Emotion in Motion: Investigating the Time-Course of Emotional Judgments of Musical Stimuli

Justin Pierre Bachorik, Marc Bangert, and Psyche Loui
Beth Israel Deaconess Medical Center and Harvard Medical School

Kevin Larke
University of California, San Diego

Jeff Berger
Sourcetone LLC, Brooklyn, NY

Robert Rowe
Steinhardt School, New York, NY

Gottfried Schlaug
Beth Israel Deaconess Medical Center and Harvard Medical School

MUSIC ELICITS PROFOUND EMOTIONS; HOWEVER, THE time-course of these emotional responses during listening sessions is unclear. We investigated the length of time required for participants to initiate emotional responses ("integration time") to 138 musical samples from a variety of genres by monitoring their real-time continuous ratings of emotional content and arousal level of the musical excerpts (made using a joystick). On average, participants required 8.31 s (*SEM* = 0.10) of music before initiating emotional judgments. Additionally, we found that: 1) integration time depended on familiarity of songs; 2) soul/funk, jazz, and classical genres were more quickly assessed than other genres; and 3) musicians did not differ significantly in their responses from those with minimal instrumental musical experience. Results were partially explained by the tempo of musical stimuli and suggest that decisions regarding musical structure, as well as prior knowledge and musical preference, are involved in the emotional response to music.

Received April 9, 2008, accepted December 10, 2008.

Key words: music, emotions, arousal, valence, time

MUSIC, BY ITS VERY NATURE, OCCURS OVER time. As such, the emotional response it elicits from listeners is inherently dynamic, and a clear understanding of exactly how and when this response unfolds is fundamental to our greater understanding of musical emotion. Though the time-course of this response (and how it is affected by musical features as they develop) has been investigated previously, it remains little understood. The present study is primarily concerned with clarifying the temporal nature of our emotional responses to musical stimuli by applying a temporally sensitive approach to measuring musical emotions.

The subject of emotion perception in music has received much interest in both theoretical and empirical studies (Juslin & Sloboda, 2001). More specifically, studies have attempted to identify the emotions elicited by music, as well as the axes along which music-induced emotions can be defined (Hevner, 1936; Sloboda, 1991; Terwogt & van Grinsven, 1991). Additional studies have investigated the emotional perception of music from a wide variety of perspectives, including the role of cultural cues in emotional perception (Balkwill & Thompson, 1999), the influence of age on emotional understanding (Cunningham & Sterling, 1988), and the physiological reactions (e.g., chills and shivers; Panksepp, 1995) that arise as a result of the emotional response to music. However, the operational definitions of "emotion," "feeling," and "affect," as well as the differences between perceived and felt emotions (Gabrielsson, 2002; Scherer, 2004), are unclear in these studies.

While autonomic and electrophysiological recordings have provided more time-sensitive biological markers for emotion perception in music (Bernardi, Porta, & Slight, 2006; Blood & Zatorre, 2001; Krumhansl, 1997; Sammler, Grigutsch, Fritz, & Koelsch, 2007; Steinbeis, Koelsch, & Sloboda, 2006), the degree to which biological markers predict the multidimensional psychological experience of musical emotion is unclear. Similarly, the growing body of neuroimaging work investigating the neural basis of music-induced emotions has implicated brain regions involved in syntax processing, sequencing, and reward and motivation

Music Perception VOLUME 26, ISSUE 4, PP. 355–364, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2009.26.4.355

perception, as well as primary sensory areas (Altenmüller, Schurmann, Lim, & Parlitz, 2002; Blood & Zatorre, 2001; Brown, Martinez, & Parsons, 2004; Koelsch, Fritz, von Cramon, Muller, & Friederici, 2006; Menon & Levitin, 2005; Mizuno & Sugishita, 2007; Pallesen et al., 2005). However, the extent to which these brain differences reflect the different axes of emotion is yet to be determined.

One question raised by both the cognitive and neuroscience studies concerns the temporal resolution of musical emotions. Several studies have examined responses to music over the course of a given musical stimulus. In an investigation using categorization and grouping techniques, Perrot and Gjerdingen (1999) reported that individuals are capable of classifying the genre to which a given piece of music belongs when exposed to as little as 250 ms of a musical excerpt. Results from this categorization study suggest that features in music may be available for recognition by the cognitive system in as little as 250 ms. In a similar study by Schellenberg, Iverson, and McKinnon (1999), the experimenters found that participants performed well above chance at matching song titles and artists to stimuli, even when the musical excerpts were only 100 ms in length. Though these studies do not directly address the time-course of the emotional response to music, they demonstrate the (relatively small) amount of stimulus exposure that individuals require in order to make cognitive decisions regarding musical content.

In a study directly investigating the effects of duration of musical stimuli on their emotional classification, Bigand, Filipic, and Lalitte (2005) employed a novel experimental paradigm in which participants were given a number of musical excerpts (represented on a computer screen as icons) and were asked to divide them into groups based upon the similarity of the emotional responses they provoked. Multidimensional scaling methods revealed that participants, when given musical excerpts only 1 s in length, grouped songs similarly to those who were given much longer excerpts (up to 250 s). However, the rapid classification of emotions (as in the above study) may not be fully representative of the musical experience, since music is of a dynamic nature and requires the integration and evaluation of continuous information over time. As many have pointed out, music perception involves a time-course that begins with an initial emotional percept and continues with subsequent appraisal, regulation, and evolution of this emotion over time (Bigand, Filipic, & Lalitte, 2005; Krumhansl, 1996; Krumhansl, 1997; Krumhansl, 2002; Vines, Krumhansl, Wanderley, Dalca, & Levitin, 2005; Vines, Krumhansl, Wanderley, & Levitin, 2006).

To that end, Schubert (2004) conducted a study in which participants continuously responded in two-dimensional emotion space to four different musical stimuli whose lower-level acoustic characteristics and higher-level musical features were modeled against participants' responses (see Schubert, 1999, for an extensive examination of this approach). Each stimulus was thus used to construct a separate model of emotional response to acoustic and musical cues, and the models' approximation of "lag-time" between causal musical/acoustic cue and emotional event ranged from zero to three seconds. The stimulus set used in this experiment was somewhat limited, however, as the four songs were all classical pieces drawn from the romantic repertoire.

More recently, Grewe, Nagel, Kopiez, and Altenmüller (2007) investigated the effects of six full-length pieces of music on participants' continuous two-dimensional ratings (as in Schubert, 2004, and Nagel, Kopiez, Grewe, & Altenmüller, 2007) as well as their subjective reports of feelings and biological signals (skin conductance and facial electromyography). This triple-measurement approach was chosen in response to some of the criticism of previous studies by Scherer (2004), who put forth the component process model of emotion. This model consists of the "emotion response triad" of physiological arousal, motor expression, and subjective feelings. Though Grewe et al. (2007) were able to find significant correlations between the stimuli's musical features and each of the three outcome measures, there were no obvious "affective events that regularly occurred across all three components in response to any of [the] stimuli" (p. 786). They also analyzed the time-series of the continuous two-dimensional emotion space responses and were able to extract some causal relationships between musical features and emotional response (e.g., beginning of a new musical section or entrance of a new leading voice).

Despite this extensive research, however, neither of the above studies reported the amount of time participants required to integrate information from musical excerpts before initiating a physical movement to indicate an emotional judgment. It is this question that motivated the present study.

In the present study we used musical stimuli of various genres and levels of familiarity and instructed participants to respond to musical excerpts by making joystick movements in an onscreen two-dimensional grid whose x- and y-axes represented emotional valence (positive/negative) and arousal (calming/stimulating) respectively (see Russell, 1980, for a discussion of "emotion space" and these axes). This approach to gathering a continuous participant response was first used by

Schubert (1996) but to our knowledge has never been used to gauge the length of time participants require to make an initial judgment about the emotional content of a piece of music. We referred to this time as the "integration time." In addition to this real-time continuous feedback, we gathered data from participants regarding their musical background and familiarity with each stimulus song for additional analysis.

We posit that affective perception of music might work as an incremental-information-integration apparatus: as soon as initial sensory information is perceived, emotional processing begins, producing hypotheses about the genre and emotional content. As processing continues, it constantly incorporates additional musical information from the sensory stream, which in turn serves to test, refine, and retest the emotional hypothesis. As perceived musical information accumulates, more information is gathered about the target emotion of the song; thus, we expect to observe a significant correlation between the "integration time" (the amount of time from the beginning of playback till the first movement) and the agreement of its direction with that of the overall judgment that participants make at the end of each musical excerpt. The earlier the movement, the more uncertainty (less correlation between the angle of the first movement and the overall emotional judgment) such a movement might entail; conversely, the later the movement (i.e., after several incremental emotional hypothesis-iterations), the higher its correlation to the overall emotional judgment (assuming that each song maintains a relatively constant emotional ambience). Furthermore, we would predict that participants possessing a stronger musical background or greater familiarity with the stimulus genre or stimulus excerpt would need fewer iterations of comparing and testing the continuous auditory stream with their emotional hypothesis; therefore their integration times would be shorter. In the present study, we test the incremental-information-integration hypothesis by obtaining continuous emotion ratings for musical stimuli. Results validate the use of two-dimensional emotion ratings and provide support for musical emotion perception as an information-dependent, incremental process.

Method

Participants

Eighty-one participants (49 females and 32 males) were recruited from the greater Boston metropolitan area via advertisements in daily newspapers. Participants ranged from 19 to 82 years of age (median age = 29). All participants completed a questionnaire regarding their musical background (including whether or not they had previously studied a musical instrument, for how long, how intensely they practiced, and whether they had taken music theory courses) and musical preferences (what music they liked to listen to now and what music they grew up with, etc.). In order to differentiate people based upon their level of musicianship, we separated the participants into two groups—musicians with substantial instrumental musical experience and individuals with minimal instrumental musical experience (MIMEs, which may be referred to as nonmusicians in other publications)—based upon their pre-trial-block questionnaire responses. Of our participants, 35 (14 males and 21 females) were musicians, and 46 (18 males and 28 females) were MIMEs. The musician group consisted both of individuals that had chosen music as their profession and were actively performing musicians and of amateur musicians that had a different profession but had substantial instrumental training and regularly played/practiced a musical instrument. The two groups differed significantly in their years of instrumental musical experience: musicians averaged 19.57 years of instrumental musical experience (SEM = 1.72) and MIMEs averaged 0.83 years (SEM = 0.19).

This study was approved by the Institutional Review Board of Beth Israel Deaconess Medical Center. All participants signed a written informed consent form prior to engaging in the experiment.

Stimuli

The stimuli consisted of 138 musical excerpts 60 s long and were drawn from 11 genres: ambient, classical, electronic, hip-hop, jazz, Latin, movie soundtracks, pop, rock, soul/funk, and world. All excerpts were taken from recordings that are publicly available for purchase. The initial categorization of these stimuli into genres was performed by two independent investigators, and their categorizations were then cross-checked with existing Billboard classifications in order to verify their accuracy. Only songs that were categorized without any uncertainty as to genre were used in this experiment. The 60 s excerpts from each song were chosen by one investigator who had an extensive musical background and was blind with regard to the research questions and hypotheses of this study. The only instructions given to this investigator were to select the 60 s excerpts that best characterized each song.

Approximately 10% of the stimuli contained vocals, and the rest were purely instrumental. Each excerpt was briefly faded in (0.50 s) at the beginning of the stimulus and faded out (0.50 s) at the end. The stimuli were divided into three blocks, ranging in length between 45 and 47 stimuli. Each block preserved the same distribution of genre and vocal/instrumental balance from the original pool of 138. The stimuli ranged in tempi between 40 and 189 beats per minute.

Procedure

Prior to testing, each participant was introduced to the concepts of emotional arousal and valence by completing an emotion-rating task using picture stimuli from the International Affective Picture System (IAPS) on arousal and valence. We selected IAPS pictures that were either extremely emotionally arousing and possessed very positive or negative valence characteristics or were completely neutral in both valence and arousal. Participants made numerical ratings of valence and arousal (each on the scale of 1 to 10) for each picture shown. Before musical testing began, participants' valence and arousal ratings of each picture were reviewed. All of our participants showed responses that were in agreement with the normed classifications of the 15 pictures (3 pictures from each of the 4 quadrants of the arousal and valence domains as well as 3 pictures that were deemed to be neutral). This picture ratings pretest was a reliability measure to ensure that subsequent emotional ratings of musical stimuli were not biased by participants' misunderstanding or unusual perception of valence and/or arousal.

Experiments were conducted using an Apple Powerbook G4 with a 15.4'' LCD screen using custom-made stimulus presentation and response recording software. Audio was presented via Altec Lansing AHP-712 headphones, and participants used a Flightstick Pro branded USB joystick to input their responses to the stimuli.

Stimuli were presented to participants in three blocks containing between 45 and 47 excerpts each. The order of presentation of the blocks and the order of presentation of excerpts within a block were randomized. For each stimulus presentation, the participant's task was the same: to use the joystick to respond, in real time, to the levels of emotional valence (defined as positive or negative emotion induced by the music) and arousal (defined as a stimulating or calming feeling induced by the music) of the music via an onscreen cursor in a two-dimensional grid (see Figure 1). To ensure that accurate measures of initial response time were gathered, participants were instructed to move the joystick as soon as they began feeling an emotional response to the music. The joystick controlled the motion of the cursor in a 640×640 resolution grid, and data about the position of the joystick and the position of the cursor was sampled with a frequency of 10 Hz. Recentering the joystick did not cause the cursor in the onscreen grid to recenter to zero; rather, it caused the cursor to stop moving such that the recorded emotional response remained unchanged until further movement. At the end of each song (after the conclusion of the continuous cursor movements), participants were asked to select one point in the two-dimensional arousal-valence space that best represented the overall arousal and valence levels for that particular song. They also were asked to rate their level of familiarity with the songs from 0-4 according to the following scale: a rating of "0" indicated a song they had never heard before, "1" a song identifiable by artist or genre, "2" a song they had heard before, "3" a song they had heard before many times passively (i.e., on the radio), and "4" a song they had heard many times actively (i.e., because they had chosen to buy the album or owned a recording of that song).

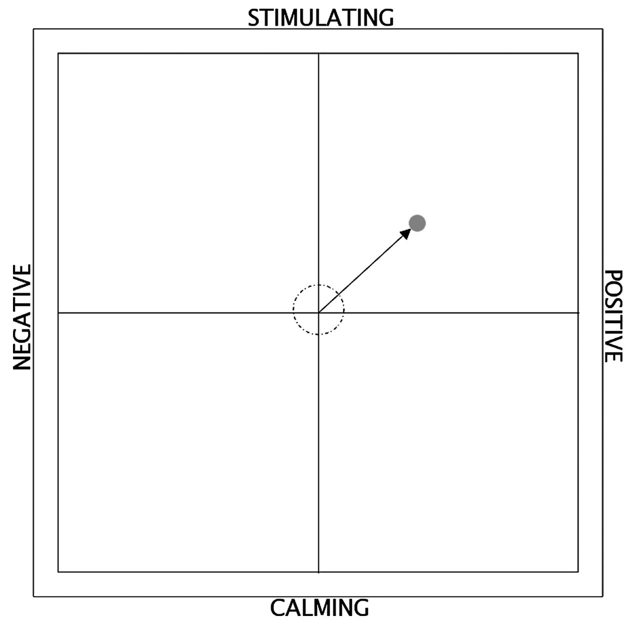

FIGURE 1. Diagram showing a possible vector of cursor motion in the two-dimensional space of valence and arousal. The dotted circle demonstrates our methodology for detecting a given participant's first motion for a given stimulus; the lag-time is the time at which the cursor first leaves the circle. Neither the dotted circle nor the arrow are present onscreen during the actual experiment.

Upon completion of one of the trial blocks (usually the first), participants were asked questions about their musical preferences, including what genres they had grown up listening to, their favorite genres, their favorite features of music, their favorite pieces, and their favorite performers/groups.

Due to technical and scheduling difficulties, some of the participants were unable to complete all three trial blocks. As our analysis showed no effects of trial block or session order, however, all data points were incorporated into our analyses in order to present the most complete data set possible, while accounting statistically for missing data by making the appropriate adjustments in degrees of freedom.

Data Analysis

Integration time (time to initial response) was measured in three different ways. The experimental software automatically recorded a timestamp and the x and y coordinates of both the joystick and the cursor whenever the participant moved the joystick above a certain "jitter" threshold of 15 pixels of joystick motion, which ensured that slight, accidental motions on the part of the participant would not be interpreted as intentional actions. (In other words, movements below the jitter threshold were not recorded by the experimental software or rendered onscreen, and all measurements presented in this paper should be interpreted as being taken above the jitter threshold.) The initial measurement of integration time was simply taken as the timestamp at which the participant moved the joystick at all, while the second and third measurements of integration time were based on the timestamp at which the onscreen cursor moved 10 and 50 pixels, respectively, away from the origin of the valence/arousal grid. These moments were calculated by taking the square root of the squares of x and y cursor position; i.e., finding the length of the vector of motion, and recording the timestamps when the vector magnitudes were greater than 10 and 50. Angles of initial motion and final judgment (see below) were calculated by taking the inverse tangent of the cursor's y position divided by its x position at the timestamps in question and translating the results to degrees.

We used repeated-measures linear regression models (i.e., mixed effects models) to evaluate the relationship between gender, musicianship, stimulus familiarity (5 categories), stimulus genre (11 categories), musical preference (6 categories) and initial integration time while accounting for the correlation among the multiple measurements made within each study participant (Fitzmaurice, Laird, & Ware, 2004). We first considered each variable independently in a series of univariate regression models and subsequently considered these variables simultaneously in a multivariate model. A compound symmetry correlation structure was assumed among multiple measurements within a single participant and empirical standard errors were used because they are known to be less sensitive to the chosen covariance structure of the repeated measures (Venables & Ripley, 1994). Analyses were carried out using the MIXED procedure in SAS v 9.1 (SAS Institute Inc., Cary, NC).

Results

Measurements of integration time were available from 81 participants who completed 1 to 3 trial blocks for a total of 7076 determinations. In a repeated-measures linear regression model accounting for the correlation among multiple measurements in the same participant, there was no evidence that initial integration time varied systematically between trial blocks, $F(2, 72) = 1.10$; $p = .33$, or by order of trial block presentation, $F(2, 72) = 0.24$, $p = .78$.

The mean initial integration time once stimulus playback began, but before participants moved the joystick at all, was 8.31 s (*SEM* = 0.10). However, once participants began to move the joystick, they did so decisively, as the mean initial integration time to move the joystick outside of a 50-pixel-radius circle about the origin was only 10.92 s (*SEM* = 0.12)—a mere 2.61 seconds later.

We used repeated-measures linear regression models to evaluate the association between initial integration time and gender, age, musicianship, stimulus genre, stimulus familiarity, and musical preference. In univariate analyses, initial integration time did not differ significantly according to gender, $F(1, 79) = 0.03$, $p = .86$, age, $F(1, 79) = 2.58$, $p = .11$, or musician status, $F(1, 79) = 0.09$, $p = .76$ (Table 1).

In contrast, increased song familiarity was associated with significantly shorter initial integration times, $F(4, 280) = 19.80$, $p < .01$ (Figure 2). For example, initial integration times for those songs that participants rated as least familiar were 3.46 s (*SEM* = 0.57) longer than songs rated as most familiar.

Similarly, initial integration times differed significantly according to stimulus genre, $F(10, 789) = 6.13$, $p < .01$ (Figure 3). Participants required an average of 7.56 s (*SEM* = 0.12) to respond to classical, electronic, jazz, and soul/funk excerpts, while participants needed on average 10.16 s (*SEM* = 0.24) to judge ambient and movie music. This was not simply due to the relative familiarity or novelty of different genres, as an analysis additionally controlling for stimulus familiarity did not change the pattern of results.

Table 1. Predictors of Initial Integration Time Considered Separately (Univariate Models) or Jointly (Multivariate Models).

		Univariate Models		Multivariate Model	
Variable	Functional Form	*F*	*p*	*F*	*p*
Gender	2 Categories	$F(1, 79) = 0.03$	.86	$F(1, 72) = 1.14$	.29
Age	Continuous, linear	$F(1, 79) = 2.58$	.11	$F(1, 72) = 3.66$	.06
Musicianship	2 Categories	$F(1, 79) = 0.09$	.76	$F(1, 72) = 0.39$	.54
Stimulus Familiarity	5 Categories	$F(4, 280) = 19.80$	< .01	$F(4, 280) = 16.40$	< .01
Stimulus Genre	11 Categories	$F(10, 789) = 6.13$	< .01	$F(10, 789) = 5.38$	< .01
Musical Preference	6 Categories	$F(5, 75) = 1.19$	.32	$F(5, 72) = 0.95$	.45

Of the 81 participants, 25% reported rock as their preferred musical genre, 17% pop, 14% rap/hip-hop, 9% orchestral music, 6% soul, and 30% reported preferring other musical genres. Given this categorization, integration times did not differ significantly based on self-reported preferred musical genre, $F(5, 75) = 1.19$, $p = .32$ (Figure 4). Interestingly, average integration times for participants that reported a preference for soul music were 3.19 s ($SEM = 1.93$) longer than the average integration time for participants that reported a preference for rock music. In a posthoc analysis based on this observed difference, we found that the integration time among those preferring soul was significantly different from those preferring "other" ($p <. 05$) or pop ($p < .05$), but not significantly different from those preferring orchestral ($p = .30$), rap/hip-hop ($p = .08$), or rock music ($p = .09$). Initial integration times did not vary appreciably between participants preferring rock, pop, rap/hip-hop, or orchestral music. Interestingly, initial integration times were significantly shorter when the stimulus genre was the same as a participant's self-reported musical preference (mean for stimulus/preference mismatch: 8.41 s; $SEM = 0.52$; mean for stimulus/preference match: 7.04 s; SEM = 0.53; $F(1, 6994) = 13.90$, $p < .01$) (Figure 5).

The above results were not significantly different when predictors were considered simultaneously in a multivariate model (Table 1).

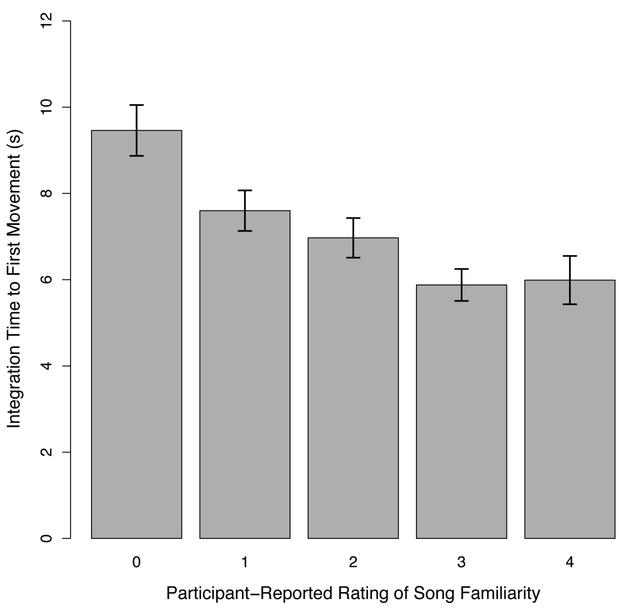

FIGURE 2. Graph of initial integration times (model predicted means +/– SEM) by song familiarity. Level of familiarity was categorized as described in the methods section in detail: a rating of "0" indicated a song they had never heard before, "1" a song identifiable by artist or genre, "2" a song they'd heard before, "3" a song they'd heard before many times passively (i.e., on the radio), and "4" a song they'd heard many times actively.

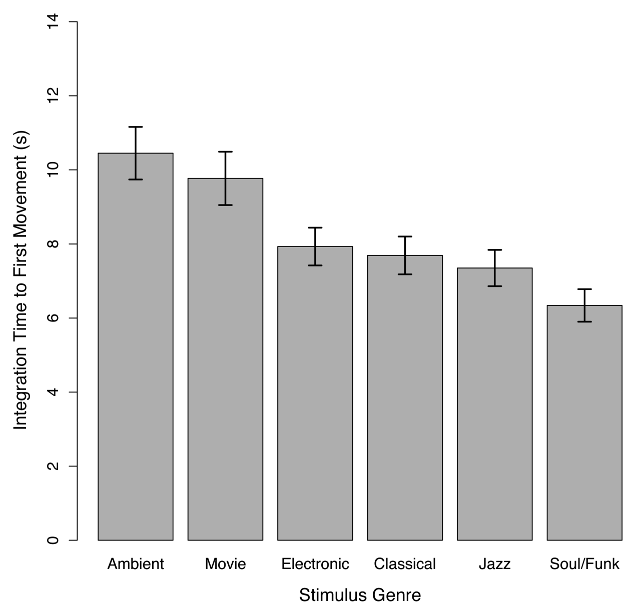

FIGURE 3. Graph of initial integration times (model predicted means +/– SEM) by stimulus genre.

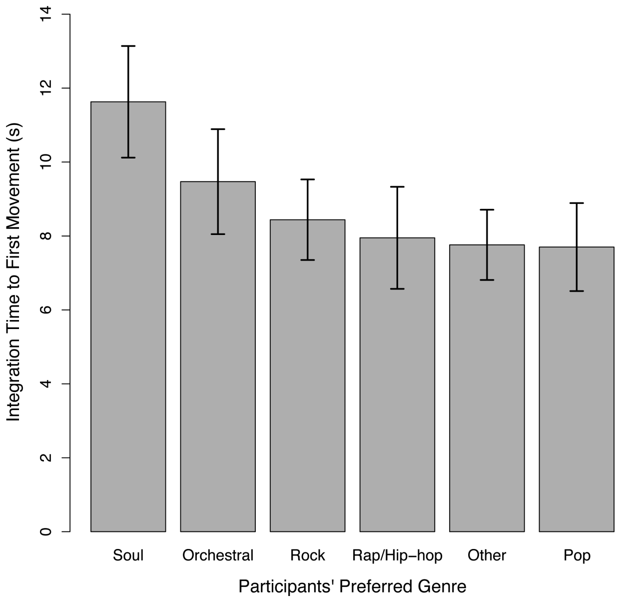

FIGURE 4. Graph of initial integration time (model predicted means +/− SEM) vs. preferred genre.

The majority (130/138) of the stimuli had a constant tempo while the other eight stimuli had a fluctuating tempo. Among the 130 stimuli with a constant tempo, we investigated whether tempo was associated with initial integration time. In a repeated-measures linear regression model, stimulus tempo was a strong and statistically significant predictor of initial integration time.

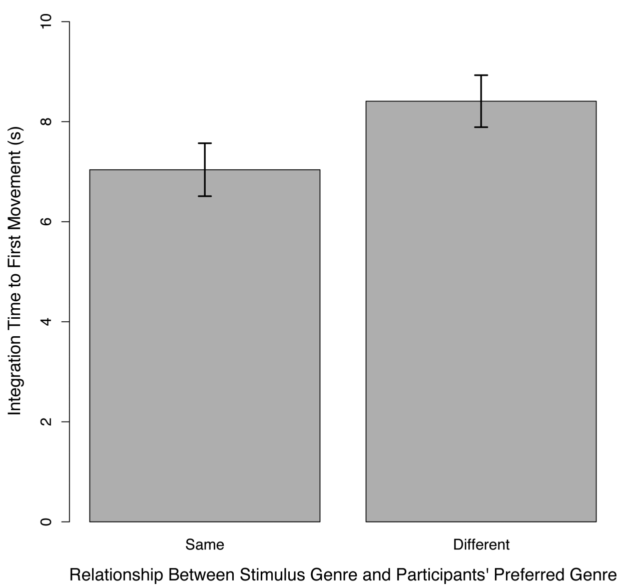

FIGURE 5. Graph of initial integration time (model predicted means +/− SEM) vs. relationship between preferred genre and stimulus genre.

Specifically, we observed a 1.40 s (SEM = 0.20; $F(1, 6581) = 53.20$, $p < .01$) decrease in initial integration time for each 1 beat per second increase in stimulus tempo. This relationship between initial integration time and BPM did not interact with musicianship, $F(1, 6580) = 0.29$, $p = .59$.

In order to investigate the relative consistency of the initial and final judgments of emotional content, we derived an initial/final angular delta measurement by measuring the absolute difference in degrees between the angle of the first movement to reach the 10-pixel threshold and the angle of the dot's final resting point. The mean angular delta between the 10-pixel threshold and the final judgment was 50.17 degrees ($SEM = 0.52$). The mean angular delta was inversely associated with integration time. Specifically, when integration time was modeled as a linear continuous variable in the context of a linear mixed model, a 1 second increase in integration time was associated with a 0.24 degree decrease in the mean angular delta, $F(1, 6547) = 5.98$, $p < .05$), suggesting that the longer participants waited to make judgments regarding the emotional content of a given stimulus, the more accurate (relative to their final judgment) their initial judgments. Results were not different when the mean angular delta between the 50-pixel threshold and the final judgment was considered instead.

Discussion

In the present study, we examined the time-course of the emotional response to music by focusing on the period of time required for initial responses. There has been much work done on understanding emotional responses to music (see Juslin & Sloboda, 2001, and Juslin & Västfjäll, 2008, for an overview), and there also is a reasonable body of literature regarding the time-course of these responses (particularly by Bigand, Filipic, & Lalitte, 2005; Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005; Grewe et al., 2007; Nagel et al., 2007; and Schubert, 1996, 1999, 2004). However, this time to initial response, which we referred to as "integration time", has not been studied in an in-depth manner until now.

The present study, therefore, offers a unique approach to examining integration time, and our findings reveal some interesting differences from those in previous studies. First, participants' response times in our study were of a different order of magnitude than those in studies that required more of a categorical decision (Bigand, Filipic, & Lalitte, 2005; Bigand, Vieillard, et al., 2005; Perrot & Gjerdingen, 1999), as the mean time to initial response in our study was 8.31 s. We believe that this is primarily due to our experimental paradigm, which required participants to make judgments that were not of a strictly categorical nature. Rather than simply classifying each stimulus as belonging to one of several listed musical genres, participants' decisions about arousal and valence may involve more cognitive appraisal processes as to which combination of several dimensions best represented their emotional response, based on a comparison between the presently-perceived stimulus and previously-experienced emotional stimuli (Bigand, Filipic, & Lalitte, 2005; Bigand, Vieillard, et al., 2005). This explanation also highlights another difference between previous investigations and the present study: judgments about emotion and affect might require more introspection on the part of the participant than judgments regarding properties inherent to the stimulus, such as genre. A given song might elicit one emotion from participant A and a completely different sentiment from participant B (see Blood & Zatorre, 2001), despite the fact that both would agree on the song's genre. This dichotomy between classification and emotion perception is corroborated by the result that our participants' integration times seem to be mediated by top-down factors, including song familiarity and participants' favorite musical genres. The latter is particularly interesting in light of the different musical preferences that might be correlated with potential personality differences (Rentfrow & Gosling, 2003), which in turn influences the time taken to make emotional judgments. It also is interesting to note that our participants responded more quickly to stimuli from their favorite musical genre; favorite genre may be a proxy for increased song familiarity, which also shortened integration times. At any rate, if participants using strictly bottom-up strategies in responding to the music, we would not expect to see such variation in response times between different types of participants. The nature of participants' responses indicates that making graded, introspective judgments about complex auditory stimuli requires much more time than assessing stimuli on a purely categorical basis.

We were surprised to find no significant differences in integration time between the musicians and nonmusicians in our study. Although this finding is in keeping with studies suggesting that emotions induced by music are influenced little by background factors, such as level of music education, age or gender (Bigand, Filipic, & Lalitte, 2005; Bigand, Vieillard, et al., 2005; Kratus, 1993; Kreutz, Bongard, & Von Jussis, 2002), there is other evidence to suggest that musicianship can play a role in the time-course of musical perception; for example, a study using the gating paradigm (Dalla Bella, Peretz, & Aronoff, 2003) found that musicians were able to recognize familiar songs faster (i.e., with fewer notes) than nonmusicians. However, the notes used in their study were not all of the same duration, and drawing conclusions about the absolute time required to make such judgments is difficult in the absence of temporal consistency. To the best of our knowledge, previous studies using the continuous-rating two-dimensional emotion space paradigm did not analyze their results on the basis of participants' musicianship, though in some cases the participant pools were balanced across musicians and nonmusicians (Grewe et al., 2007; Schubert, 2004). It is possible that the present study suffered from a lack of power, or too much noise in the dataset; furthermore, actively recruiting musicians (as opposed to identifying them posthoc) might have yielded better balance between the musician and MIME groups both in terms of absolute number of participants (our two groups differed in size by 11) and in terms of music training and experience.

On another note, our study also revealed that the longer participants took to listen to the musical excerpts before moving (i.e., the longer they integrated the continuous auditory information), the higher the correspondence between the angle of the initial movement and the final overall judgment of the musical excerpt. We suggest that affective listening to streaming music might work as an incremental-information-integration apparatus—i.e., as soon as initial sensory information comes in, the processing starts producing hypotheses about the emotional effect. As integration time increases and processing continues, it continually incorporates additional musical information from the sensory stream, which in turn serves to refine the emotional hypothesis.

We also found that the variability in integration time across different songs was linked to the varying tempo of each stimulus; faster songs elicited shorter integration times. This finding is in keeping with the incremental-information-integration apparatus posited above; presumably, faster songs deliver sensory information at a higher rate and therefore require less listening time prior to evaluation. It also is likely that differences in response lags are caused by higher-level structural properties of individual musical excerpts (Grewe et al., 2007; Hevner, 1936; Schubert, 2004; Sloboda, 1991; Steinbeis et al., 2006), though such analyses are beyond the scope of the present study. A theoretical analysis of the first 20 s or so of each song stimulus, including its harmonic structure (Piston & DeVoto, 1987), melodic processes (Narmour, 1990), and rhythmic structure might yield valuable insight as to how participants' response times are affected by higher-level factors. Also, while the present study assumes a two-dimensional interpretation of emotional space, further analyses and experiments could incorporate different perspectives, such as the separation of arousal into the subtypes of tension arousal and excitatory arousal (Ilie & Thompson, 2005).

In this study, we showed that individuals require between 8.31-11.00 s (depending on the measurement being used) to begin formulating emotional judgments regarding musical stimuli, and that this integration time can be modulated by participant-specific variables including familiarity with the musical excerpt and preference for the excerpt's genre. Future studies might consider these results and use stimuli at least 11.00 s in length in measuring the time-course of emotional responses to music. Our results highlight the elements of music that drive emotions and begin to put these emotion determinants in a more temporally sensitive context. Music is a temporal art; by using this method we can begin to understand, at a behavioral level, how and when such temporal features lead to the development of musical emotion.

Author Note

This work was supported by a research grant from Sourcetone, LLC, given to Beth Israel Deaconess Medical Center to support research on music and emotions. The authors have no conflict of interest and no financial interests in Sourcetone, LLC.

We would like to thank L. Cuddy, S. Koelsch and two other reviewers for their helpful comments on this manuscript.

Correspondence concerning this article should be addressed to Gottfried Schlaug, Department of Neurology, Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Avenue, Boston, MA 02215. E-MAIL: gschlaug@bidmc.harvard.edu

References

ALTENMÜLLER, E., SCHÜRMANN, K., LIM, V. K., & PARLITZ, D. (2002). Hits to the left, flops to the right: Different emotions during listening to music are reflected in cortical lateralisation patterns. *Neuropsychologia*, *40*, 2242-2256.

BALKWILL, L. L., & THOMPSON, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17,* 43-64.

BERNARDI, L., PORTA, C., & SLEIGHT, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart, 92*, 445-452.

BIGAND, E., FILIPIC, S., & LALITTE, P. (2005a). The time course of emotional responses to music. *Annals of the New York Academy of Science, 1060*, 429-437.

BIGAND, E., VIEILLARD, S., MADURELL, F., MAROZEAU, J., & DACQUET, A. (2005b). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion, 19*, 1113-1139.

BLOOD, A. J., & ZATORRE, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences, 98*, 11818-11823.

BROWN, S., MARTINEZ, M. J., & PARSONS, L. M. (2004). Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport*, *15,* 2033-2037.

CUNNINGHAM, J. G., & STERLING, R. S. (1988). Developmental change in the understanding of affective meaning in music. *Motivation and Emotion, 12,* 399-413.

DALLA BELLA, S., PERETZ, I., & ARONOFF, N. (2003). Time course of melody recognition: A gating paradigm study. *Perception and Psychophysics*, *65*, 1019-1028.

FITZMAURICE G. M., LAIRD N. M., & WARE J. H. (2004). *Applied longitudinal analysis*. Hoboken, NJ: Wiley-Interscience.

GABRIELSSON, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, Special issue 2001-2002*, 123-147.

GREWE, O., NAGEL, F., KOPIEZ, R., & ALTENMÜLLER, E. (2007). Emotions over time: Synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion, 7,* 774-788.

HEVNER, K. (1936). Experimental studies of the elements of the expression in music. *American Jounal of Psychology, 48*, 246-268.

ILIE, G., & THOMPSON, W. F. (2005). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, *23*, 319-329.

JUSLIN, P. N., & SLOBODA, J. A. (2001). *Music and emotion: Theory and research.* Oxford: Oxford University Press.

JUSLIN P. N., & VÄSTFJÄLL D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31,* 559-575.

Koelsch, S., Fritz, T., von Cramon, D. Y., Müller, K., & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping, 27*, 239-250.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music, 21*, 3-19.

Kreutz, G., Bongard, S., & Von Jussis, J. (2002). Cardiovascular responses to music listening: The effects of musical expertise and emotional expression. *Musicae Scientiae, 6*, 257–278.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K.282: Segmentation, tension, and musical ideas. *Music Perception, 13,* 401-432.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology, 51*, 336-353.

Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science, 11*, 45-50.

Menon, V., & Levitin, D. J. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *Neuroimage, 28*, 175-184.

Mizuno, T., & Sugishita, M. (2007). Neural correlates underlying perception of tonality-related emotional contents. *Neuroreport, 18*, 1651-1655.

Nagel, F., Kopiez, R., Grewe, O., & Altenmüller, E. (2007). EMuJoy: Software for continuous measurement of perceived emotions in music. *Behavioral Research Methods, 39,* 283-290.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.

Pallesen, K. J., Brattico, E., Bailey, C., Korvenoja, A., Koivisto, J., Gjedde, A., et al. (2005). Emotion processing of major, minor, and dissonant chords: A functional magnetic resonance imaging study. *Annals of the New York Academy of Science, 1060*, 450-3.

Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception, 13,* 171-207.

Perrot, D., & Gjerdingen, R. O. (1999). Scanning the dial: An exploration of factors in the identification of musical style [Abstract]. *Proceedings of the 1999 Society for Music Perception and Cognition* (p. 88). Evanston, IL: SMPC.

Piston, W., & DeVoto, M. (1987). *Harmony*. New York, NY: WW Norton.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84*, 1236-1256.

Russell, J. A. (1980). A circumflex model of affect. *Journal of Personality and Social Psychology, 39*, 1161-1178.

Sammler, D., Grigutsch, M., Fritz, T., & Koelsch, S. (2007). Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, *44*, 293-304.

Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin and Review, 6,* 641-646.

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research, 33,* 239-251.

Schubert, E. K. (1996). Continuous response to music using a two dimensional emotion space. *Proceedings of the 4th International Conference of Music Perception and Cognition* (pp. 263-268). Montreal, Canada: ICMPC.

Schubert, E. K. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology, 51,*154-165.

Schubert, E. K. (2004). Modeling perceived emotion with continuous musical features. *Music Perception, 21*, 561-585.

Sloboda, J. A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, *19*, 110-120.

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience, 18*, 1380-1393.

Terwogt, M. M., & van Grinsven, F. (1991). Musical expression of moodstates. *Psychology of Music, 37,* 99-109.

Venables W. N., & Ripley B. D. (1994). *Modern applied statistics with S-Plus*. New York, NY: Springer.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2005). Dimensions of emotion in expressive musical performance. *Annals of the New York Academy of Science, 1060*, 462-466.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition, 101*, 80-113.