Genre identification of very brief musical excerpts

By: Sandra T. Mace, Cynthia L. Wagoner, David J. Teachout, Donald A. Hodges

Abstract:

The purpose of this study was to examine how well individuals were able to identify different music genres from very brief excerpts and whether musical training, gender and preference played a role in genre identification. Listeners were asked to identify genre from classical, jazz, country, metal, and rap/hip hop excerpts that were 125, 250, 500, or 1000 ms in length. Participants ($N = 347$), students recruited from three college campuses in the southeast region of the USA, were found to be quite successful in identifying the genre of brief excerpts, even at 125 ms. Length of excerpt significantly affected participants' ability to identify genre with longer time lengths leading to greater accuracy. Classical, metal, and rap/hip hop excerpts were correctly identified more often than were country or jazz excerpts. Further, there were many distinct interactions across lengths among genres. Musical training did not affect participants' ability to identify excerpts overall or by length, but training was found to affect genre identification: those with training were better able to identify classical and jazz excerpts while those without training were better able to identify rap/hip hop excerpts. Gender did not affect participants' ability to identify excerpts overall or by length, but gender was found to affect genre identification: males were better able to identify metal excerpts. Preference did affect participants' ability to identify excerpts; most favorite genres were identified more accurately than all other genres and least favorite genres were identified less accurately than all other genres. In general, these findings support a primary conclusion that people are adept at identifying particular genres when presented with excerpts that are one second or less.

Keywords: brief excerpt | excerpt length | gender | genre | musical training | preference

Article:

Nearly everyone has had the experience of rapidly scanning through radio stations, hunting for a favorite kind of music (Gjerdingen & Perrott, 2008; Perrott & Gjerdingen, 1999). It often sounds something like this:

Fzzz—♪ — stock mar—qrzfpt— — zzzk%&#akk— —two teasp—

xyxyxsbqrpt—trip to Paris—

Amazingly, most listeners can perform this search mission very deftly, needing only the merest snippet to determine, first, whether a station is playing music and, second, whether the genre being broadcast is one that is desirable at the moment. Most listeners are able to make a choice in a much shorter time period than that provided by the scan function of most radios. Previous researchers (Ashley, 2008; Bigand, Filipic, & Lalitte, 2005; Peretz, Gagnon, & Bouchard, 1998) found listeners were able to determine the valence of musical excerpts in as little as one-eighth of a second (i.e., 125 ms), which makes sense from an evolutionary standpoint.

Our brains are wired with rapid response mechanisms so that we can react instantly to perceived danger. When people encounter a sudden, unexpected sound, there is an immediate reaction called variously in the literature acoustic startle reflex (Yeomans & Frankland, 1996), auditory startle response (Müller et al., 2003), or auditory brainstem response (Hall, 2007). This is a rapid brainstem reflex that includes such motor responses as an eye blink, a facial grimace, head turning, increased muscle tension in the arms, and so on. Initial reactions can be registered in as little as 5 to 6 ms (Hall, 2007) and motor responses can occur from 14 to 22 ms after the presentation of an auditory stimulus (Kofler et al., 2001; Yeomans & Frankland, 1996). Such a brainstem reflex is also a rudimentary system that provides only minimal information, enough for an immediate response to avoid danger (LeDoux, 1998).

A somewhat slower pathway, running from the medial geniculate nucleus in the thalamus to the auditory cortex, provides more information (LeDoux, 1995). This cortical pathway is necessary for interpreting additional information about a complex stimulus. There are two, at least partially distinct, pathways for the 'what' and 'where' aspects of a sound (De Santis, Clarke, & Murray, 2007). In two experiments, the mean peak latency of a stimulus as it reached the auditory cortex was 45 ms (Howard et al., 2000) and neural signals indicating multisensory gestalts, combining auditory and somatosensory (skin touch) information, were obtained at 50 ms (Murray et al., 2005). Thus, this 'slower pathway' is still extremely rapid.

A sudden sound activates both fast and slow pathways. Stimulus information is first sent from the inner ear via the auditory nerve through lower brainstem way stations to the auditory thalamus (Li, Stutzmann, & LeDoux, 1996). At the auditory thalamus, the pathways divide; the fast pathway sends information directly to the amygdala, which is concerned with emotional processing, especially of fear. The slower pathway runs from the auditory thalamus to the primary auditory cortex, from there to the auditory association cortex, and then back to the amygdala. The fast and slow pathways converge in individual cells in the lateral amygdala. Information received at the lateral amygdala via the auditory cortex arrives approximately 20 ms after signals from the fast pathway are registered.

Fast and slow pathways are active during music listening (Huron, 2006; Sloboda & Juslin, 2001). Certain unexpected features in the music, such as a cymbal crash or sudden modulation, may cause a startle reaction via the fast pathway. More sophisticated recognition of musical elements, say conscious awareness of a shift from major to minor, comes via the slower, cortical pathway. Each layer, from brainstem to auditory thalamus to primary auditory cortex to auditory association cortex, adds more information to the mix. Even beyond these are additional areas, such as frontal lobe and limbic regions that are concerned with evaluation and appreciation of musical experiences. Lee, Skoe, Kraus, & Ashley (2008) demonstrated that musical training affected auditory brainstem responses, and Diethe, Teodoru,

Furl, and Shawe-Taylor (2009) used magnetoencephalography to demonstrate that while it is feasible to determine which genre a listener is attending to from brain activations, it is not a reliable method. However, very little research has examined musical judgments based on these response mechanisms.

In a set of experiments concerning music and emotion in a brain-damaged patient, Peretz, Gagnon and Bouchard (1998) also included data from four normal controls (i.e., mean age: 41.2 years, minimal musical training). From prior research, half of 32 musical excerpts— orchestral, piano with orchestra, or piano solo—were shown to evoke a sense of happiness and half to evoke a sense of sadness. These excerpts were presented to the four control participants always from the beginning but lasting only 500 ms and increasing by a second each time to a length of 7.5 seconds. There were a total of 256 trials (i.e., 32 excerpts each presented at eight different lengths: 0.5, 1.5, 2.5, 3.5, 4.5, 5.5, 6.5 and 7.5 seconds). Excerpts were presented in a random order and participants were asked to rate each one on a happy–sad continuum of 10 points.

Researchers found the participants were nearly as consistent at rating the 500 ms excerpts as they were at rating the longer excerpts. For the happy excerpts, participants gave the shortest excerpt a mean rating of slightly more than 7, while the 4.5 second excerpt had an average rating of slightly more than 8, with all other lengths rated between 7 and 8. For the sad excerpts, the 500 ms excerpt was rated just below 4, the 3.5 second excerpt at 3, with all other lengths in between. Because a previous experiment in this series had confirmed tempo and mode to be prime indicators of emotion, the investigators determined the 500 ms happy excerpts contained 4.4 events and the sad excerpts contained 1.7 events on average. Thus, it is possible listeners used event density as a prime clue for determining emotional valence. Fourteen of the 16 excerpts contained sufficient information to determine mode (i.e., a tonic triad or major third), while only five of the 16 sad excerpts did so. Thus, mode may have contributed useful information also, especially in the happy excerpts.

To confirm the results obtained with the 500 ms excerpts were not due to some artifact such as prior exposure, an extension of the study was administered with two control groups of 10 naive subjects each. The first group of naive subjects heard the excerpts presented in random order as before, only this time each excerpt was also presented in a 250 ms version. As before, participants judged emotional valence of the 250 ms sample nearly the same as all the other lengths. The second group of naive participants heard each of the 32 excerpts presented progressively, by increments of 250 ms, beginning with 250 ms and continuing systematically to 3 seconds. Although there were slightly larger variations in emotion ratings from the shortest to longest excerpts, all the results taken together indicated that listeners could make immediate judgments of emotional valence along a happy–sad dimension.

Ashley (2008) used a subset of the excerpts used by Peretz et al. (1998) at even shorter time courses: 25, 50, 100, 250, 500 and 1000 ms. In this case, participants used 10-point scales to rate the excerpts along the following dimensions: happy–sad, slow–fast, dark–bright, bass–treble, heavy–light, thick–thin, and also as major–minor. Many judgments of valence and mode made at 50 ms corresponded with those of longer time courses and nearly all the categories reached stability by 100 ms.

Bigand et al. (2005) asked musically trained and untrained listeners to place 27 musical excerpts in geometric space, where the vertical axis indicated arousal level, the horizontal axis indicated emotional valence and a third axis separated pieces with broad melodic contours from those that proceeded harmonically or by broken arpeggios. Participants did this twice, once with the excerpts at approximately 25 seconds in length and again at 1 second lengths. There was no

distinction in geometric solutions between musically trained and untrained listeners. Also, multidimensional scaling was only weakly affected by length of excerpt; emotional responses did not differ between 1-second and 25-second excerpts.

In a follow-up experiment, Bigand et al. (2005) asked participants to listen to excerpts starting at 250 ms and increasing by 500 ms to 1 second, 2 seconds, 5 seconds and approximately 20 seconds. Participants were asked to indicate on a subjective scale how 'moving' the excerpts sounded. Excerpts were chosen so that half would be considered strongly moving and the other half less moving. One group of musically trained and untrained subjects listened to classical excerpts, while a second group did the same for pop/rock examples. They were told to:

> imagine that they were looking for moving (emotional) music on a radio. Shifting from one radio station to another would be an ecological situation that was similar to the experiment. The critical question of the experiment was to specify how many milliseconds they needed to decide whether the excerpts they were listening to were highly moving. (Bigand et al., 2005, p. 434)

Results were highly consistent for both musically trained and untrained listeners and for both classical and popular genres. In all cases, judgments made with the 250 ms excerpts were the same as judgments made for the longer excerpts.

Perrott and Gjerdingen (1999) presented a touchstone paper that likened the rapid identification of music genres to scanning a radio dial. As described in the published version (Gjerdingen & Perrott, 2008), they asked 52 university students to listen to 400 examples in five excerpt lengths (i.e., 250 ms, 325 ms, 400 ms, 475 ms, and 3 seconds), with four vocal and four instrumental excerpts of each of 10 genres (i.e., blues, classical, country, dance, jazz, Latin, pop, R&B, rap, and rock). For 3-second excerpts, listeners' choices agreed with categories assigned by music companies about 70% of the time. Because shorter excerpts were extracted from the 3-second excerpts, researchers could compare listeners' choices for shorter excerpts with the categories they assigned to the longest samples. Although there was considerable variation among genres, recognition scores were above chance for all genres at all excerpt lengths. Researchers found little or no effects of training, but listeners were more accurate with instrumental than vocal excerpts.

From these few experiments it is apparent that listeners can make very rapid judgments about the music they hear. Whether these judgments are made subcortically (Peretz et al., 1998) or require cognitive processing (Bigand et al., 2005) remains in doubt. Also, the time course of the presumably more sophisticated judgment of genre identification has been investigated in only one study (Gjerdingen & Perrott, 2008). The purpose of the present study was to extend this study. None of Gjerdingen and Perrott's participants were music majors. They did conduct a pilot study with a small group of music students (N unspecified) and found no significant difference in the results compared to non-music majors. They also did not find a gender effect, but as two-thirds of their listeners were female, they acknowledge that their data set was not ideal. They collected the original data for their experiment in 1999, and acknowledged that several of their genres are no longer viable labels. Finally, they did not investigate the issue of preference.

Therefore, the purpose of the present study was to examine how well individuals were able to identify different music genres from very brief excerpts. Musically trained and untrained

listeners were asked to identify the genre from classical, jazz, country, metal and rap/hip hop excerpts that were 125, 250, 500, or 1000 ms in length. Specific research questions were:

1. How well were participants able to identify the correct genre for each length of excerpt?
2. Were there significant differences in accuracy by excerpt length and by genre?
3. Were possible differences among excerpt length or genre affected by musical training?
4. Were possible differences among excerpt length or genre affected by gender?
5. Did preferences affect participants' identification accuracy?

Methods

Participants

Participants for the present experiment were 347 students recruited from three college campuses in the southeast region of the USA. Table 1 provides additional demographic data. For this study, age was not a factor of interest. Data included in the present study were from participants 18 years or older. Participants were undergraduate students, thus, most of them were between 18 and 24 years of age. Music training was operationally defined as having 3 or more years of formal music instruction. Approximately 72% of the participants reported that they were trained musicians.

Procedures for creating the stimulus CD and test form

Five genres were selected for the study: classical, jazz, country, metal and rap/hip hop. When considering genres, opera and pop were originally included. However, during initial sampling, opera was easily confused with classical when using samples one second or shorter. Pop contained elements of jazz, rap/hip hop, country, or metal within the brief excerpts that were used. Thus, to eliminate confounding cross-genre variables, the decision was made to eliminate opera and pop from the final genre list.

An ecological approach was used in choosing the format of music files for the stimulus CD. In a survey conducted by Alloy Media and Marketing (2009), researchers found that 74% of college students report ownership of an MP3 player. Whether downloaded as MP3 files, or transferred from CD to computer to MP3 player, MP3 encoded music files are apparently the most prevalent form used by the population in this study. Therefore, downloading MP3 files and using no subsequent compression was the most ecologically valid approach to creating the stimulus CD.

Ten examples of each genre were selected in the following manner. After determining the five genres to be used, investigators selected recordings from lists maintained by iTunes, a softwarebased online digital media store operated by Apple, Inc. The iTunes store was chosen because of its place as the leader in music sales with the population tested. As of the first quarter of 2009, iTunes purchases accounted for 25% of sales in the music market. The iTunes store outsold all other individual outlets with the second-place outlet accounting for 14% of sales in the music market, which includes sales of MP3 and compact discs ('Digital music increases share of overall music sales volume in the US,' n.d.). Classical recordings came from the iTunes

Essentials play list 'Classical: 20th Century', skipping any vocal works. For purposes of this study, examples chosen for the classical genre represent western art music, rather than that of the Classical music period generally understood to be art music composed between 1750 and 1820. Examples from the 20th century iTunes list were found to represent multiple composers, styles and instrument combinations in the classical genre. Also, because music from the other genres came from the 20th century, choosing classical music from the 20th century avoided the problem of having listeners choose those items only because they were not modern. That is, a Baroque piece might be identified not so much because it represented classical music, but because it did not fit the sounds of the other four 20th-century styles. Jazz recordings were selected from the iTunes Essentials play list 'Jazz 101'. Because vocal jazz music was determined to be confused with country, vocal jazz recordings were omitted from consideration, and one jazz/pop-crossover and one fusion recording were eliminated. For rap/hip hop and country, the 'Top Songs' iTunes lists of the week were used. Starting with number 1, recordings were chosen in the order listed. Artist repetition was avoided by skipping to the next recording on the list and no single artist was represented more than once within a genre. Metal recordings were selected from the 'Heavy Metal' featured albums list, under the genre heading of 'Rock'. Albums were selected according to most recent release dates, and the first track on each album was selected. All recordings were purchased between 10 and 22 September 2008. Because these lists are based on the most popular purchases transacted on iTunes, the lists change frequently. For a complete list of recordings used in the present study, see Table 2.

Table 1. Demographic data on participants

	Males	Females	Total
Gender	129	218	347
	Trained[a]	Untrained	
Training	251	91	342[b]

Notes: [a] Participants classified as 'trained' included those having 3 or more years of formal music training. [b] Because some participants did not provide training information, this number does not match the total number of participants ($N = 347$).

To assure the selection of examples was representative of genre, each recording was previewed by two of the researchers and a volunteer, who were all professional musicians familiar with the genres included in the present study. To be included, recordings needed to meet the following criteria: (1) employ instruments commonly associated with the genre; (2) employ stylistic performance practices (e.g., articulation, note duration, and phrasing) associated with the genre; and (3) neither the music nor the artist would be considered crossover. A list of random numbers was used to determine the starting point within each recording. For many pieces chosen, the introductory material could have caused confusion as a result of crossover and minimal dynamic levels. Therefore, the range of random numbers was set at 30 to 240, to reflect item selections beginning no less than 30 seconds and no more than 4 minutes from the beginning of any track. If the starting time for the excerpt exceeded the length of a particular recording, that random number was skipped and the next random number was used. A spreadsheet was kept with the random numbers used for the starting point and the length of the excerpt.

Four excerpts, one at each of the lengths to be examined in the present study (i.e., 125, 250, 500, and 1000 ms) were prepared from each of the 50 recordings, using Sound Forge (Sony Creative Software, 2008) on a Dell 690 computer. The resultant 200 excerpts included 10

excerpts for each of the four lengths, for each of the five genres. Each excerpt began at a different place in the recording according to the list of random numbers. That is, excerpts shorter than 1 second were not chosen from within the 1 second excerpt. Thus, no learning effect was likely to have taken place.

Once the 200 excerpts were created, another random number list was created to organize them into 10 groups of 20 excerpts each. This was done so listeners could complete groups of 20 followed by a break, rather than completing all 200 items at once. Participants were allowed 4 seconds to determine the genre and mark the answer sheet for each item. To allow participants time to focus on each new group, 4 additional seconds of silence were inserted between each group of 20 items.

Each of the 10 groups of 20 items was created to have an equal number of genres and an equal number of excerpt lengths. Once the excerpts were placed into the 10 groups, another randomization placed the 20 excerpts of each group and the 10 groups into their final order for the recording. In the recording, 8 seconds of silence were inserted to allow for turning the answer sheet to continue from group 4 to group 5. In addition to the 200 items, 3 second samples of each genre were used as practice examples. Recordings for practice examples were taken from tracks other than those used for the 200 items.

Instructions for taking the Genre Identification Test were written and then recorded. These instructions and the 200 excerpts were burned onto a CD. Each sound file was normalized such that once set, the intensity control of the stereo amplifier did not need to be adjusted. Two different pilot-test administrations resulted in minor revisions before creating the final CD. A bubble-sheet test form was developed and piloted in the same pilot administrations. Using Remark Classic OMR software (Gravic, 2009), a final version was developed that could be machine scored.

Test procedures

During test administration, all participants were provided pencils and test forms. At campuses A and B, participants enrolled in a music appreciation course completed the task during a regular class meeting. Participants at campus C completed the task during a recital-attendance class held in a performance hall or during a regular class meeting. Consent information sheets and a brief description of the project were provided prior to administration of the test. Participants completed a demographics section of the test form providing the following information: (a) gender; (b) whether they were 18 years of age or older; (c) consent to use data for research; (d) whether they had three years or more of formal music training (e.g., read music, took lessons on an instrument, etc.); (e) most preferred music genre (from the five genres included in this study); and (f) least preferred music genre. To satisfy Institutional Review Board guidelines, data from participants younger than 18 were not used. One 3-second example of each genre was played and the recorded instructions provided correct answers before proceeding with the test items.

The sound systems used at campuses A and B were permanently installed multi-component stereos and stereo speakers. The sound system at campus C included a Yamaha CDX-490 CD player, a NAD S300 Amplifier, and two JBL Professional monitors (i.e., speakers). Playback equipment used at all campuses provided clarity of sound quality and more than sufficient amplification.

Following data collection, test forms were scanned using a National Computer Systems (NCS) Pearson Opscan 6 (NCS is now owned by Scantron Corporation). Because the software

has limited capacity for data analysis, it was necessary to create multiple answer keys to obtain data necessary for answering stated research questions. Cronbach's Alpha coefficient used to determine that reliability for the Genre Identification Test was .90.

Table 2. Music selections list

Genre	Song	Artist
Classical	Fanfare for the Common Man	Copland
	Adagio for Strings from String Quartet, Op. 11	Barber
	Piano Concerto in F Major: III	Gershwin
	Concierto de Aranjuez: II	Rodrigo
	Trois Gymnopedies: I	Satie
	Short Ride in a Fast Machine	Adams
	The Planets, Op. 32: IV	Holst
	The Unanswered Question	Ives
	Festive Overture, Op. 96	Shostakovich
	Symphony No. 1 in D Major, Op. 25: I	Prokofiev
Jazz	Take Five	Dave Brubeck
	So What	Miles Davis
	In the Mood	Glenn Miller
	Bedtime	King Cole Trio
	Chill	Clear Voyage
	In A Sentimental Mood	John Coltrane and Duke Ellington
	Sing, Sing, Sing	Benny Goodman
	Goodbye Pork Pie Hat	Charles Mingus
	Straight No Chaser	Thelonius Monk
	Body & Soul	Coleman Hawkins
Country	Johnny & June	Heidi Newfield
	All I Want To Do	Sugarland
	Just a Dream	Carrie Underwood
	Don't Think I Don't Think About It	Darius Rucker
	Do You Believe Me Now	Jimmy Wayne
	Should've Said No	Taylor Swift
	She Never Cried in Front of Me	Toby Keith
	Everybody Wants to Go to Heaven	Kenny Chesney
	Troubadour	George Strait
	Gunpowder & Lead	Miranda Lambert
Metal	Troops of Doom	Sepultura
	Walk With Me in Hell	Sacrament
	Demon's Kiss	Blue Oyster Cult
	Holy Wars ... the Punishment Due	Megadeth
	Night Songs	Cinderella
	Welcome to the Jungle	Guns N' Roses
	Star Spangled Banner	Exodus
	Heavy Metal Thunder	Saxon
	Madhouse	Anthrax
	Rock You Like A Hurricane	Scorpions
Rap/hip hop	Whatever You Like	T.I.
	Swagga Like Us	Jaz-Z & T.I.
	In the Ayer	Flo Rida
	Can't Believe It	T-Pain
	Got Money	Lil Wayne & T-Pain
	Dangerous	Kardinal Offishall
	Swing	Savage & Solilja Boy
	My Life	The Game
	What Them Girls Like	Ludacris
	Lolli Lolli (Pop that Body)	Three 6 Mafia

Results

The purpose of the present study was to determine the extent to which listeners could identify the genre of very brief examples across five music genres and four excerpt lengths. The results for 347 participants who completed the Genre Identification Test are presented in Table 3. Regarding research question 1, participants were able to identify the genre of brief excerpts very well. Even at 125 ms, participants performed considerably above chance level (i.e., 20%).

Table 3. Accuracy of genre identification

Category	N	M	SD	% correct
All genres at 125 ms	347	27.06	5.66	54% of 50 items
All genres at 250 ms	347	37.95	4.31	77% of 50 items
All genres at 500 ms	347	41.39	3.83	83% of 50 items
All genres at 1000 ms	347	44.55	3.73	89% of 50 items
All genres at all lengths	347	151.43	14.95	76% of 200 items

Additional details on test performance are presented in Table 4. Participants were most accurate in rap/hip hop 1000 ms (94%), classical 1000 ms (93%), metal 1000 ms (92%), and rap/hip hop 500 ms (91%). Performance was between 80 – 90% accurate in seven categories (i.e., classical 500 ms, jazz 500 and 1000 ms, country 1000 ms, metal 250 and 500 ms, rap/ hip hop 250 ms). Performance was below 50% accuracy in only one category (i.e., rap/hip hop 125 ms). Participants were able to make correct judgments at 125 ms at very impressive levels for classical – 70%, jazz – 53%, country – 56%, and metal – 51%. While they were less successful for rap/hip hop at 125 ms – 40%, such an accuracy level is still well above chance.

For research questions 2, 3, 4, and 5, a conservative alpha level of .01 was established to control for Type I error. For all data that violated the sphericity assumption, using Mauchley's Test, the Greenhouse–Geisser correction was used, and degrees of freedom were expressed with noninteger values. The purpose of the second research question was to explore whether significant differences in accuracy identification existed due to excerpt length and genre. A 4×5 (Length $\times$ Genre) repeated-measures ANOVA was performed, with excerpt length and genre both serving as within-subjects variables. Significant main effects were found for excerpt length, $F (2.22, 763.38) = 2514.81$, $p = .000$, $h2 = .88$ and for genre, $F (3.71, 1282.39) = 54.98$, $p = .000$, $h2 = .14$. Significant main effect differences for excerpt length and genre were examined through post-hoc pairwise comparisons of the means using Bonferroni's adjustment. Differences were found among all excerpt lengths with accuracy increasing significantly ($p < .001$) for each successive excerpt length (i.e., 125 ms M = 5.41, SD = 1.13; 250 ms M = 7.59, SD = 0.86; 500 ms M = 8.28, SD = 0.77; 1000 ms M = 8.91, SD = 0.75). Regarding genre, classical (M = 7.90, SD = 0.99), metal (M = 7.80, SD = 0.94), and rap/hip hop (M = 7.77, SD = 0.98) excerpts were identified significantly more accurately ($p = .000$) than were country (M = 7.22, SD = 1.25) or jazz excerpts (M = 7.18, SD = 1.18). No significant differences in accuracy were found among classical, metal, and rap/hip hop excerpts or between country and jazz excerpts.

A statistically significant two-way interaction effect was found between excerpt length and genre, $F (10.08, 3487.55) = 212.26$, $p = .000$, $h2 = .38$. As Figure 1 shows, accuracy of genre identification is distinct for each length being tested. Particularly noteworthy is that classical decreased from being the most accurately identified genre at 125 ms to being the least accurately identified genre at 250 ms. Simultaneously, rap/hip hop increased from being the least accurately

identified genre at 125 ms to being most accurately identified genre at 250 ms. Further, rap/hip hop continued as the most accurately identified genre at 500 ms and 1000 ms.

Table 4. Percentage of correct Genre Identification Test items according to genre and excerpt length

Category	N	M	SD	% correct of items
Classical overall	347	31.58	3.98	79% of 40
Classical 125 ms	347	6.99	1.67	70% of 10
Classical 250 ms	347	6.79	1.44	68% of 10
*Classical 500 ms	347	8.53	1.20	85% of 10
*Classical 1000 ms	347	9.27	1.05	93% of 10
Jazz overall	347	28.72	4.73	71% of 40
Jazz 125 ms	347	5.31	1.97	53% of 10
Jazz 250 ms	347	7.20	1.60	72% of 10
Jazz 500 ms	347	8.12	1.20	81% of 10
*Jazz 1000 ms	347	8.09	1.32	81% of 10
Country overall	347	28.86	4.99	72% of 40
Country 125 ms	347	5.65	2.03	56% of 10
Country 250 ms	347	7.66	1.61	77% of 10
Country 500 ms	347	7.01	1.62	70% of 10
*Country 1000 ms	347	8.55	1.30	85% of 10
Metal overall	347	31.20	3.77	78% of 40
Metal 125 ms	347	5.08	1.60	51% of 10
*Metal 250 ms	347	8.27	1.16	83% of 10
*Metal 500 ms	347	8.64	1.27	86% of 10
*Metal 1000 ms	347	9.21	0.99	92% of 10
Rap/hip hop overall	347	31.06	3.92	78% of 40
Rap/hip hop 125 ms	347	4.03	1.83	40% of 10
*Rap/hip hop 250 ms	347	8.51	1.36	85% of 10
*Rap/hip hop 500 ms	347	9.09	0.96	91% of 10
*Rap/hip hop 1000 ms	347	9.43	0.95	94% of 10

Note: Rows with * indicate accuracy between 80% and 94%.
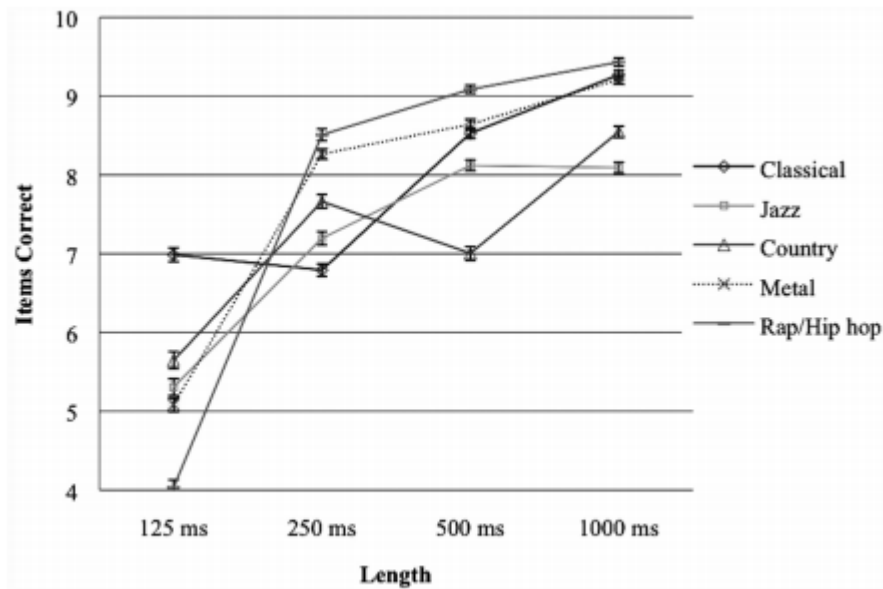
Figure 1. Accuracy of genre identification across excerpt lengths

The purpose of the third research question was to examine whether significant accuracy identification differences regarding excerpt length by genre were due to musical training. A three-way repeated-measures ANOVA was performed, with training serving as a between subjects variable and excerpt length and genre both serving as within-subjects variables. The between-subjects main effect of training, examining the overall accuracy identification difference between those with training (M = 7.60, SD = 0.78) and those without training (M = 7.56, SD = 0.73), was found to be non-significant, F (1, 340) = .585, p = .445, h2 = .00. The two-way interaction effect for excerpt length by training was also found to be non-significant, F (2.20, 748.97) = 4.261, p = .012, h2 = .01.

The two-way interaction effect for genre by training was found to be significant, F (3.76, 1278.92) = 21.72, p = .000, h2 = .06. Univariate analyses revealed those with training were significantly more accurate in identifying jazz excerpts, F (1, 340) = 28.29, p = .000, h2 = .08, while those without training were significantly more accurate in identifying rap/hip hop excerpts, F (1, 340) = 21.05, p = .000, h2 = .06. Figure 2 shows the two-way interaction for genre by training. Further, the three-way interaction effect for genre by length by training was found to be significant, F (10.13, 3445.41) = 3.30, p = .000, h2 = .01. Univariate analyses revealed those with training were significantly more accurate in identifying classical 1000 ms excerpts, F (1, 340) = 8.81, p = .003, h2 = .03, jazz 125 ms excerpts, F (1, 340) = 28.84, p = .000, h2 = .08, jazz 250 ms excerpts, F (1, 340) = 15.24, p = .000, h2 = .04, and jazz 500 ms excerpts, F (1, 340) = 15.06, p = .000, h2 = .04. Those without training were significantly more accurate in identifying rap/hip hop 125 ms, F (1, 340) = 10.81, p = .001, h2 = .03, rap/ hip hop 250 ms, F (1, 340) = 15.63, p = .000, h2 = .04, and rap/hip hop 500 ms, F (1, 340) = 20.06, p = .000, h2 = .06.

The purpose of the fourth research question was to examine whether significant accuracy identification differences regarding excerpt length by genre were due to gender. A three-way repeated-measures ANOVA was performed, with gender serving as a between-subjects variable and excerpt length and genre both serving as within-subjects variables. The between subjects main effect of gender, examining the overall accuracy identification difference between males (M = 7.60, SD = 0.78) and females (M = 7.56, SD = 0.73), was found to be non-significant, F (1, 345) = .23, p = .630, h2 = .00. The two-way interaction effect for excerpt length by gender was also found to be non-significant, F (2.22, 765.77) = 4.13, p = .013, h2 = .01. However, the two-way interaction effect for genre by gender was found to be significant, F (3.74, 1288.67) = 7.44, p = .000, h2 = .02. Figure 3 shows the interaction for genre by gender. Univariate analyses revealed males to be significantly more accurate in identifying metal excerpts, F (1, 345) = 7.49, p = .007, h2 = .02. The three-way interaction effect for genre by length by gender was also found to be significant, F (10.05, 3468.66) = 2.35, p = .009, h2 = .01. Univariate analyses revealed males to be significantly more accurate with identifying jazz 250 ms excerpts, F (1, 345) = 8.97, p = .003, h2 = .03, and metal 125 ms excerpts F (1, 345) = 14.68, p = .000, h2 = .04, while females were significantly more accurate with identifying rap/hip hop 250 ms, F (1, 345) = 10.82, p = .001, h2 = .03.

The purpose of the fifth research question was to determine whether or not preference played a role in participants' ability to identify genre accurately. For each participant, a mean score was calculated from results of all excerpts not designated as being the most preferred genre and was labeled 'other than most preferred.' A paired-samples t-test was performed to determine whether a significant difference existed between participants' most preferred genre scores and the mean scores of their other-than-most-preferred genres. From the results, participants' identification of most preferred genre (M = 7.2, SD = 0.93) was significantly more accurate than

identification of their other-than-most-preferred genres (M = 7.48, SD = 0.31), t (339) = 9.08 (two-tailed), p = .000, d = .49.
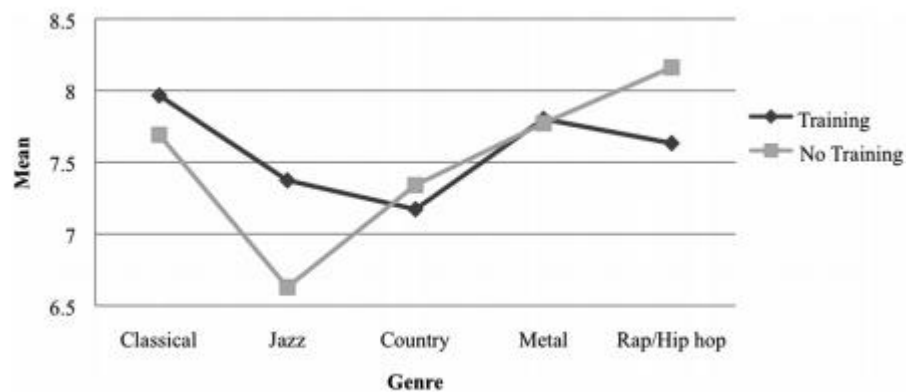

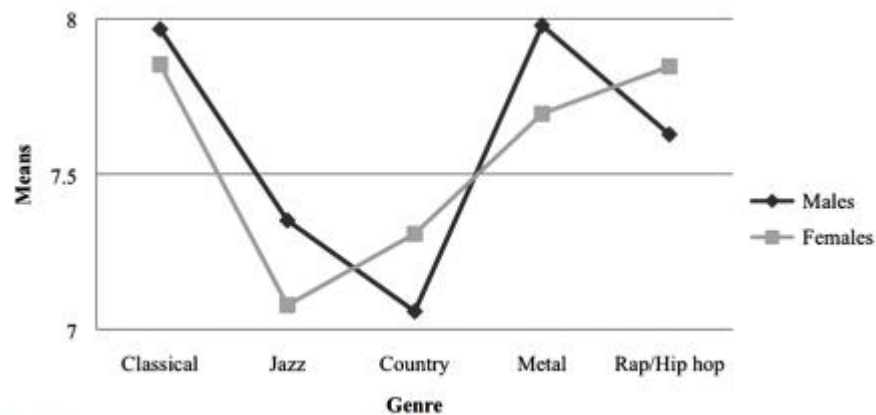**Figure 2.** Two-way interaction effect of genre by training

Figure 3. Two-way interaction effect of genre by gender

Similarly, a mean score was calculated from results of all excerpts not designated as being the least preferred genre and was labeled 'other than least preferred.' A paired-samples t-test was performed to determine whether a significant difference existed between participants' least preferred genre scores and the mean scores of their other-than-least-preferred genres. From the results, participants' identification of least preferred genre (M = 7.41, SD = 1.07) was found to be significantly less accurate than identification of their other-than-least-preferred genres (M = 7.61, SD = 0.79), t (341) = -3.72 (two-tailed), p = .000, d = .21.

Discussion

Anecdotally, listeners have an amazing ability to make very rapid judgments about the music they are hearing. For example, while scanning through radio stations, many listeners make choices upon hearing extremely brief snippets of music selections. Biologically, this ability can be accounted for as a function of the auditory brainstem response (Hall, 2007). Previous researchers have demonstrated listeners can make judgments about the emotional quality of a musical excerpt that is as brief as 250 ms (Bigand et al., 2005; Peretz et al., 1998) or even 50 ms (Ashley, 2008). Because only Gjerdingen and Perrot (2008) have investigated the rapid

identification of genres from very brief excerpts, the purpose of the present study was to extend their study.

Participants in the present study were university students; both trained (n = 251) and untrained (n = 91). They listened to excerpts from classical (instrumental), jazz, country, metal and rap/hip hop genres. All five genres were represented by 10 excerpts each of 125, 250, 500, and 1000 ms presented in random and counterbalanced order for a total of 200 test items. Participants also indicated their most and least preferred genre in order to determine whether preference affected accuracy of judgment.

According to the results, participants were very successful at identifying genres of very brief excerpts. Overall, they were correct 76% of the time. At 1000 ms, accuracy reached 89% for all genres combined. Even at 125 ms, listeners performed well above chance: classical (70%), jazz (53%), country (56%), metal (51%), and rap/hip hop (40%). Clearly, these participants were able to make fairly accurate judgments on the basis of extremely limited information.

In general, the length of excerpt made a significant difference in performance, with longer time lengths expectedly leading to greater accuracy. Regarding genre differences, classical, metal and rap/hip hop were more easily identified than were country or jazz. Although it is difficult to provide concrete reasoning for the results, perhaps the amplified nature of metal and rap/hip hop produces uniquely associated timbres, while classical has a historically grounded and recognizable set of timbres. Jazz and country, on the other hand, include a substantially wider range of possible associated timbres. From the interaction effect of genre by length, it is evident that the accuracy of genre identification is distinct for each length being tested. Replication studies and/or data from additional participants are needed to determine whether the distinct results of the present study can be supported.

No overall significant difference was found between those with musical training and those without. Similarly, identification accuracy across excerpt length was not affected by training. However, identification accuracy of genre was found to be affected by musical training, with trained musicians significantly better able to identify classical and jazz excerpts of varying lengths and those untrained significantly better able to identify rap/hip hop excerpts of varying lengths. Perhaps familiarity might be at play such that trained musicians may be more likely to have encountered classical and jazz music while those untrained may be more likely to have encountered rap/hip hop music.

Many of the results for gender are similar to those regarding training. No overall significant difference was found between males and females. Similarly, identification accuracy across excerpt length was not affected by gender. However, identification accuracy of genre was found to be affected by gender, with males significantly better able to identify jazz 250 ms and metal 125 ms excerpts and females significantly better able to identify rap/hip hop 250 ms excerpts. In general, fewer significant differences occur as a result of gender than of training. Those few gender differences found in the present study were somewhat unexpected and will require additional research to substantiate or refute.

Preference was found to have played a role in participants' ability to identify genre of very brief excerpts accurately. Participants' most preferred genre was identified significantly more often than all other genres; and their least preferred genre was identified significantly less often than all other genres. Perhaps such a result was due to familiarity.

These findings support a primary conclusion that people are very adept at identifying particular music genres when presented with excerpts that are 1 second in length or less. The literature on auditory brainstem response provides a biological explanation for how this is

possible. In previous studies, researchers found listeners can make emotional judgments very rapidly and the current results extend this ability to the identification of specific genres. Based on Ashley's (2008) findings that listeners can make judgments of valence and mode as rapidly as 100 ms, and in some cases in 50 ms, it is tempting to speculate that genre identification takes slightly longer. That is, perhaps one can first make judgments of emotional valence and then move on to more discrete categorizations. Perhaps, too, valence judgments are made via the fast path analysis and genre identification is handled via the slower pathway (LeDoux, 1995, 1998). However, further studies are needed to test these hypotheses.

In general, the results of the present study support and extend the findings of Gjerdingen and Perrott (2008). In both studies, listeners were able to recognize the genre of very brief excerpts. In Gjerdingen and Perrott, the shortest excerpts were 250 ms, while in the current study they were 125 ms. Gjerdingen and Perrott found no differences due to gender or training, but did not examine either variable fully. In the current study minor differences were found due to gender and to training. Gjerdingen and Perrot did not investigate the role of preference; the current study found that preference did play a significant role.

One way of accounting for the rapid genre identification abilities of the participants in the present study, and for the roles of gender, training and preference, may be found in the Preference for Prototypes Theory (PPT). First proposed by Whitfield and Slatter (1979) and further developed by Martindale and Moore (1988), PPT arises out of work in cognitive psychology (e.g., Martindale, 1984; Posner & Keele, 1968). The basic idea is that people classify objects by matching incoming information with previously formed schema or prototypes. Preferred categories are represented in more well-defined, discrete prototypes than non-preferred categories. For example, over time, listeners place different features of the music they hear into categories that gradually form recognizable prototypes. For a person raised in a home where country music is frequently heard, musical features such as the twang of a steel guitar or particular vocal qualities become associated with the label country music.

Reber, Schwartz, and Winkielman (2004) proposed that the more fluently one can process aspects of a contemplated object (e.g., a painting or a musical experience), the more positive their aesthetic response. This underlying cause has been demonstrated with literature (Piters & Stokmans, 2000), paintings (Farkas, 2002; Hekkert & Wieringen, 1990), and furniture (Whitfield & Slatter, 1979). Smith and Melara (1990) and Repp (1997) confirmed the effect of processing fluency on musical choices.

PPT is not universally supported (e.g., see Boselie, 1991), however, others (Hekkert & Snelders, 1995; Martindale, Moore, & West, 1988; North & Hargreaves, 2000; Whitfield, 1983, 2000, 2009) argue that it is a viable theory and is compatible with the dominant arousal-mediating theory of Berlyne (1971, 1974). Moreover, PPT does fit data from the current study rather nicely. First, there is an evolutionary basis for PPT; it would have been greatly to our ancestors' advantage to identify prey or predators rapidly. Early on, we placed the shape, color, sounds, and so on, of each animal encountered into various categories. Matching a novel stimulus such as a growl or a chirp to preformed categories allowed for faster identification.

Second, PPT can account for gender, training, and preference. That is, it is reasonable to hypothesize that males and females, trained and untrained listeners, and those who prefer a particular genre will hear certain sounds more frequently than others and will thus have developed more crystallized categories for certain sounds than for others. For example, a highly trained female violinist will have developed stronger schema for classical music than an untrained male who listens to a great deal of rap. If both are presented with a brief burst of

orchestral music, he might rapidly decide 'not rap', without being able to place it into the classical music genre as quickly as she could, because it fits her preferred prototype. Of course, there is an extensive literature on music preferences that more fully explicates how we make musical choices (e.g., Abeles & Chung, 1996; Hargreaves, North, & Tarrant, 2006; Lamont, 2009; North & Hargreaves, 2008; Rentfrow & Gosling, 2003, 2006).

While administering the Genre Identification Test, it was remarkable that many participants were observed whispering the name of the artist or even the titles of some of the examples. Both musically trained and untrained listeners identified specific titles and artists. For instance, one young man wrote on his paper the decade that a metal artist recorded a particular piece. Another looked up from his test form and named the composition, adding, 'Nice choice!' Several participants identified rap/hip hop artists, and after one classroom administration, participants asked whether specific artists and songs were included in the examples. Indeed, all artists and titles the participants questioned had been included. This shows an amazing capacity for individuals to identify not only music genre, but also artist and composition in a very short period of time.

Many participants were quite interested in knowing how well they did on the test, and were disappointed that they would not be able to discern their individual scores. Frequently, participants commented about the length of the test, with about half saying they tired toward the end and the other half reporting that they felt they had improved at identification as they progressed through the 200 items.

For the current study, vocal pieces were eliminated from the genres of classical and jazz. To determine the effect of vocals on genre recognition, a future study could be designed to determine accuracy in identification of vocal music, including genres such as opera, country, metal and musical theater (Broadway). Another focus for such studies could be that of specific performance areas to determine accuracy in identification of specific compositions with a high degree of familiarity. For example, one might choose standard piano compositions from which to create a similar stimulus recording for participants who are accomplished pianists. Participants could be asked to identify titles of specific pieces from brief music examples. Numerous additional studies could be designed to explore this fascinating and quite extraordinary ability to make rapid musical judgments based on extremely brief snippets of music. The results of such a line of research would have both practical applications, as in the burgeoning area of Music Information Retrieval (Downie, 2008), and would also lead to a deeper understanding of human musical processing.

References

Abeles, H., & Chung, J. (1996). Responses to music. In D. Hodges (Ed.), Handbook of music psychology (2nd ed.; pp. 285–342). San Antonio, TX: IMR Press.
Alloy Media and Marketing. (2009). Totally wired campus – the class of 2013 gets high 'tech' grades: Alloy media and marketing. Retrieved 25 April 2010, from http://74.125.45.132/search?q=cache:KLkzD8g5HgJ:www.alloymarketing.com/investor_relations/news_releases/doc/Alloy-Harris-CETECH11-1209FINALPR.doc+The+2009+Alloy+College+Explorer+study&cd=1&hl=en&ct=clnk&g l=us&client=firefox-a

Ashley, R. (2008). Affective and perceptual responses to very brief musical stimuli. Paper presented at the International Conference on Music Perception and Cognition, , 25–29 August, Sapporo, Japan.

Berlyne, D. (1971). Aesthetics and psychobiology. New York: Appleton-Century-Crofts.

Berlyne, D. (1974). Studies in the new experimental aesthetics. New York: Wiley.

Bigand, E., Filipic, S., & Lalitte, P. (2005). The time course of emotional responses to music. Annals of the New York Academy of Sciences, 1060, 429–437.

Boselie, F. (1991). Against prototypicality as a central concept in aesthetics. Empirical Studies of the Arts, 9(1), 65–73.

De Santis, L., Clarke, S., & Murray, M. (2007). Automatic and intrinsic auditory 'what' and 'where' processing in humans revealed by electrical neuroimaging. Cerebral Cortex, 17, 9–17.

Diethe, T., Teodoru, G., Furl, N., & Shawe-Taylor, J. (2009). Sparse multiview methods for classification of musical genre from magnetoencephalography recordings. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P. Eerola (Eds.), Proceedings of the 7th Triennial Conference of the European Society for the Cognition of Music, Jyväskylä, Finland. Digital music increases share of overall music sales volume in the US. (n.d.). Retrieved 16 August 2010, from http://www.npd.com/press/releases/press_090818.html

Downie, J. (2008). The music information retrieval evaluation exchange (2005–2007): A window into music information retrieval research. Acoustical Science and Technology, 29(4), 247–255.

Farkas, A. (2002). Prototypicality-effect in surrealist paintings. Empirical Studies of the Arts, 20, 127–136.

Gjerdingen, R., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. Journal of New Music Research, 37(2), 93–100.

Gravic. (2009). Remark Classic OMR: Version 3.0. Malvern, PA. Gravic Inc. Hall, J. (2007). New handbook of auditory evoked responses. Boston, MA: Pearson Education.

Hargreaves, D., North, A., & Tarrant, M. (2006). Musical preference and taste in childhood and adolescence. In G. McPherson (Ed.), The child as musician, (pp. 135–154). Oxford: Oxford University Press.

Hekkert, P., & Snelders, H. (1995). Prototypicality as an explanatory concept in aesthetics: A reply to Boselie (1991). Empirical Studies of the Arts, 13(2), 149–160.

Hekkert, P., & Wieringen, P. (1990). Complexity and prototypicality as determinants of appraisal of cubist paintings. British Journal of Psychology, 81, 483–495.

Howard, M., Volkov, I., Mirsky, R., Garell, P., Noh, M., Granner, M., et al. (2000). Auditory cortex on the human posterior superior temporal gyrus. The Journal of Comparative Neurology, 416, 79–92.

Huron, D. (2006). Sweet anticipation: Music and the psychology of expectation. Cambridge: The MIT Press. Kofler, M., Müller, J., Wenning, G., Reggiani, L., Hollosi, P., Bösch, S., et al. (2001). The auditory startle reaction in Parkinsonian disorders. Movement Disorders, 16(1), 62–71.

Lamont, A. (2009). Musical preferences. In S. Hallam, I. Cross, & M. Thaut (Eds.), The Oxford handbook of music psychology (pp. 160–168). Oxford: Oxford University Press.

LeDoux, J. (1995). Emotion: Clues from the brain. Annual Review of Psychology, 46, 209–235.

LeDoux, J. (1998). The emotional brain. In J. Jenkins, K. Oatley, & N. Stein (Eds.), Human emotions: A Reader (pp. 98–111. Malden, MA: Blackwell.

Lee, K., Skoe, E., Kraus, N., & Ashley, R. (2008). The effect of musical training on the subcortical processing of musical intervals. Paper presented at the International Conference on Music Perception and Cognition, 25–29 August, Sapporo, Japan.

Li, X., Stutzmann, G., & LeDoux, J. (1996). Convergent but temporally separated inputs to lateral amygdala neurons from the auditory thalamus and auditory cortex use different postsynaptic receptors: In vivo intracellular and extracellular recordings in fear conditioning pathways. Learning and Memory, 3: 229–242.

Martindale, C. (1984). The pleasures of thought: A theory of cognitive hedonics. Journal of Mind and Behavior, 5(1), 49–80.

Martindale, C., & Moore, K. (1988). Priming, prototypicality, and preference. Journal of Experimental Psychology: Human Perception and Performance, 14, 661–670.

Martindale, C., Moore, K., & West, A. (1988). Relationship of preference judgments to typicality, novelty, and mere exposure. Empirical Studies of the Arts, 6(1), 79–96.

Müller, J., Kofler, M., Wenning, G., Seppi, K., Valls-Solé, J., & Poewe, W. (2003). Auditory startle response in cervical dystonia. Movement Disorders, 18(12), 1522–1526.

Murray, M., Molholm, S., Michel, C., Heslenfield, D., Ritter, W., Javitt, D., et al. (2005). Grabbing your ear: Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. Cerebral Cortex, 15, 963–974.

North, A., & Hargreaves, D. (2000). Collative variables versus prototypicality. Empirical Studies of the Arts, 8(1), 13–17.

North, A., & Hargreaves, D. (2008). The social and applied psychology of music. Oxford: Oxford University Press.

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. Cognition, 68, 111–141.

Perrott, D., & Gjerdingen, R. (1999). Scanning the dial: An exploration of factors in the identification of musical style. Paper presented at the meeting of the Society for Music Perception and Cognition, 15 August, Evanston, IL.

Piters, R., & Stokmans, M. (2000). Genre categorization and its effect on preference for fiction books. Empirical Studies of the Arts, 18, 159–166.

Posner, M., & Keele, S. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353–363.

Reber, R., Schwartz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? Personality and Social Psychology Review, 8, 364–382.

Rentfrow, P., & Gosling, S. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. Journal of Personality and Social Psychology, 84(6), 11236–1256.

Rentfrow, P., & Gosling, S. (2006). Message in a bottle: The role of music preferences in interpersonal perception. Psychological Science, 17(3), 236–242.

Repp, B. (1997). The aesthetic quality of a quantitatively average music performance: Two preliminary experiments. Music Perception, 14, 419–444.

Sloboda, J., & Juslin, P. (2001). Psychological perspectives on music and emotion. In P. Juslin & J. Sloboda (Eds.), Music and emotion: Theory and research, (pp. 71–104). Oxford: Oxford University Press.

Smith, J., & Melara, R. (1990). Aesthetic preference and syntactic prototypicality in music: 'Tis the gift to be simple. Cognition, 34, 279–298.

Sony Creative Software. (2008). Sound Forge 8. Madison, WI. Sony Creative Software Inc.

Whitfield, T. (1983). Predicting preference for familiar, everyday objects: An experimental confrontation between two theories of aesthetic behaviour. Journal of Experimental Psychology, 3, 221–237.

Whitfield, T. (2000). Beyond prototypicality: Toward a categorical-motivation model of aesthetics. Empirical Studies of the Arts, 18(1), 1–11.

Whitfield, T. (2009). Theory confrontation: Testing the categorical-motivation model. Empirical Studies of the Arts, 27(1), 43–59.

Whitfield, T., & Slatter, P. (1979). The effect of categorization and prototypicality on aesthetic choice in a furniture selection task. British Journal of Psychology, 70, 65–75.

Yeomans, J., & Frankland, P. (1996). The acoustic startle reflex: Neurons and connections. Brain Research Reviews, 21(3), 301–314.