-choice task, subjects responded to Likert-type scales measuring the likeability, affect, and activity of each stimulus in rating set A. This set of means, shown in Table 3, was obtained through an independent set of ANOVAs. In each case, a one-way repeated measures statistical design with

five layers was initially conducted to assess performance in the control group relative to the experimental conditions. This was followed by a 2 X 2 repeated measures factorial (excluding the control group) to determine any potential main effects of interactions due to format and mood. Any significant effects that emerged were further evaluated through a set of Tukey's HSD post-hoc comparisons ($p < .05$).

Table 3
Mean Stimulus Ratings of Likeability, Affect and Activity as a Function of Visual Stimulus Affect and Format

	Positive Video	Negative Video	Positive Montage	Negative Montage	No Visual (Control)	Statistical Results
Likeability 1=like, 7=dislike	3.00	5.20*	2.35*	6.05*	3.03	F(4,120)=85.068, p<0.0001; Tukey's HSD=0.66
Affect 1=positive, 7=negative	2.60*	5.95*	1.53*	6.53*	3.40	F(4,120)=179.490, p<0.0001; Tukey's HSD=0.61
Activity 1=active, 7=passive	2.60*	5.95*	2.55*	2.60*	4.10	F(4,120)=61.520, p<0.0001; Tukey's HSD=0.73

Asterisk (*) denotes significant difference from the control, $p < .05$.

Relative to the control group, both negative formats were significantly less well liked while the positive montages received significantly higher likeability ratings [$F(4,120)=85.068$, $p < 0.0001$]. There was no significant difference between the positive video condition and control group in terms of likeability. As expected, the secondary ANOVA revealed a significant main effect of mood on how likable the stimulus was; positive audiovisual material was liked

more than negative audiovisual material [$F(1,30)=146.05$, $p<.05$]. No effects of format were observed. No effects of format were observed.

The expected pattern of results was also found for the affective ratings of the stimulus [$F(4,120)=179.490$, $p<0.0001$]. Compared to the control groups, positive stimuli conditions across both formats received significantly more positive ratings while negative stimuli conditions received significantly more negative ratings. This significant main effect of visual stimulus mood on the perceived affect of the overall stimulus [$F(1, 30)=575.15$, $p<.0001$] was confirmed in the subsequent 2 x 2 ANOVA and was an important measure of the experiment's construct validity. Once again, video format exerted no significant effects.

In terms of the activity level of the stimuli, all audiovisual conditions differed significantly relative to the control group [$F(4,120)=61.520$, $p<0.0001$]. Both positive conditions and the negative montage condition received higher activity ratings than the control, whereas the negative video condition was perceived as significantly more passive. The secondary 2 x 2 repeated measures ANOVA revealed a main effect of format [$F(1,30)=85.17$, $p<.0001$], which supports the adjective-choice trend in which montages were perceived as significantly more active as compared to videos. A main effect of mood [$F(1,30)=80.23$, $p<.0001$] showed that positive stimuli were perceived as more active than negative stimuli on the 7-point Likert-type scale. Lastly, a significant interaction between format and mood [$F(1,30)=79.24$, $p<.0001$] showed a greater difference between formats within the negative as opposed to the positive visual conditions. The negative video condition therefore received the most passive ratings of all. Overall, the stimulus activity rating corroborated the adjective selection data in that montages were perceived as more active than videos.

Collectively, the analyses of both the adjective-choice and Likert-type ratings affirmed the construct validity of the experimental stimuli used in this study.

Musical Dependent Measures

Past literature addressing the relationship between music's emotionality and structural components helped formulate predictions for the current study (e.g. Bruner, 1990; Gabrielsson & Lindstrom, 2001). The primary hypothesis was that the mood of the visual stimulus would influence perception of the ambiguous music's structural qualities. It was expected that tunes paired with positive visuals would be perceived as having a faster tempo, more regular rhythm, better flow, higher tonality, more harmonious quality, and higher volume than in the no visual condition. The reverse effects were predicted for conditions paired with the negative visual stimulus. A secondary hypothesis considered the role that visual format plays in affecting musical perception. For example, montages may be inherently choppy than videos, skewing the music's perception by associating it with negative acoustical qualities. These acoustical dimensions were assessed using Likert-type scales in rating set C and were analyzed using the same statistical procedures as the Likert-type stimulus dependent measures. These findings are related with reference to Bruner's (1990) division of musical structure into the components of time, pitch, and texture (see Table 4).

Table 4
Mean Ratings of Acoustical Dimensions as a Function of
Visual Stimulus Affect and Format

	Positive Video	Negative Video	Positive Montage	Negative Montage	No Visual (Control)	Statistical Results
Tempo 1=fast, 7=slow	4.10	4.39	3.34*	3.37*	4.59	F(4,124)= 9.45, p<0.0001; Tukey's HSD=0.71
Tonality 1=tonal, 7=atonal	2.80	2.85	2.54	2.49	2.90	No Significant Differences
Harmony 1=harmonious, 7=discordant	2.83	2.73	2.56	3.02	2.54	No Significant Differences
Rhythm 1=regular, 7=irregular	2.32*	2.46*	1.83*	1.98*	2.59	F(4,124)= 3.09, p<0.05; Tukey's HSD=0.11
Flow 1=flowing, 7=choppy	2.66	2.59	2.95	3.15*	2.39	F(4,124)= 2.34, p<0.05; Tukey's HSD=0.74
Loudness 1=loud, 7=soft	3.88	3.88	3.40*	3.60	4.00	F(4,124)= 2.44, p<0.05; Tukey's HSD=0.59

Asterisk (*) denotes significant difference from the control, p<.05.

Temporal Qualities: Musical time refers primarily to a tune's tempo and rhythm. For tempo, both the positive and negative montages were perceived as significantly faster than the no visual control, while the positive visual material showed no significant differences in ratings

[F(4,124)=9.45, $p < 0.0001$]. Furthermore, the 2 x 2 ANOVA, a main format effect, showed that montages, regardless of mood, were rated as having faster tempi than both video conditions [F(1,31)=29.2, $p < .05$]. No effects of mood were observed.

All visual conditions differed significantly from the control group in terms of perceived rhythm [F(4,124) = 3.09, $p < .03$]. Regardless of mood or format, the inclusion of visual information resulted in higher rhythmic ratings. Further analyses through the 2 x 2 ANOVA showed significant main effects in both mood and format [F(1,31)=2.5, $p < .03$ and F(1,31)=10.9, $p < .002$, respectively]. In contrast to the other audiovisual conditions, positive stimuli were perceived as having more regular rhythms than negative stimuli, and montages were perceived as more regular than videos. When flow, an additional temporal music dimension, was assessed as a dependent measure, only the negative montage condition differed significantly from the control, being rated as choppy [F(4,124)=2.34, $p < .05$]. Again, a main effect of format emerged in which montages were perceived as significantly less flowing, in other words choppy, than videos [F(1,31)=6.08, $p < .02$]. No effects of mood were found.

Pitch Qualities: Musical pitch was examined via the pitch-related dimensions of tonality and harmony, neither of which yielded any significant effects through the overall ANOVA's. The means in all experimental conditions were comparable to one another.

Texture Quality: Loudness represented the final structural category of musical texture. Compared to the no visual control, positive montages were rated as significantly louder [F(4,124)=2.44, $p < 0.05$]. Additionally, in a main effect of format, videos were perceived as softer than montages regardless of affect [F(1,31)=5.68, $p < .02$]. Once again, no effects of mood were observed.

Overall, the musical dependent measures strongly supported both hypotheses that mood and format of the visual information would significantly affect the perception of ambiguous music. Although this effect was more strongly evident for some measures than others, the findings suggest that visual material significantly alters a tune's perceived temporal and textural acoustical qualities as a result of visual format and, at times, mood.

Visual Dependent Measures

In reference to the four audiovisual conditions, rating set D allowed for further assessment on the construct validity of the stimuli, but also ensured that participants attended to visual as well as audio information throughout the experiment. The no visual control condition did not warrant any response on this fifth and final rating set, hence a 2 x 2 repeated measure factorial ANOVA was used for each Likert-type rating to examine effects due to stimulus mood and format. Similar to the other dependent measures, a set of Tukey's HSD post-hoc comparisons ($p < .05$) was used to further assess any significant effects. The mean ratings for the four visual measures are shown in Table 5.

Visual Affect: The overall ANOVA revealed a significant main effect of mood on perceived affect for the visual stimulus [$F(1, 40) = 194.39, p < .001$], which confirmed the expectation that positive stimuli would be rated as significantly more positive and negative stimuli as more negative. A significant interaction between visual format and mood, [$F(1,40)=21.95, p < .001$] indicated that positive montages were rated significantly more positive than videos, while negative videos received more positive ratings than montages. Overall, the montage stimuli had more extreme affective ratings – more negative in the negative mood condition and more positive in the positive mood condition – than their affective video counterparts.

Table 5
Mean Visual Stimulus Ratings as a Function of Visual Stimulus Affect and Format

	Positive Video	Negative Video	Positive Montage	Negative Montage	Statistical Results
Visual Affect 1=positive, 7=negative	2.46	5.97	1.73	6.10	<i>Mood Main Effect:</i> F(1,40)=194.39; <i>Format x Mood:</i> F(1, 40)=21.95
Visual Likeability 1=like, 7=dislike	2.78	5.00	2.24	5.93	<i>Mood Main Effect:</i> F(1,40)=87.95; <i>Format x Mood:</i> F(1, 40)=20.64
Visual Theme 1=meaningful 7=unmeaningful	3.85	2.37	2.93	2.02	<i>Mood Main Effect:</i> F(1,40)=35.26; <i>Format Main Effect:</i> F(1, 40)=8.29
Fit of Visual and Music 1=well, 7=poor	3.05	2.68	3.29	3.88	No Significant Effects

All main effects and interactions reported are significant at the $p < .05$ level.

Visual Likeability: A similar pattern of results was found for visual stimulus likeability. Once again there was a main effect of mood [F(1,40)=87.95, $p < .001$] in the predicted direction; positive visuals were liked significantly more than negative ones. A significant interaction between format and mood [F(1,40)=20.64, $p < .001$] also showed that for positive stimuli, montages were liked more than videos, but for negative stimuli, videos were liked more than

montages. In other words, the montages were both more liked and disliked than the video conditions.

Visual Theme: Mean scores of the visual theme for all four conditions, from most to least meaningful, were as follows: negative montage, negative video, positive montage and lastly, positive video. Both mood and format produced main effects in the thematic ratings of the visual material [$F(1,40)=35.26, p<.001$ and $F(1,40)=8.29, p<.001$, respectively], although no interaction occurred. The main effect of mood indicated that negative stimuli were perceived as more meaningful than positive stimuli. Furthermore, the effect of format showed a higher thematic meaning for montages compared to videos.

Audio Visual Fit: Finally, no significant effects were found in the ANOVA assessing the fit between the visual and musical information (see Table 5). Since none of the conditions varied significantly from one another, this finding indicates that all audiovisual pairings were of equal fit— which eliminated concerns about certain visual manipulations of tunes being better suited for certain pairings. All of the means were closer to the “very well” Likert-anchor rating than the other extreme of “very poor.” This shows that no pairings were more suitable than others; each was seen as at least somewhat appropriate, while none were perceived as misfits.

In sum, these visual dependent measures not only provide further construct validity of the stimulus material and pairings, but also a more detailed description of how the different visual stimuli were perceived.

Discussion

For researchers and non-researchers alike, the complex interplay between music and film has been of great interest. Past research has identified many reliable effects of music on visual

stimuli in terms of cognitive processing (e.g., Cohen, 2005), comprehension (e.g., Cohen, 2003), emotional response (e.g., Thayer & Levenson, 1983), and even memory functioning (e.g., Boltz, Schulkind & Kantra, 1991). Explorations of the reverse relationship assessing how visual information influences musical stimuli are sparse and inconclusive (e.g., Geringer et al., 1996, 1997). The present experiment aimed to examine the issue of how music is influenced by visual information in terms of emotional and cognitive processing. The findings both converge with and extend the existing literature by showing that the perception of music is not only influenced by visual accompaniment (the primary hypothesis), but that specific visual qualities (i.e. format) also affect perception (the secondary hypothesis). Analysis of the dependent measures supports the role of visual information in how music is affectively and structurally perceived and the impact of the visual format on perceptual processes. This research confirms the validity and potential benefits of further explorations into visual stimuli's influence on music, which could lead to a whole new level of understanding cognitive processes. This also has direct implications for societal interests such as entertainment, marketing, and education.

Stimulus Checks

In order to make these claims and consider the theoretical and practical implications in greater detail, construct validity must first be established across numerous dependent measures. The use of ambiguous tunes, which could be interpreted by subjects as either negative or positive, paired with affectively manipulated visuals provided a methodologically rigorous means in which to assess the effects of visual information on musical perception. Many of the dependent measures confirmed that the stimuli – ambiguous tunes and positively or negatively charged visuals – were in fact perceived in the manner they were pre-tested and designed to be interpreted.

Visual stimulus mood manipulation was confirmed by the affective ratings of the audiovisual pairs and visual material alone. Compared to the no visual control group, the inclusion of positive or negative visual information elicited significantly more positive and negative ratings, respectively. The ambiguity of the musical stimulus was also validated through the stimulus affect rating of the no visual conditions (resulting in a music-only measure) that were perceived to be more neutral (mean score of 3.40 on a 7-point scale) than the audiovisual conditions. The music-only condition also received a more neutral likeability rating (mean score of 3.02 on a 7-point scale) than all but the positive video condition. The ambiguous quality of the musical stimuli was also exemplified by the ratings of audiovisual fit in which no pairing was viewed as either extreme – very well or very poorly. Across the experiment, each tune was paired with each visual condition and yet a strong-to-moderate visual and tune fit was consistently found (i.e. receiving an average rating of 2.68-3.88 where 1=fit well on a 7-point scale). This shows that none of the pairings were particularly disconcerting to the participants and the lack of extreme responses proves that this was in fact due to the tune's ambiguity and not a matched affect between the visual and musical information.

One way in which the visual mood can influence the musical stimulus is by providing interpretive value through thematic meaning. Interestingly, while both positive and negative stimuli were thematically meaningful in terms of mean comparisons, negative stimuli were significantly more meaningful than positive. These ratings may reflect the nature of the images chosen, in that negative images of human and animal suffering, violence, and natural disasters, carried a stronger message than positive visuals like beautiful scenery and expressions of pleasant feelings. The relative strength of the negative visual information may be connected to an adaptive survival or defense mechanism in which an attentional bias toward negative or

potentially threatening stimuli provides information relevant to self-preservation (e.g., Öhman, 2002).

The adjective-choice task corroborates the stimulus check from the affective ratings, in that mood manipulations elicited adjective selections matching their intended affect. Furthermore, montages were more frequently characterized by active adjectives, whereas videos elicited more passive adjectives. Stimulus activity measures showed the same results. These format effects are probably due to the rapid influx of stimuli in the montages – approximately 130 images changing every 700 milliseconds – as compared to the slow-moving images of the video. Activity measures also displayed a main effect of mood in which participants saw positive visuals as more active and negative visuals as more passive. This trend is similar to past findings that correlate positive emotions with fast tempo and upbeat, rhythmic music (e.g., Bruner, 1990; Gabrielsson & Lindstrom, 2001). Across the various dependent measures of this study, the validation of stimulus manipulations and the emergence of predictable trends provided the necessary construct validity to ensure the soundness of the study's significant findings and implications.

Visual Mood's Effect on the Interpretation of Music: Exploring the Primary Hypothesis

Compared to past literature addressing the relationship between auditory and visual stimuli, the present research was primarily concerned with whether or not visual material influences the perception of music in a systematic fashion. This study offers some of the first substantial evidence that the previously well-documented ways in which music influences visual assessment can work in reverse – with visuals affecting music. Specifically, the primary hypothesis explored the influence of affectively charged visual stimuli on the perceived emotionality of naturally ambiguous tunes.

While the stimulus dependent measure unequivocally showed that the inclusion of mood-manipulated visuals affected the overall ratings of the stimulus, the acoustical ratings were the only direct measures of musical perception within audiovisual pairs. Extensive previous research has substantiated the link between specific acoustical dimensions and emotional perceptions, finding that fast tempo, high degrees of tonality and harmony, regular rhythms, uninterrupted flow, and greater loudness are associated with positive emotions, while the reverse qualities are associated with negative emotions (e.g., Scherer, 1977; Bruner, 1990; Cohen, 2000; Gabrielsson & Lindstrom, 2001). Relying on ratings of these affectively salient musical dimensions, the current study found that visual information had a significant affect on the perception of music.

Two of the temporal qualities – rhythm and flow – supported the predicted visual mood effects. Positive visual conditions were perceived as being significantly more rhythmic than their negative counterparts. Furthermore, compared to the control group, negative montages conveyed significantly less sense of musical flow – a positive mood characteristic. That only the negative condition and not all montages had this effect suggests that it was the visual stimulus' mood that led to the choppy or more negative perception of the music. It is also possible that the inherent choppiness of the montages compared to videos may have enhanced the emotional impact of the negative montage audiovisual pairs (this idea will be further explored in considering the secondary hypothesis regarding the overall format effect).

Loudness, the only textural quality, displayed similarly supportive effects. Despite the consistent presentation volume, tunes paired with positive montages were evaluated as significantly louder than the control group. The association between a positive acoustical trait

and positive visual reached significance for the montage format alone, suggesting a combined influence of visual mood and format similar to the effect of flow on music perception.

The emergence of systematic effects of visual accompaniment was clearer for these temporal and textural qualities as compared to the two pitch dimensions – tonality and harmony – which revealed no significant effects. One potential explanation is that subjects had more difficulty grasping the more technical nature of these qualities and the terminology used (despite the definitions provided in the instructions). Additionally, past literature notes that temporal dimensions have some of the strongest associations with emotionality (e.g., Juslin, 1997; Lindstrom, 1997). Therefore both experiment-specific and generalized aspects of the acoustical qualities probably contributed to the effects observed. An alternative explanation, which would need further testing, is that there is an attentional bias toward pitch-related dimensions that makes evaluating them through accompanying stimuli less effective overall. Future research could investigate musical attention trends in the context of visual accompaniment to explore whether or not individuals engage their limited cognitive resources in a systematic fashion with regards to musical perception.

Regardless of the interpretation of the null effects, it should be noted that there were no significant effects that contradicted the primary hypothesis. Considering the novelty of this area of research, this represents a promising step toward a systematic understanding of visual stimulus' effect on musical perception. As seen in the flow and loudness measures previously discussed, the format or structural characteristics of the visual stimulus are also important in understanding the effects of visual accompaniment. This exploration of the effect of visual mood on music not only provided support for the primary hypothesis, but also showed that visual affect may not be enough to explain the impact of visual accompaniment. The secondary

hypothesis, which considers visual format, is also essential to developing a context in which to interpret the current work.

Visual Format's Effect on the Interpretation of Music: Exploring the Secondary Hypothesis

The secondary hypothesis was confirmed across numerous dependent measures, showing that different types of visual information produce different effects on music. It was predicted that inherent differences between the two visual formats – consistently changing randomized images within the montages and flowing cohesive video stories – would create different effects of visual accompaniment. This hypothesis was directly supported by the main effects of format found for all acoustical ratings, with the exclusion of the null effects of pitch qualities discussed above. Across these musical measures, montage audiovisual pairs were significantly louder, choppy, more rhythmic, and faster than video conditions—regardless of their mood manipulation. Although these acoustical trends point in different affective directions (all positively skewed except for the negative association with flow), findings strongly support the hypothesized impact of structural visual characteristics on music perception.

Hypothetical explanations for how format effects are found would help determine how visual influence works. Future studies can test the effect of structural elements more directly by holding the emotionality or substantive material constant across various formats. For example, if montages were created using still images from the stimulus videos, then the format elements could be isolated. Still, the validity of the mood manipulations in the present study allows for preliminary observations about the significant ways in which varying visual formats can affect music.

Montages were probably characterized as more “choppy” – the only negatively associated musical dimension – as a result of the abrupt shifts from one unassociated (beyond

mood congruence) image to another in the montage conditions, as compared to the streaming video images. The constantly changing montage visuals are also likely responsible for the increased rhythm regularity and tempo ratings. The presentation speed for the montages was synchronized to the average tune-beat duration (700 milliseconds), which may have highlighted and standardized the rhythm, while the video stimuli were created by extracting the original audio but leaving the remaining elements in their original form. The montage speed also allowed for the display of approximately 130 images in the same ninety-second period during which the video portrayed a limited range of visuals. This may have led to the perception of faster tune tempos for montage pairings.

These acoustical trends and explanations corroborate format effects found for activity (Likert-ratings and adjective choice distribution), as the structural properties of the experimental montages are inherently more active than videos. Lastly, the increased amount and variance of visual information provided in the experimental montages may have been crucial in eliciting significantly more thematic meaning from the montages than the videos. While the photos did not offer a cohesive story, there were greater opportunities for the subject to create their own meaning than with the more narrowly defined themes of the video clips that led to this format effect. As a set, these findings support the hypothesis that it is not just the affect of visual information that significantly affects a tune's perception, but also the characteristics of the visual format.

Theoretical and Practical Implications

The present findings both converge with and extend previous research examining the influence of visual information on music behavior. The experiments of Geringer and colleagues (1996, 1997) are most similar to the present design, although improvements were made, such as

the use of verifiably unfamiliar and ambiguous stimulus material and more clearly defined cognitive tasks. The present analyses of the stimulus dependent measures corroborated one of the central findings of Geringer and colleagues, that visual plus musical conditions were more emotionally meaningful than music only conditions. Additionally, the current work supported past findings that participants presented with audiovisual performance clips or music-only stimuli performed better on cognitive music tasks than subjects in the programmatic audiovisual condition (*Fantasia* video clip). The tendency toward neutrality on acoustical ratings for the music-only control group suggests that this group performed better than the audiovisual conditions because of the ambiguity of the experiment tunes. Although this perceptual difference only reaches significance for rhythm, the trend lends support to past conclusions that visual information can have a distracting effect on music perception. By this logic, the inclusion of visual accompaniment can have an enhancing effect if the visual stimulus is constructed to provide information relevant to the musical goals. This idea is supported by the existence of the reverse relationship—that music acts as a schema for interpreting ambiguous visual information (Boltz, 2001; Cohen, 2005).

Within the present context, visual accompaniment can add schematic value by eliciting emotion, providing thematic information, or highlighting common associations that provide an interpretative framework for the musical stimulus. Present and past research (e.g., Yarbough & Hendel, 1991; Gillespie 1997) suggests that audiovisual presentations like music videos could potentially utilize visual information to elicit stronger and more lasting responses to the music independent of the visuals. The implications of an enhancement effect of visual accompaniment are pertinent to a range of industries, from educational to marketing. In the case of music videos,

the perceptions, interpretations and personal connections generated by the interpretative framework of the carefully constructed visuals could boost music popularity and sales.

To better understand both the theoretical and practical applications of the effects of visual information on music, the style of cognitive processing must also be considered. The widely used Elaboration Likelihood Model (Petty & Wegener, 1999) makes the distinction between central and peripheral route processing. The former is characterized by higher cognitive elaboration resulting in stronger and more lasting attitudinal judgments, while the latter requires lower cognitive involvement and leads to more heuristically-based evaluations. The characteristics that encourage specific placement along the elaboration continuum are frequently used as marketing tactics to achieve the desired level of cognitive processing and persuasion. A visual stimulus can be designed to act as either a musical distracter or enhancer, or to increase the likelihood of either processing route being utilized. Goldberg and colleagues (1993) found that music videos lacking closure required more interpretive effort, encouraged the central processing route and reduced the chance of wear-out. The present research, in addition to past studies on music videos (e.g. Hall et al., 1986; Blanchard-Field et al., 1986), suggests that somewhat ambiguous visuals encourage central processing. Simplistic and narrowly defined visual material, by contrast, is more conducive to the peripheral route. The format effects found in the current study – montages providing more meaning and receiving more extreme evaluative ratings – suggest that the experimental montages would more readily engage central processing, while the videos encourage peripheral processing. It must be noted that the present findings posit format trends associated with structural elements that characterized the experimental montages and videos, which is not to say that any montage or video would affect perceptual processing in exactly the same way. Nevertheless, these findings on format manipulation are of great practical value for

music advertising or promotional efforts seeking to influence which type of evaluative processing route is used.

The present study not only has implications for the ELM, but also for the Congruence-Associationist Model (e.g., Cohen 2000, 2001, 2005) – the only model that directly addresses the influence of music on the perception of visual information. Cohen’s model proposes that the processing of auditory and visual stimuli across all film domains is integrated using both short-term and long-term memory to construct an appropriate visual narrative. While the present study examines the reverse relationship, it still lends support to Cohen’s model. The fact that musical and visual stimuli were confirmed as affectively ambiguous and charged, respectively, without any perceived mismatch in the visual and musical fit suggests that the two information sources are jointly perceived and processed. Subjects also aligned their musical perception to the visual mood in the rhythm rating most likely to assist them in making sense of the visual information. Within each format group, the positive audiovisual pairs were perceived as having a more regular rhythm – a positive affective characteristic – whereas the negative conditions were seen as significantly less rhythmic. The Congruence-Associationist Model also raises the question of how much and in what ways musical narratives use visual cues across different dimensions. Future studies could shed light on this by requiring participants to provide a musical narrative after exposure to various audiovisual stimuli. This would help illuminate the relationship between specific visual aspects and subsequent musical understanding.

Another area that deserves attention is the effect of visual information on musical memory. Research on the remembrance of visual information supports and refines the Congruence-Associationist Model to incorporate auditory and visual mood-congruence as the determining factor for the encoding and memory process, particularly at the first level of cross-

modal analysis (Boltz, Schulkind, and Kantra, 1991; Boltz, 2004). Boltz (2004) explains that when mood congruence is high, joint encoding prompts a high level of integration between audio and visual information, whereas incongruent stimuli result in less integration. Additionally, when the narrative construction takes place at the next level of short-term memory, mood congruence further validates the affective properties of the character and events, enhancing the joint encoding of the audio and visual information in long-term memory. This leads to the reverse question: Is music encoded and remembered differently as a result of visual pairing? If the ambiguous tunes were in fact jointly encoded in the subject's memory with the affectively charged visuals, the music would likely be systematically misremembered as having acoustical qualities associated with the visual mood, suggesting enhanced effects of visual information on musical memory. For example, songs paired with positive visual stimuli would be misremembered as faster, more rhythmic, louder, and higher pitched, whereas negative audiovisual conditions would be misremembered as having more negatively associated acoustical characteristics. If visual accompaniment resulted in decreased performance on memory tasks without any effect of mood, it would be suggestive of an independent encoding process and confirm that visual material acts as a distracter for musical stimuli. Future research could test this by using the same stimulus material and sets as the present experiment, but instead of completing perceptual ratings after each condition, following a distracter task of meaningful time delay, participants would be given a surprise recognition task of the tunes independent of any visual information.

The present research shows that the type of visual accompaniment in terms of affective and structural qualities affects the perception of music. It lays the groundwork for understanding the effects of visual stimuli on music and suggests how future research can contribute to this

important area of study. The theoretical and practical applications of this work make it pertinent to researchers and non-researchers alike. The present study focused on psychological principles that would support forthcoming explorations in additional fields. Specifically, the music education field could better promote musical understanding by implementing teaching tools or styles that incorporate visual information. Furthermore, marketers of almost any product – including music – could work on increasing the impact of their promotional material by influencing the potential consumer’s perceptual processing. The hope is that the current research will motivate continued explorations into the dynamic relationship between the visual and auditory domains, and thus enhance the numerous areas that stand to benefit from further research.

REFERENCES

- Adams, B. L. (1994). *The effects of visual/aural conditions on the emotional response to music*. Unpublished doctoral dissertation, Florida State University, Tallahassee.
- Alpert, J. I., & Alpert, M. I. (1990). Music influences on mood and purchase intentions. *Psychology & Marketing*, 7 (2) 109-133.
- Alpert M. I., & Alpert, J. I. (2005). Purchase occasion influence on the role of music in advertising. *Journal of Business Research*, 58, 369-376.
- Blanchard-Fields, F., Coon, R. C., & Matthews, R. C. (1986). Inferencing and television: developmental study. *Journal of Youth and Adolescence*, 15, 453, 459.
- Bolivar, V. J., Cohen, A. J., & Fentress, J. C. (1994). Semantics and formal congruency in music and motion pictures: Effects on the interpretation of visual action. *Psychomusicology*, 13, 28-59.
- Boltz, M. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception*, 4, 427, 454.
- Boltz, M. (2004). The cognitive processing of film and musical soundtracks. *Memory and Cognition*, 32, 1194-1205.
- Boltz, M. (2004) Film and musical soundtracks. Unpublished.
- Boltz, M. G., Schulkind, M., & Kantra, S. (1991) Effects of background music on the remembering of filmed events. *Memory & Cognition*, 19, 593-606.
- Bharucha, J. J. (1994). Tonality and expectation. In R. Aiello & J. A. Slobada (Eds.), *Musical perceptions* (pp. 213-239). New York: Oxford University Press.
- Bruner, G. (1990). Music, Mood, and Marketing. *Journal of Marketing*, 54, 94-104.
- Cassidy, J., & Sims, W. (1991). Effects of special education labels on peers' and adults' evaluations of a handicapped youth choir, *Journal of Research in Music Education*, 39, 23-34.
- Center for the Study of Emotion and Attention [CSEA-NIMH] (2001). *The international affective picture system: Digitized photographs*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
- Chopra, Y. (producer and director). (1998). *Darr: a violent love story* [film]. Available from Eros International.

- Clynes, M. (1975). Communication and generation of emotion through essentic form. In L. Lennart (Ed.) *Emotions-Their parameters and measurement* (pp. 561-601). New York, NY: Raven Press.
- Clynes, M. (1977). *Sentics: The touch of emotions*. Garden City, NY: Anchor Press/Doubleday.
- Clynes, M. (1980). The communication of emotions theory of sentics. In R. Poutchik & H. Kellerman (Eds.) *Emotion, theory, research, and experience* (pp. 271-300). New York, NY: Academic Press, Inc.
- Clynes, M., & Nettheim, N. (1982). The living quality of music. In M. Clynes (Ed.) *Music, mind, and brain* (pp. 47-82). New York, NY: Plenum Press.
- Cohen, A. J. (1993). Associationism and musical soundtrack phenomena. *Contemporary Music Review*, 9, 163-178.
- Cohen, A. J. (1995). One-trial memory integration of music and film: A direct test. Paper presented at the annual Meeting of the Canadian Acoustical Association, Quebec.
- Cohen, A. J. (1998). The functions of music in multimedia: A cognitive approach. Proc. 5th International conference on music perception and cognition. *ICMPC Seoul*, 5, 13-20.
- Cohen, A. J. (2000). Film music: Perspectives from cognitive psychology. In J. Buhler, C. Flinn & D. Neumeyer (Eds.). *Music and Cinema* (pp. 360-377). Hanover: Wesleyan University.
- Cohen, A. J. (2001). Music as a source of emotion in film. In Juslin P. & Sloboda, J. (Eds.). *Music and emotion* (pp.249-272). Oxford: Oxford University Press.
- Cohen, A. J. (2002). Music cognition and the cognitive psychology of film structure. *Canadian Psychology*, 43, 215-232.
- Cohen, A. J. (2005). How music influences the interpretation of film and video: Approaches from experimental psychology. In R.A. Kendall & R. W. Savage (Eds.), *Selected Reports in Ethnomusicology: Perspectives in Systematic Musicology*, 12, 15-36.
- Cupchik, G., & Saltzman, M. (1999). Reception and memory for serious and light rock music videos. *Empirical Studies Of The Arts*, 17, 59-72.
- Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. New York, NY: Academic Press.
- Davidson, J. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, 21, 103-113.
- DeNora, T. (1999). Music as a technology of the self. *Poetics*, 27, 31-56.

- Gabrielsson, A., & Lindstrom, E. (2001). The influence of musical structure on emotional expression. In Juslin P. & Sloboda, J. (Eds.), *Music and emotion* (pp.223-248). Oxford: Oxford University Press.
- Geringer, J., Cassidy, J., & Byo, J. (1996). Effects of music with video on responses of nonmusic majors: an exploratory study. *Journal of Research in Music Education*, 44, 240-251.
- Geringer, J., Cassidy, J., & Byo, J. (1997). Nonmusic majors' cognitive and affective responses to performance and programmatic music videos. *Journal of Research in Music Education*, 45, 221-233.
- Gillespie, R. (1997). Ratings of violin and viola vibrato performance in audio-only and audiovisual presentations. *Journal of Research in Music Education*, 45, 212-220.
- Goldberg, M. E., Chattopadhyay, A., Gorn, G. J., & Rosenblatt, J. (1993). Music, music videos, and wear out. *Psychology & Marketing*, 10 (1), 1-13.
- Gorn, G. J. (1982). The effects of music in advertising on choice behavior: A classical conditioning approach. *Journal of Marketing*, 46, 94-101.
- Grossberg, S. (1995). The attentive brain. *American Scientist*, 83, 438-449.
- Gundlach, R. H. (1935). Factors determining the characterization of musical phrases. *American Journal of Psychology*, 47, 624-644.
- Hall, J., Miller, C., & Hanson, J. (1986). Music Television: A perceptions study of two age groups. *Popular Music and Society*, 4, 17-28.
- Henkin, R. I. (1955). A factorial study of the components of music. *Journal of Psychology*, 39, 161-181.
- Henkin, R. I. (1957). A reevaluation of a factorial study of the components of music. *Journal of Psychology*, 43, 301-306.
- Hevner, K. (1935a). The affective character of the major and minor modes in music. *American Journal of Psychology*, 47, 103-118.
- Hevner, K. (1935b). Expression in music: A discussion of experimental studies and theories. *Psychological Review*, 47, 186-204.
- Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, 48, 246-268.
- Hevner, K. (1937). The affective value of pitch and tempo in music. *American Journal of Psychology*, 49, 621-630.

- Iwamiya, S. (1994). Interaction between auditory and visual processing when listening to music in an audiovisual context: 1. Matching 2. Audio quality. *Psychomusicology*, 13, 133-154.
- Jarre, J.M. (2006). Industrial Revolution Overture. On *The symphonic Jean-Michel Jarre*. [CD]. Silva America records.
- Juslin, P. N. (1997). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgment policy. *Musicae Scientiae*, 1, 225-256.
- Kastner, M. P., & Crowder, R. G. (1990). Perceptions of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception*, 8, 189-202.
- Kellaris, J. J., & Cox, A. N. (1989). The effects of background music in advertising: A reassessment. *Journal of Consumer Research*, 16 (1), 113-118.
- Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Oxford, UK: Oxford University Press.
- Lang, P.J., Bradley, M. M., & Cuthbert, B. N. (2001). *International affective picture system (IAPS): Instruction Manual and Affective Ratings. Technical Report A-5*. Gainesville, FL. The Center for Research in Psychophysiology, University of Florida.
- Leigh, J. H. (1991). Information processing differences among broadcast media: Review and suggestions for research. *Journal of Advertising*, 20 (2), 71-75.
- Lindstrom, E. (1997). Impact of melodic structure in emotional expression. In A. Gabrielsson (Ed.) *Proceedings of the third triennial ESCOM conference, Uppsala, June 1997 (pp.292-297)*. Uppsala, Sweden: Uppsala University.
- Lipscomb, S. D., & Kendall, R. (1994). Perceptual judgment of the relationship between musical and visual components in film. *Psychomusicology*, 13, 60-98.
- Mack, A., & Rock, I. (1998). *Inattention blindness*. Cambridge, MA: MIT Press.
- Magidson, M. (producer) & Fricke, R. (director). (1992). *Baraka* [film]. Available from Mpi Home Video.
- Maman, B. (2006). Traverser le Temps. On *Benza!* [CD]. Tam-A-Do Academy.
- Marshall, S., & Cohen, A. (1988). Effects of musical soundtracks on attitudes toward animated geometric figures. *Music Perception*, 6, 95-112.
- Merriam, A. (1964). *The anthropology of music*. Chicago, IL: Northwestern University Press.
- Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: Chicago University Press.

- McKennis, L. (1997). The Mummer's Dance. On *The Book of Secrets*. [CD]. Quinlan Road Records
- Nielzen, S., & Cesarec, Z. (1982). Emotional expressions of music as a function of musical structure. *Psychology of Music*, 10 (2), 7-17.
- Oldfield, M. (1996). The Chamber. On *Songs of Distant Earth*. [CD]. Reprise Records.
- Öhman, A. (2002). Automaticity and the amygdala: nonconscious Responses to emotional faces. *Current Directions in Psychological Science* 11 (2), 62–66.
- Park, C. W. & Young, S. M. (1986). Consumer response to television commercials: The impact of involvement and background music on brand attitude formation. *Journal of Marketing Research*, 23, 11-24.
- Petty, R., & Wegener, D. The elaboration likelihood model: current status and controversies. In draft of chapter to appear in: Chaiken, S & Trope, Y. (Eds.), *Dual process theories in social psychology* (pp. 1-34). New York: Guilford Press.
- Rigg, M. G. (1940). Effect of register and tonality upon musical mood. *Journal of Musicology*, 2, 49-61.
- Rigg, M. G. (1964). The mood effects of music: a comparison of data from four investigators. *Journal of Psychology*, 58, 427-438.
- Roman, K., & Mass, J. (1976). *How to Advertise* (pp. 50). New York, NY: St. Martin's Press.
- Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 6, 1161-1178.
- Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, 1, 331-346.
- Scott, L. M. (1990). Understanding jingles and needledrop: A rhetorical approach to music in advertising. *Journal of Consumer Research*, 17, 223-236.
- Sekular, R. & Blake, R. (1994). Perception. New York: McGraw-Hill, Inc.
- Sloboda, J., & Juslin, P. (2001). Psychological perspectives on music and emotion. In Juslin P. & Sloboda, J. (Eds.), *Music and emotion* (pp. 71-104). Oxford: Oxford University Press.
- Sloboda, J., & O'Neill, S. (2001). Emotion in everyday listening to music. In Juslin P. & Sloboda, J. (Eds.), *Music and emotion* (pp.415-429). Oxford: Oxford University Press.
- Swanwick, K. (1973). Musical cognition and aesthetic response. *Psychology of Music*, 1, 7-13.

- Thayer, J., & Russell, M. L. (1983). Differential effects of induced mood on recall of positive, negative, and neutral words. *British Journal of Clinical Psychology*, 22, 163-172.
- Thompson, W. F., & Robitaille, B. (1992). Can composers express emotions through music? *Empirical Studies of the Arts*, 10, 79-89.
- Tiergarten (1985). [Recorded by Tangerine Dream]. On *Le Parc* [LP]. Relativity Records.
- Vinovich, G. S. (1975). *The communicative significance of musical affect in eliciting differential perception, cognition and emotion in sound-motion media messages*. Unpublished doctoral dissertation, University of Southern California
- Watson, K. B. (1942). The nature and measurement of musical meanings. *Psychology Monographs*, 54 (2), 1-43.
- Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology*, 13 (4), 241-257.
- Yarbrough, C., & Hendel, C. (1993). The effect of sequential patterns on rehearsal evaluations of high school and elementary students, *Journal of Research in Music Education*, 41, 246-257.

APPENDIX A
Final Tune Stimuli

The Chamber by Mike Oldfield
Mummer's Dance by Loreena Mckennit
Industrial Revolution Overture by Jean Michel Jarre
Tiergarten by Tangerine Dream
Traverser le Temps by Benza Maman

APPENDIX B

Summary of the Film Clips Used as Experimental Stimuli

Positive Video A: Nature Scene

Water from a smooth waterfall tumbles down a rock face with mist rising to the sky. A rainbow appears amongst the mist. The film switches to a different, more intense waterfall with panoramic and close-up shots. The sequence ends with an aerial shot above a peaceful and large body of water with clouds reflecting from its surface. (From *Baraka*, 1992, produced by Mpi Home Video, 16101 South 108th Avenue, Orland Park, IL 60467).

Positive Video B: Romantic Bollywood Clip

The scene opens with a woman playfully dancing in front of a country campfire in the mountains when her lover comes and steals her away. The action jumps to the woman standing on a later repairing their house. She suddenly falls, but her lover is right below to catch her. They share a romantic embrace on the pier beside their house, which overlooks a beautiful pond in the mountains. The action jumps to a flirtatious moment in a bathtub before the couple embrace once more on a bed. (From *Darr: A Violent Love Story*, 1998, produced by Eros International Inc., 550 County Avenue, Secaucus, NY 07094).

Negative Video A: Chicken Factory

Baby chickens in a processing plant are squeezed together and fall helplessly from one conveyer belt to another. Workers inspect their wings and fling them into a Funnel where different workers brand their beaks using a heated metal instrument. The suffering of the chicks is noted. (From *Baraka*, 1992, produced by Mpi Home Video, 16101 South 108th Avenue, Orland Park, IL 60467).

Negative Video B: Homeless Conditions

A homeless couple sleep under a bridge as cars pass by. The scene turns to another homeless person is a lean-to and then to another homeless person sleeping while breathing heavily. Next, a homeless mother is seen covering her children with rags as she mumbles to herself. A homeless child is then shown unsuccessfully begging passers for money. The scene ends with a group of homeless people crouched in a public space, beneath an overhang, seeking shelter from the cold and wind. (From *Baraka*, 1992, produced by Mpi Home Video, 16101 South 108th Avenue, Orland Park, IL 60467).

APPENDIX C

Experiment Musical and Visual Pairings Sets

Tune Stimuli Key:

1=Chamber	4=Traverser le Temps
2=Tiergarten	5=The Mummer's Dance
3=Industrial Revolutions Overture	

Visual Stimuli-Video Key:

Positive video A=nature scene	Positive video B=romantic Bollywood clip
Negative video A=chicken factory	Negative video B=Homeless conditions

Visual Stimuli-Montage Key:

Positive montage A=version 1	Positive montage B=version 2
Negative montage A=version 2	Negative montage B=version 2

Set A:

Tune 1 – no visual
Tune 2 – negative montage A
Tune 3 – negative video A
Tune 4 – positive montage B
Tune 5 – positive video A

Set B:

Tune 5 – negative montage B
Tune 4 – negative video A
Tune 3 – no visual
Tune 2 – positive video A
Tune 1 – positive montage B

Set C:

Tune 3 – positive montage B
Tune 4 – positive video B
Tune 1 – negative montage B
Tune 2 – negative video A
Tune 5 – no visual

Set D:

Tune 3 – negative montage B
Tune 4 – negative video B
Tune 5 – positive montage B
Tune 1 – positive video B
Tune 2 – no video

Set E:

Tune 1 – no visual
Tune 5 – positive video B
Tune 4 – positive montage A
Tune 3 – negative video B
Tune 2 – negative montage B

Set F:

Tune 3 – positive montage A
Tune 2 – negative video B
Tune 5 – no visual
Tune 4 – positive video A
Tune 1 – negative montage A

Set G:

Tune 2 – positive montage B
Tune 5 – negative montage B
Tune 3 – positive video B
Tune 1 – negative video B
Tune 4 – no visual

Set H:

Tune 5 – negative video B
Tune 1 – positive montage A
Tune 4 – negative montage A
Tune 2 – positive video B
Tune 3 – no visual

Set I:

Tune 2 – no visual

Tune 1 – positive video A

Tune 3 – negative montage A

Tune 5 – positive montage A

Tune 4 – negative video A

Set J:

Tune 4 – no visual

Tune 5 – negative montage A

Tune 2 – positive montage A

Tune 1 – negative video A

Tune 3 – positive video A

APPENDIX D
Affective Response to Stimuli Used in Pretests and Experiment Rating Set A

Pretest

1	2	3	4	5	6	7	8	9
strongly like				neutral				strongly dislike

1	2	3	4	5	6	7	8	9
very positive				neutral				very negative

1	2	3	4	5	6	7	8	9
very active				neutral				very passive

Rating Set A

1	2	3	4	5	6	7
strongly like			neutral			strongly dislike

1	2	3	4	5	6	7
very positive			neutral			very negative

1	2	3	4	5	6	7
very active			neutral			very passive

APPENDIX E
Adjective Choice Task for Experiment Rating Set B

Positive Affect/High Arousal

Exciting
Delightful
Happy
Pleasing
Passionate
Triumphant
Cheerful

Positive Affect/Low Arousal

Satisfying
Serene
Calm
Soothing
Relaxing
Tranquil
Sleepy

Negative Affect/High Arousal

Frustrating
Annoying
Distressing
Tense
Tragic
Angry
Frightening

Negative Affect/Low Arousal

Boring
Gloomy
Mournful
Melancholy
Sad
Depressing
Miserable

*Words presented in randomized order and without category labels

APPENDIX F
Musical Stimuli Ratings for Experiment Rating Set C

1 very fast	2	3	4 neutral	5	6	7 very slow
1 very tonal	2	3	4 neutral	5	6	7 very atonal
1 very regular rhythm	2	3	4 neutral	5	6	7 very irregular rhythm
1 very loud	2	3	4 neutral	5	6	7 very soft
1 very harmonious	2	3	4 neutral	5	6	7 very discordant
1 very flowing	2	3	4 neutral	5	6	7 very choppy

APPENDIX G
Visual Stimuli Ratings for Experiment Rating Set D

Visual and Tune Fit:

1	2	3	4	5	6	7
very well						very poorly

Visual Theme:

1	2	3	4	5	6	7
very meaningful						very unmeaningful

Visual likeability:

1	2	3	4	5	6	7
strongly like						strongly dislike

Visual affect:

1	2	3	4	5	6	7
strongly positive						strongly negative