PLINK: "THIN SLICES" OF MUSIC

CAROL L. KRUMHANSL
Cornell University

SHORT CLIPS (300 AND 400 MS), TAKEN FROM POPULAR songs from the 1960's through the 2000's, were presented to participants in two experiments to study the detail and contents of musical memory. For 400 ms clips, participants identified both artist and title on more than 25% of the trials. Very accurate confidence ratings showed that this knowledge was recalled consciously. Performance was somewhat higher if the clip contained a word or part-word from the title. Even when a clip was not identified, it conveyed information about emotional content, style and, to some extent, decade of release. Performance on song identification was markedly lower for 300 ms clips, although participants still gave consistent emotion and style judgments, and fairly accurate judgments of decade of release. The decade of release had no effect on identification, emotion consistency, or style consistency. However, older songs were preferred, suggesting that the availability of recorded music alters the pattern of preferences previously assumed to be established during adolescence and early adulthood. Taken together, the results point to extraordinary abilities to identify music based on highly reduced information.

Received August 3, 2009, accepted January 3, 2010.

Key words: music memory, meta-memory, popular music, emotion, style

THE CAR RADIO SCANS FROM STATION TO STATION and you recognize a song from decades ago. The title, artist, and lyrics flood into consciousness; perhaps the album cover, the appearance of the artist, or a personal anecdote. Or, if you don't recognize the song, you might immediately recall its era, musical genre, emotional content, how you danced to it, or its social significance. How detailed are these memories? What attributes of songs can be recalled? How is musical memory organized and which attributes are recalled even if the song is not recognized? And, what is the relationship between musical preferences and other aspects of musical memory? These questions were addressed in two experiments by presenting listeners with short (300 and 400 ms) clips from popular songs. In order to study effects of release date, the songs were taken from the 1960's through the 2000's. Listeners were asked to name the artist and title and indicate their confidence in the identifications. They were also asked about the decade of release, and judged the emotional content and style of the songs. After this, listeners were presented with long (15 s) clips for ratings of recognition and liking. They also judged these long clips for emotional content and style, which can be compared with those responses for the short clips as a way of assessing emotion and style consistency.

These questions about musical memory seem obvious, but surprisingly little research has been done on them despite the remarkable expansion of psychological research on music in the last few decades. Perhaps this is because the primary emphasis has been on the cognition of more abstract properties of musical structure (particularly scale, harmony, and meter), with many experiments using materials composed for that purpose and/or presented in timbre-neutral sounds (such as piano). As an example, take a typical study on whether listeners have knowledge of the structure of the major scale. Melodies would be composed for the experiment that conform to the scale and the way tones are typically sequenced in a major key. On a trial one melody would be presented, and then a second melody with one or more tones changed to tones outside the scale. The task is to detect the changed tone(s) and if it is done correctly (as it would be for both musicians and nonmusicians) it is concluded that listeners know the structure of the scale. The findings of a large number of studies demonstrate that these abstract descriptors of musical structure are useful not only in the technical analysis of music, but that they also function as cognitive frameworks for perceiving, remembering, and performing music. Thus, they have cognitive reality. However, these studies do not bear on the long-term effects of sustained exposure to recorded music, which offers a unique opportunity to study human memory.

There were three starting points for this study. The first is the concept of "thin slices" made well known by Gladwell's (2005) *Blink*, from which the title derives. Gladwell describes interesting cases in which, despite very reduced information, experts are able to do such things as judge authenticity of art, predict double faults in tennis, and

Music Perception VOLUME 27, ISSUE 5, PP. 337–354, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2010.27.5.337

spot likely customers. These are not cases in which these "thin slice" judgments are better than judgments with more complete information, but cases in which experts have good intuitions, although the basis for them is largely unconsciousness. The name "thin slices," however, goes back to a review by Ambady & Rosenthal (1992; see also Ambady, Bernieri, & Richeson, 2000). They compiled results from 44 studies showing that short exposures of behavior led to fairly accurate assessments along such dimensions as trustworthiness, physician proficiency, teacher adequacy, existence of deception, and supervisor ratings of therapists. Many other studies have been conducted since seeking to determine what kinds of social judgments can be made and on what basis (see, for example, Kraus & Keltner, 2009; Pentland, 2008). The question here is what kinds of information can be derived from "thin slices" of music (very short clips), particularly when the clips go unrecognized and the responses are based purely on impressions.

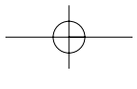
The second starting point was a series of well-known studies on the capacity of visual memory. Nickerson (1965) presented a set of 200 black and white photographs subsequent to which pairs of photographs were presented for forced-choice old/new judgments. Shepard's (1967) study used 612 colored pictures of objects. Standing, Conezio, and Haber (1970) showed 2560 photographs, and Standing (1973) showed 10,000 pictures. Recognition accuracy was high in all these studies, even when the retention interval was as long as a few days and the presentation time was as short as one minute. Recognition accuracy depended somewhat on these factors, but under most conditions accuracy was 90% or higher. Thus, visual recognition memory seems to be virtually limitless. Visual memory also contains a great amount of detail. Brady and colleagues (2008) presented 2,500 pictures of objects, which took 5.5 hrs. The test pairs were highly similar (e.g., bureau with drawer open or closed, telephone in different orientations). The forced-choice old/new judgments were still highly accurate (approximately 87%). The present study examines the detail of musical memory by determining if very short clips of music can be identified. This requires that listeners match the clips to a quite detailed representation of the sounds of the songs in memory and then retrieve the artist and title. From the data it is also possible to get a very rough estimate of the number of songs that are as familiar as those used in the experiment.

The most influential starting point, however, was Gjerdingen and Perrott's (2008) long-awaited article on genre classification of very short excerpts. (It was first presented at the 1999 meeting of the Society for Music Perception and Cognition, SMPC.) Gjerdingen and Perrott downloaded excerpts of songs in 10 genres of music (as categorized by the music business), with eight songs per genre. The original segments were 3000 ms in duration, and eight short segments of durations 250, 325, 400, and 475 ms were extracted from these; half were vocal and half purely instrumental. College undergraduates were asked to judge each clip as to which of ten industry-defined genres the song belonged. They also judged the genre of the 3000 ms clips. They argued that because genre classification is highly individual, the judgments of the short segments should be compared to the listener's own judgments of the long segments. However, their judgments could also be compared with the music industry's classification.

The results were striking. The main measure they presented was the percent of time a listener's genre judgments for the short segments agreed with their own judgments for the long segments. The percent of self-agreement was over 40% for the 250 ms segments and increased to over 50% for the 475 ms segments. (The judgments also corresponded to the industry's classification 75% of the time, so if they had used the industry classifications for comparison the results would still be impressive.) Self-agreement varied considerably across styles from above 70% for classical music to around 20% for blues. As they note, this may not accurately represent how recognizable the different genres are, but may reflect the categories used. For example, classical may be more distinguishable than the other nine categories (blues, country, dance, jazz, Latin, pop, R&B, rap, and rock) which, as popular music genres, would be expected to overlap more with one another. In addition, some classifications might include others (e.g., dance might include rock and Latin).

What enables listeners to classify these very short segments? They point to timbre, which they define as "the spectral and rapid time-domain variability in the acoustic signal," as being one likely source of information. They note that these short segments might also include very restricted information about melodic, bass, harmonic, and rhythmic features. They speculate that together some combination of this information may be sufficient to convey information about genre. The percent of self-agreement for instrumental music was higher than vocal music, which is somewhat surprising given that vocal qualities would intuitively be quite distinctive. They suggest one reason for this finding might be that vocal qualities may in fact be less informative about style per se, or recording engineers' emphasis on the vocals may mask the instrumentals.

The Gjerdingen and Perrott (2008) study also raised two theoretical issues that are pertinent to the present



experiments. The first is the nature of genre. They emphasize that listeners are not likely to agree with one another and conclude that genre classification is, in the final analysis, subjective. It would depend, for example, on the social categories associated with the musical categories. Moreover, the classification of musical genres would depend on age, gender, and other demographic factors. Boundaries between genres would be expected to shift over time as different genres are innovated or emerge from older genres. For these reasons, they used self-agreement as their main measure of classification consistency.

To mitigate some of these problems associated with genre boundaries, the present study used broader musical categories, which will be called styles. The style categories were based on a study by Rentfrow and Gosling (2003), who collected music preferences from over 3,500 individuals. The question addressed by their method (factor analysis) was if an individual liked one style of music, what other styles would they like? This method yielded four categories: (1) reflective/complex (classical, jazz, blues, folk), (2) intense/aggressive (alternative, rock, heavy metal), (3) upbeat/conventional (country, pop, religious, soundtracks), and (4) energetic/rhythmic (rap/hip-hop, soul/funk, electronica/dance). These classifications also may be useful because they are based on behavioral data that may predict listening and consumer choices better than industry-defined genres that are motivated by marketing considerations. Also, these preferences are associated with stable personality traits (Rentfrow & Gosling, 2006). For example, Rentfrow and Gosling used the Big Five Inventory (BFI; John & Srivastava, 1999) to measure personality traits. Openness was associated with preference for both reflective/complex and intense/rebellious music. Extraversion and agreeableness were associated with preference for upbeat/conventional and energetic/rhythmic music.

The second theoretical issue that was raised in the Gjerdingen and Perrott (2008) paper is the idea of a *plasticity period* (sometimes called the *reminiscence bump*), during which lifelong musical preferences are established. This plasticity period generally is located during the period from puberty through courtship and early adulthood. The obvious example is the "baby-boomers" attachment to 1960's era music (of course, it may be better music—a possibility that will receive some serious consideration later). An older study (Holbrook & Schindler, 1989) finds good support for this idea. Holbrook and Schindler used recordings of popular music ranging from 1932 to 1986 and listeners up to the age of 85. They calculated how old each participant was at the date of each song's release. Against this was plotted the rated musical preference. The results followed a very clear inverted U-shaped pattern with a peak at about the age of 24 years, with lower preferences for songs released before or after that age.

Holbrook and Schindler (1989) suggested two classes of explanations for a plasticity period: *intrinsic* and *extrinsic.* That the plasticity period is believed to occur during puberty and early courtship raises the possibility that it is related to hormonal and other physiological and physical changes that occur during that age. This would be an example of an intrinsic determinant. Drawing on the example of birdsong, a number of authors (see Wallin, Merker, & Brown, 2000) have thus suggested that music plays a role in sexual selection because it is used as a device for courtship. Extrinsic determinants focus, instead, on the music experience of the listeners. Among explanations they offered for the apparent effects of adolescence is the increased role of music in social life, the time when political and other ideological commitments are crystallized, as well as the establishment of an "in-group" (North & Hargreaves, 1999). Most likely, if there is such a thing as a plasticity period, intrinsic and extrinsic determinants would both be at play. The present experiments address how musical preferences depend on release date relative to age by selecting songs over a five decade range, and considers the effect of release date on a variety of measures in addition to musical preference.

Bartlett and Snelus (1980) presented songs that were popular during the six decades from 1920 through 1970 to middle-aged (37–57) and "elderly" (60–76) participants. The songs were transcribed and played on the piano. The middle-aged participants did not recognize the songs from the decades before they were born and when they were young children. Neither group's data showed a peak for the songs popular during adolescence and early childhood as might be predicted from Holbrook and Schindler's (1989) results, but it should be remembered these were recognition judgments rather than preference judgments. Another finding was that the accuracy of participants' temporal judgments (estimated decade of popularity) required recognition of the songs from melody or title cues. The authors concluded that the temporal judgments depended on recognizing the particular song.

In contrast, Schulkind, Hennis, and Rubin (1999) found effects consistent with the inverted U-shaped curve. The younger subjects (18–21) would be on the increasing side of the function, whereas the older subjects (66–71) would be on the decreasing side, predicting interactions with age. These interactions were found on measures of familiarity, combined recall (artist, title, and cued-recall of lyrics), general memories of the decade of the song, and emotionality. An important difference from the Bartlett and Snelus (1980) study is that this study used excerpts

from recordings that would be expected to contain richer cues for recognition, particularly timbre (the "sound color" of the music).

The present experiments also asked listeners to judge the emotional content from the short clips. Because of the large number of possible emotion terms, participants were restricted to a small set of choices. Juslin (2000) showed that performers can successfully communicate five basic emotions: (1) happiness, (2) sadness, (3) tenderness, (4) fear, and (5) anger. That is, when performers were instructed to play music with these emotional intentions, listeners were able to decode them fairly accurately. Juslin's "lens model" (see also Brunswick, 1956) posits that each emotion is conveyed by a variety of acoustic cues, such as loudness, rate, and articulation, which in combination serve to communicate the intended emotion. Juslin and Laukka (2003) showed that similar cues are used in emotional speech. Owing to this research, these five emotion categories were offered as the choices in the experiments.

Two previous studies have found that the emotional content of music is conveyed by brief sections of classical music. Peretz, Gagnon, and Bouchard (1998) presented listeners with excerpts ranging from 0.5 to 7.5 s. Participants rated, on a 10-point scale, how happy or sad the excerpts were. Even at the shortest durations the happy and sad pieces were reliably discriminated, with only a modest increase out to 7.5 s. Another experiment showed the discrimination extended down to .25 s. An interesting feature of this study is that it compared the emotion judgments of the normal participants with those of I. R., an amusic following brain damage. Even though I. R. had severe difficulties recognizing songs, her judgments or happy versus sad were spared.

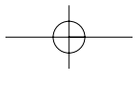
The second study (Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005) also used excerpts of classical music. Participants in two experiments were asked to group together 27 long excerpts according to whether they were similar in terms of their *own felt emotion.* (For discussion of how felt emotion is distinguished from emotion represented by the music, for example, rapid tempo representing happiness, see Krumhansl, 1997.) In their first experiment, the excerpts averaged 30 s and in the second experiment were 1 s in duration. Similarity judgments of pairs of the 1 s excerpts were collected in the third experiment. All three experiments found the same two dimensions using multidimensional scaling: valence and arousal, which have been found in many studies of musical emotion dating back to Hevner (1935a, 1935b). The third dimension was whether or not the excerpt began with melodic material or harmonic material (chords).

In addition to the emotional content and style judgments, listeners in the present experiments were asked to identify the artist, title, and release decade of the songs from which the short segments (300 and 400 ms) were drawn. Judgments of emotional content and style might be based on subtle and impressionistic cues of which the participants are likely unaware, but recall of artist and title calls for very explicit knowledge. A related question is whether participants have *meta-cognition* concerning these identifications; that is, did they know whether or not their responses were correct? For this, participants rated how confident they were of their identifications. To see whether identification accuracy depended on the location of the excerpt in the song, half the clips came from the chorus and half from another section of the song, the idea being that the chorus might be more "catchy" or contain a "hook." Hyman and Rubin's (1990, Rubin, 1995) study of Beatles lyrics found better recall when the line shared content words with the title (other factors were primacy and repetition). Consequently, each clip was coded as to whether it contained a word or part-word from the title.

To summarize, this article reports two experiments that presented short clips from popular songs with release dates ranging over the five decades from the 1960's to the present. In Experiment 1, the clips were 400 ms in length; in Experiment 2 they were 300 ms in length. For each clip, participants were asked to name artist and title, and indicate their confidence in the accuracy of their identifications. They also judged decade of release, emotional content (from among five choices), and style (from among four categories). After the main experiments, they also rated longer (15 s) segments for recognition and liking and (in Experiment 2) emotional content and style. The advantage of asking all these questions in combination, and determining which attributes degrade together for the shorter durations, was to be able to study the contingent relationships between these aspects of musical memory. One possibility, for example, is that knowing the artist and title is prerequisite to accurately judging release decade, emotional content, and style. Another possibility is that judgments of emotional content and style (and perhaps decade of release) might be made based on impressions from the short clips even when it is not identified.

Experiment 1

Preliminary testing found that 400 ms clips yielded a reasonable level of identification accuracy. For some clips, identification of both artist and title was correct and the recognition was rapid and certain. Other clips were



unrecognizable. Thus, for clips of this length it would be possible to look at the various response measures (release decade, emotional content, and style) as a function of whether or not the clip was correctly identified. This was done in order to explore the kinds of information that depend on correct recognition. To look at this in more detail, the data were then analyzed separately for those trials on which participants were sure they did not recognize the song ("1") and those trials when they correctly named both artist and title ("6"). The short (400 ms) clips will be referred to as SC400, trials on which participants were certain they did not recognize the clip will be referred to as SC400 "1" trials, and trials on which participants named both artist and title will be called SC400 "6" trials.

Method

PARTICIPANTS

Twenty-three university students participated in the experiment for course credit. Their average age was 20.5 years (range 18–23). Most had formal music instruction averaging 10.3 years (range 0–29, summing across all instruments and voice, e.g., 4 piano, 6 violin = 10 years total). They listened to popular music for an average of 21.7 hours per week (range 7–58). They reported listening to reflective/complex music (blues, jazz, classical, folk) an average of 5.7 hours per week, intense/aggressive music (rock, alternative, heavy metal, punk) 8.8 hours per week, upbeat/conventional music (pop, country, religious) 2.8 hours per week, and energetic/rhythmic music (rap, hip-hop, soul/funk, dance/electronica) 4.5 hours per week.

STIMULUS MATERIALS

A list of songs was compiled from *Rolling Stone*, Billboard, and Blender lists of top songs "of all time." To this list were added a number of more recent songs that would be familiar to the college-aged participants but would not be listed in these sources because they were released relatively recently. This resulted in a total of 52 songs.

Of the 52 songs, 28 were chosen for the experiment based on the availability of recordings; they are listed in Table 1. Three clips were made from each of these songs using Audacity software: (1) a 400 ms short clip (SC400) from the chorus, (2) a 400 ms short clip (SC400) from another part of the song, and (3) a 15 s long clip (LC) that was used to test recognition and liking of the songs after the main experiment. It was noted which short clips contained a word or part-word from the title of the song.

PROCEDURE

The 56 short clips (chorus and non-chorus for each of 28 songs) were randomly ordered and each was presented three times with approximately 2 s between presentations. Participants answered the following questions for each short clip:

(1) Artist
(2) Title
(3) Confidence in identification ("1" = "don't recognize at all," "2" = "pretty sure don't recognize," "3" = "might possibly recognize," "4" = "recognize but couldn't name either artist or title," "5" = "can name either artist or title," "6" = "can name both artist and title")
(4) Decade the song was released (60's, 70's, 80's, 90's, 00's)
(5) Emotion conveyed by the song ("1" = "happiness," "2" = "sadness," "3" = "anger," "4" = "fear," "5" = "tenderness")
(6) Style ("1" = "reflective/complex" (blues, jazz, classical, folk), "2" = "intense/aggressive" (rock, alternative, heavy metal, punk), "3" = "upbeat/conventional" (pop, country, religious), "4" = "energetic/rhythmic" (rap, hip-hop, soul/funk, dance/electronica))

Following this session, the 15 s long clips (LC) of the 28 songs were presented in a random order. The participants answered two questions: Do you recognize it? Do you like it? Both answers were made on a scale of 1–6 (1 = "least" and 6 = "most").

Lastly, participants filled out a questionnaire about their music background: the extent of their musical training, whether they are currently performing, and their listening habits. In addition, they gave some demographic information: age, gender, age of parent, number and relative age of siblings, as well as indicating the kind of music they heard in the home environment.

Results and Discussion

IDENTIFICATION OF 400 MS SHORT CLIPS (SC400)

Initial inspection of identification confidence ratings for the 400 ms short clips (SC400) showed that most were either "1" ("don't recognize at all") or "6" ("can name both artist and title"). It was possible to check the accuracy of the "6" responses because the participants were also asked to identify the artist and title. Of those who said "6," 95.2% correctly named both, 1.4% correctly named only artist or title, and 3.4% were incorrect on both artist and song title. Of the trials on which participants correctly named both artist and song title, 96.6% said "6" and the remaining 3.4% underestimated their accuracy. In other words,

TABLE 1. Table of Songs Ordered by 400 ms Short Clip (SC400) Identification.

Artist	Title	SC400 Recog	Chorus	Title Word	Actual Release	Judged Release	LC Recog	LC Liking
Britney Spears	Baby One More Time	5.6	C	Y	1998	1997.2	6.0	2.4
Aretha Franklin	Respect	5.6	C	Y	1967	1973.7	5.6	4.8
Britney Spears	Baby One More Time	5.5	NC	N	1998	1996.3	6.0	2.5
Red Hot Chili Pepper	Californication	5.3	C	Y	1999	1993.3	5.7	4.7
The Police	Every Breath You Take	5.2	C	N	1983	1985.4	5.6	4.2
Michael Jackson	Thriller	4.9	C	N	1982	1988.0	5.8	4.4
The Beatles	Help	4.9	C	Y	1965	1970.2	5.6	4.7
Nirvana	Smells Like Teen Spirit	4.7	NC	N	1991	1992.8	5.7	4.0
The Police	Every Breath You Take	4.6	NC	N	1983	1984.1	5.7	4.3
Eagles	Hotel California	4.3	NC	N	1976	1979.3	5.1	4.6
The Beatles	Imagine	4.2	NC	N	1971	1975.4	4.6	5.3
Journey	Don't Stop Believing'	4.1	C	N	1981	1984.6	5.0	4.0
Jimi Hendrix	Purple Haze	4.1	NC	N	1967	1973.7	5.7	4.5
Outcast	Hey Ya!	4.0	C	Y	2007	1997.2	5.5	4.3
The Ramones	Blitzkrieg Bop	3.9	NC	N	1976	1983.3	4.5	3.7
Coldplay	Viva La Vida	3.8	NC	N	2008	1995.0	5.4	4.0
Led Zeppelin	Stairway to Heaven	3.7	C	Y	1971	1979.8	4.9	5.1
Bob Dylan	Mr. Tambourine Man	3.5	C	Y	1965	1976.8	4.6	4.4
Aretha Franklin	Respect	3.2	NC	N	1967	1975.0	5.6	4.8
The Beatles	Imagine	3.2	C	Y	1971	1978.0	4.6	5.3
Jimi Hendrix	Purple Haze	3.1	C	N	1967	1978.0	5.7	4.5
Madonna	Like a Virgin	3.0	C	Y	1984	1983.3	5.0	3.4
Madonna	Like a Virgin	2.9	NC	N	1984	1989.8	5.1	3.3
Louis Armstrong	What a Wonderful World	2.8	C	Y	1968	1976.7	5.4	4.7
Rolling Stones	Satisfaction	2.8	C	Y	1965	1978.5	5.6	4.7
Rolling Stones	Satisfaction	2.8	NC	N	1965	1977.2	5.6	4.6
The Clash	London Calling	2.8	C	N	1979	1986.7	3.5	3.4
Will Smith	Gettin' Jiggy With It	2.6	NC	N	1997	1996.3	5.7	3.7
The Ramones	Blitzkrieg Bop	2.4	C	Y	1976	1988.0	4.6	3.7
Amy Winehouse	Rehab	2.4	NC	N	2006	1996.7	4.9	3.5
Bob Marley	No Woman No Cry	2.3	NC	N	1974	1985.4	4.8	4.3
Outcast	Hey Ya!	2.2	NC	N	2007	1999.8	5.5	4.3
Eagles	Hotel California	2.2	C	Y	1976	1987.2	5.4	4.6
Bob Marley	No Woman No Cry	2.2	C	N	1974	1977.6	4.9	4.5
Simon & Garfunkle	Bridge Over Troubled Water	2.1	C	N	1970	1982.8	4.8	3.9
Led Zeppelin	Stairway to Heaven	2.1	NC	N	1971	1976.3	4.9	5.1
Red Hot Chili Pepper	Californication	2.0	NC	N	1999	1988.5	5.7	4.8
Will Smith	Gettin' Jiggy With It	2.0	C	N	1997	1994.1	5.7	3.5
Amy Winehouse	Rehab	2.0	C	N	2006	1995.4	4.9	3.6
Journey	Don't Stop Believing'	2.0	NC	N	1981	1989.3	5.1	3.9
Katy Perry	I Kissed a Girl	2.0	C	N	2008	1993.7	5.0	2.4
The Beatles	Help	1.9	NC	N	1965	1981.1	5.5	4.6
Simon & Garfunkle	Bridge Over Troubled Water	1.9	NC	N	1970	1985.4	4.9	4.0
Queen	Bohemian Rhapsody	1.8	NC	N	1975	1985.9	5.6	5.3
Nirvana	Smells Like Teen Spirit	1.7	C	N	1991	1993.3	5.8	4.1
Katy Perry	I Kissed a Girl	1.7	NC	N	2008	1995.4	5.0	2.6
Bob Dylan	Mr. Tambourine Man	1.7	NC	N	1965	1977.6	4.5	4.3
Guns N' Roses	Sweet Child O' Mine	1.7	NC	N	1987	1990.7	5.0	4.2
Queen	Bohemian Rhapsody	1.7	C	N	1975	1986.3	5.7	5.3
Louis Armstrong	What a Wonderful World	1.6	NC	N	1968	1986.7	5.0	4.7
Guns N' Roses	Sweet Child O' Mine	1.6	C	N	1987	1988.0	5.1	4.3
Coldplay	Viva La Vida	1.6	C	N	2008	1990.2	5.4	3.9
Michael Jackson	Thriller	1.5	NC	N	1982	1991.1	5.7	4.3
The Clash	London Calling	1.5	NC	N	1979	1983.7	3.6	3.3
U2	Beautiful Day	1.4	C	N	2000	1989.8	5.4	3.8
U2	Beautiful Day	1.4	NC	N	2000	1991.1	5.6	4.0

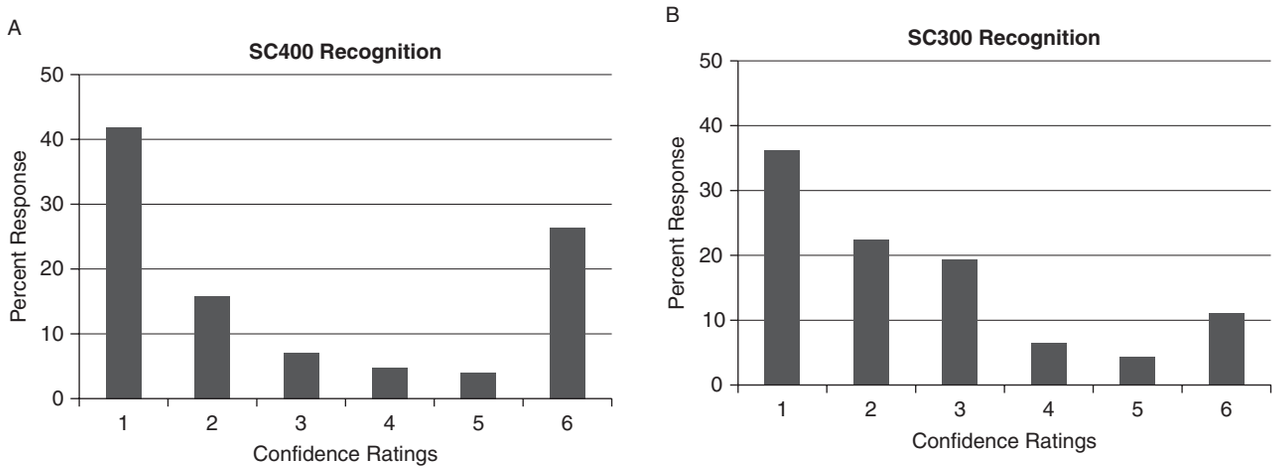


FIGURE 1. (A) Distribution of confidence ratings about identification ("1" = "don't recognize at all"; "6" = "can name both artist and title") for 400 ms short clips (SC400). (B) Distribution of confidence ratings for 300 ms short clips (SC300).

the confidence ratings most often corresponded with their artist and title identifications. On the few trials (4%) when participant identified only artist or title, 84% of the time they identified only the artist and 16% of the time they identified only the title.

The confidence scores were adjusted according to artist and title responses, so the "5" and "6" scores were true measures of identification accuracy. Figure 1a shows the strongly bimodal distribution of confidence ratings adjusted in this way. Performance was strong, with correct identifications 26.4% of the time. Table 1 lists the clips in order of SC400 identification confidence.

Chorus clips were identified slightly more often than non-chorus clips (31.1% vs. 21.7% "6"responses). The average confidence scores were 3.12 and 2.73, respectively, but the difference was not significant, $F(1, 54) = 1.32$, $p = .256$. Clips that contained a word or part-word from the title, however, were better recognized than clips that did not (44.1% vs. 21.0% "6" responses). The average confidence scores were 3.77 and 2.67, respectively, and the difference was significant, $F(1, 54) = 8.38$, $p = .006$. SC400 identification confidence did not correlate with the release date of the song, $r(54) = -.07$, $p = .59$.

JUDGED DECADE OF RELEASE

Judged release dates were computed by transforming the participants' judged decade into the midpoint of that decade (i.e., 1980's was coded as 1985), and then calculating the average for each song. The correlation between actual release date and judged release date for all clips was highly significant, $r(54) = .89$, $p < .0001$.

The correlations also were computed separately for trials with "1" ("don't recognize at all") and "6" ("can name both artist and title") confidence ratings. For SC400 "1" trials, the correlation between judged and actual release date was quite strong, $r(54) = .59$, $p < .0001$, as shown in Figure 2a. This was so even though they were certain they did not recognize the song. As would be expected, the SC400 "6" trials correlated more strongly with actual release date, $r(54) = .98$, $p < .0001$, as shown in Figure 2b (Figure 2c shows the data for Experiment 2 with 300 ms clips.) For "6" trials, the accuracy of the judged release dates is notable, as well as the degree to which the slope of the regression line approximates true accuracy.

JUDGED EMOTIONAL CONTENT

For each clip, the total number of participants choosing each of the five emotions was tabulated and then converted into percentiles. For example, suppose for one clip that the number of participants choosing each emotion was distributed: happiness = 0, sadness = 16, anger = 0, fear = 0, tenderness = 8. Then the profile would be 0, .67, 0, 0, .33. These profiles were then correlated with the emotion profiles of the LC emotion judgments (from the participants in Experiment 2), computed in the same way. The correlations will be called a measure of *emotion consistency*. The correlations were significantly different from 0, $t(55) = 8.84$, $p < .0001$. The distribution of correlations was strongly skewed to the left, so a Wilcoxon signed-rank test also was computed, rendering a value of $W = 696$, $p < .0001$.

The correlations also were computed separately for SC400 "1" and SC400 "6" trials. The distributions are shown in Figure 3a and 3c. They were both significantly different from 0, $t(55) = 5.40$, $p < .0001$, Wilcoxon

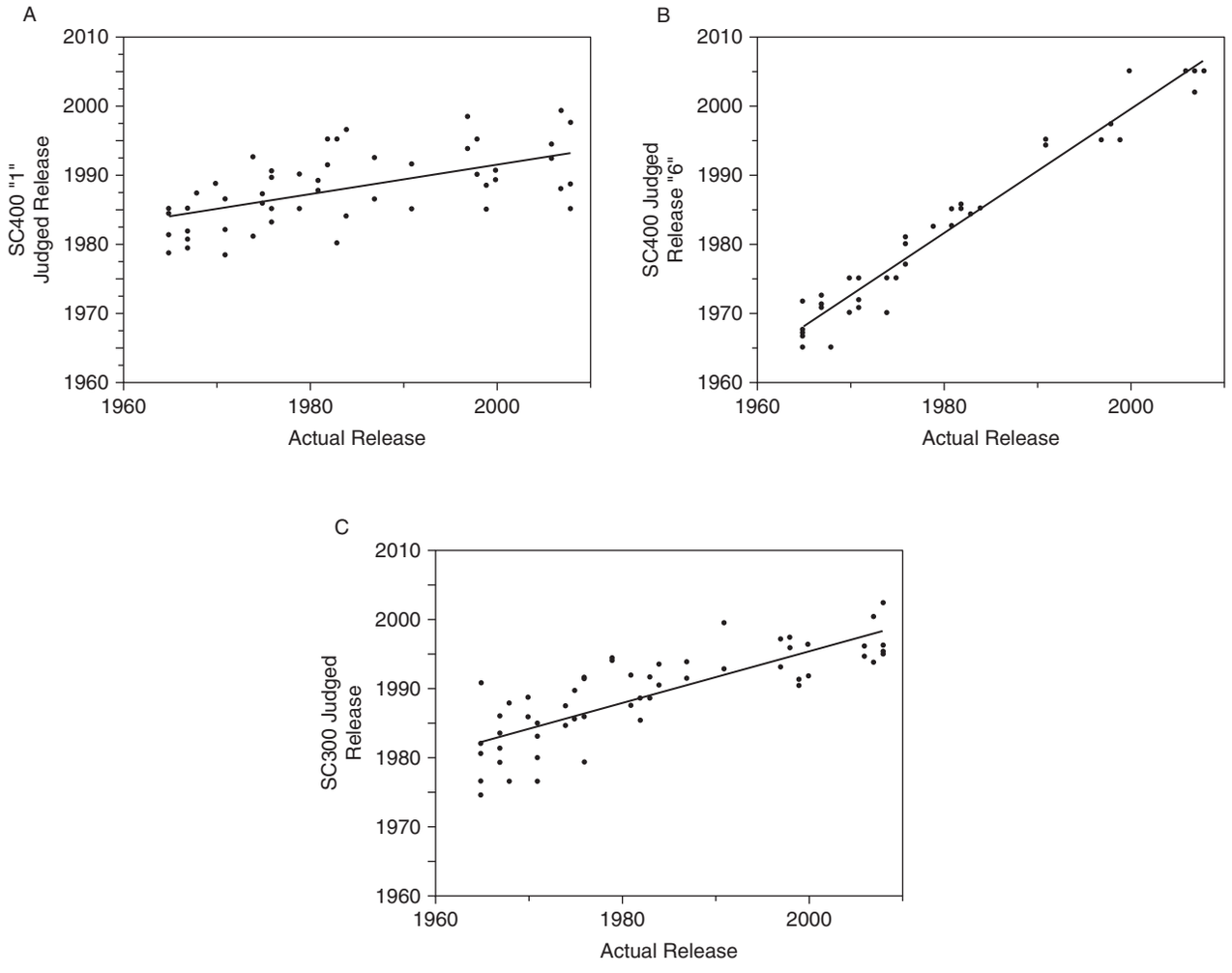


FIGURE 2. (A) Correlation between judged release date and actual release date on SC400 "1" ("don't recognize at all") trials. (B) Correlation between judged release date and actual release date on SC400 "6" ("can name both artist and title"). (C) Correlation between judged release date and actual release date on SC300 trials.

signed-rank = 525, $p < .0001$; $t(49) = 8.70$, $p < .0001$, Wilcoxon signed-rank = 565, $p < .0001$, respectively. The two kinds of trials (SC400 "1" and SC400 "6") differed significantly from one another, $t(49) = 2.56$, $p = .014$, although the difference was not as large as one might have expected given the contrast between the extremes of identification. Emotion consistency did not correlate with actual release date for the SC400 "1" trials, but did correlate for the SC400 "6" trials, $r(48) = .36$, $p = .009$), with more consistency for recent songs.

The choice of sadness ratings decreased with actual release date, $r(54) = -.36$, $p = .006$) as did tenderness, $r(54) = -.38$, $p = .004$. In contrast, anger increased with actual release date, $r(54) = .41$, $p = .002$. Neither of the other two emotions correlated with actual release date.

JUDGED STYLE

For style judgments, the total number of participants choosing each of the four styles was tabulated, converted into percents to make the style profile, and correlated with the LC style profiles. The correlations will be taken as a measure of *style consistency*. The correlations were significantly different from 0, $t(55) = 8.84$, $p < .0001$, Wilcoxon signed-rank = 750, $p < .0001$.

The correlations also were computed separately for SC400 "1" and SC400 "6" trials. The distributions of the correlations are shown in Figure 4a and 4c. They were both significantly different from 0, $t(55) = 7.28$, $p < .0001$, Wilcoxon signed-rank = 637, $p < .0001$; $t(49) = 8.70$, $p < .0001$; Wilcoxon signed-rank = 593, $p < .0001$, respectively. The two kinds of trials (SC400 "1" and

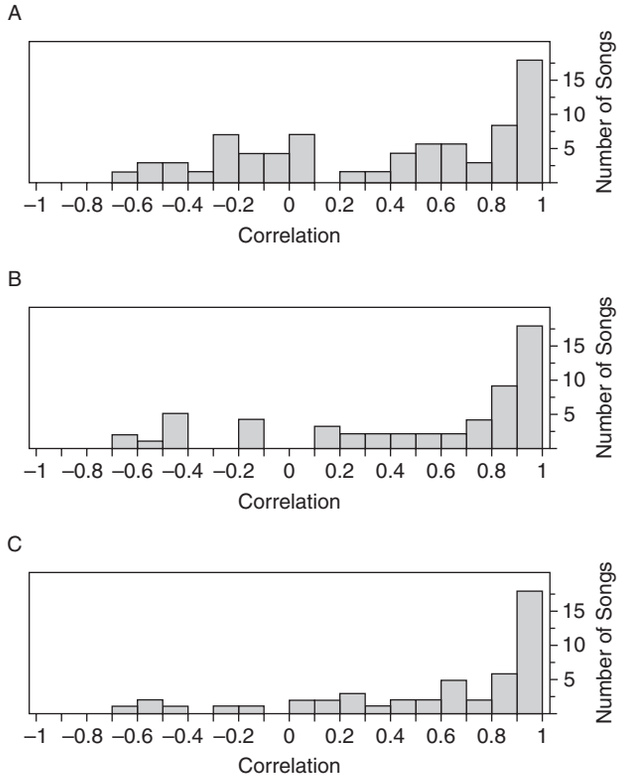


FIGURE 3. (A) Distribution of correlations between emotion profiles on SC400 "1" ("don't recognize at all") trials and LC (15 s). (B) Distribution of correlations between emotion profiles on SC300 trials and LC trials. (C) Distribution of correlations between emotion profiles on SC400 "6" ("can name both artist and title") trials and LC trials.

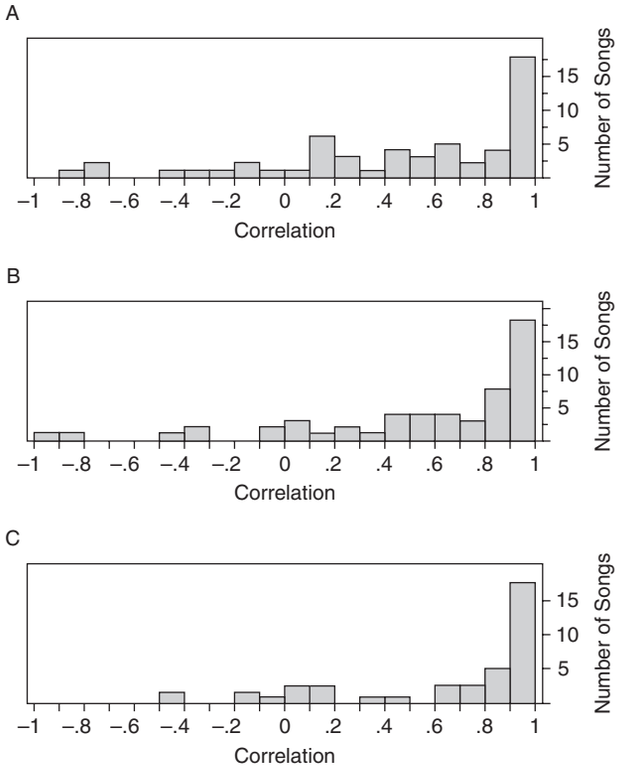


FIGURE 4. (A) Distribution of correlations between style profiles on SC400 "1" ("don't recognize at all") trials and LC (15 s) trials. (B) Distribution of correlations between style profiles on SC300 trials and LC trials. (C) Distribution of correlations between style profiles on SC400 "6" ("can name both artist and title") trials and LC trials.

SC400 "6") also were significantly different from one another, $t(49) = 2.46$, $p = .017$, but again the modest size of this effect should be noted. Style consistency did not correlate with release date for either SC400 "1" or SC400 "6" trials, although there was an increasing trend for the former, $r(54) = .23$, $p = .064$.

The choice of reflective/complex styles (blues, jazz, classical, and folk) decreased with the actual release date, $r(54) = -.53$, $p < .0001$. In contrast, energetic/rhythmic styles (rap, hip-hop, soul/funk, dance, electronica) increased with actual release date, $r(54) = .37$, $p = .005$. Neither of the other two styles correlated with actual release date.

To summarize briefly, the SC400 clips were identified with both artist and title more than 25% of the time and, when identified, the judged release date was remarkably accurate. Participants' confidence in their identifications closely matched their true accuracy. The judged release date correlated significantly with actual release date even without identification, although the correlation was weaker. This suggests that release date is linked to artist and title information (or can be inferred from them), but that there are also impressions of release date in the absence of identification. Similarly, participants apparently were able to get impressions of emotional content and style even when the short clips were not identified.

Experiment 2

The primary manipulation in Experiment 2 was reducing the duration of the clips from 400 ms to 300 ms. This provides a test of the replicability of the effects found in Experiment 1. In addition, if the level of accuracy on clip identification is lower, as would be expected, then the results can be compared with the trials in Experiment 1 on which the participant was sure they did not recognize the clip (SC400 "1") and the trials on which the participant named both artist and title correctly (SC400 "6"). Intermediate results would be expected. Another reason was to see which attributes of the clips degraded together

in order to explore which appear to be associated in memory either by direct links or inference. Another motivation for Experiment 2 was to obtain emotion and style judgments for the long clips that can be compared with those judgments for the short clips. (These data have already been used in the analysis of Experiment 1 results.) Finally, the questionnaires now included questions about the participants' current musical influences. The short (300 ms) clips will be referred to as SC300.

Method

PARTICIPANTS

Thirty-six university students participated in the experiment for course credit. Their average age was 21.1 years (range 19–25). They had formal music instruction for an average of 10.0 years (range 0–32, summing across all instruments and voice). They currently listened to popular music for an average of 19.8 hours per week (range 1–70). They reported listening to reflective/complex music an average of 5.7 hours per week, intense/aggressive music 6.4 hours per week, upbeat/conventional music 4.0 hours per week, and energetic/rhythmic music 3.9 hours per week.

STIMULUS MATERIALS

The clips used in Experiment 1 were edited with Audacity software so that they were 300 ms in duration.

PROCEDURE

The procedure was the same as Experiment 1 except that they also judged the emotion and style of the long (15 s) clips (LC). The questionnaire, in addition to asking about music background, asked participants to indicate the strength of their musical influences (on a scale from 1 to 6): music of peers, music on radio/web/etc., music from lessons, music of parents, and music of siblings.

Results and Discussion

IDENTIFICATION OF 300 MS SHORT CLIPS (SC300)

As was true for Experiment 1, SC300 confidence ratings were bimodal, although the pattern was less clear than in Experiment 1. That the small increase at "6" is really a mode is reinforced by the accurate assessment of confidence associated with the "6" ratings. Of those who said "6", 90.3% correctly named both, 1.8% correctly named only artist or title, and 8.2% were incorrect on both artist and song title. Of the trials on which participants correctly named both artist and song title, 93.2% said "6" and the remaining 6.8% underestimated their accuracy. In other words, participants' confidence judgments were quite consistent with their accuracy. On the few trials (2.3%) when participants identified only artist or title, 66.7% of the time they identified only the artist and 33.3% of the time they identified only the title. The confidence scores were adjusted according to artist and title accuracy, resulting in the distribution of confidence ratings shown in Figure 1b. Performance was considerably worse than in Experiment 1 with only 11.1% of trials correct on both artist and title.

No effect was found of whether the clip came form the chorus, $F(1, 54) = 0.02$, *n.s.*, or whether the clip contained a word or part-word of the title, $F(1, 54) = 0.001$, *n.s.* SC300 identification confidence did not correlate with the release date of the song, $r(54) = .13$, *n.s.*

JUDGED DECADE OF RELEASE

Judged decade of release judgments were preprocessed as in Experiment 1. The correlation between the judged release date and the actual release date for all clips was highly significant, $r(54) = .81$, $p < .0001$, as can be seen in Figure 2b. It was stronger than that for SC400 "1" trials and weaker than that for SC400 "6" trials.

JUDGED EMOTIONAL CONTENT

As in Experiment 1, emotion profiles were found for each SC300 and then correlated with the emotion profiles for the LC judgments. Fig ure 3b shows the distribution of correlations. They were significantly different from 0, $t(55) = 6.07$, $p < .0001$, Wilcoxon signed-rank = 638, $p < .0001$. These correlations were only marginally different from SC400 "1" trials, $t(55) = 1.71$, $p = .047$, and not significantly different from SC400 "6" trials, $t(49) = 1.16$, $p = .254$. Emotion consistency did not depend on actual release date.

As in Experiment 1, the choice of sad and tenderness decreased with actual release date, $r(54) = -.39$, $p = .032$ and $r(54) = -.44$, $p = .001$, respectively, the choice of anger increased with actual release date, $r(54) = .41$, $p = .002$; the correlations with the other two emotions were not significant.

JUDGED STYLE

Style profiles were found for each SC300 and correlated with the style profiles from the LC judgments of the four styles. Figure 4b shows the distribution of correlations. They were significantly different from 0, $t(55) = 8.90$, $p < .0001$, Wilcoxon signed-rank = 687, $p < .0001$. They were not significantly different from SC400 "1," $t(55) = 1.41$, $p = .165$, or SC400 "6," $t(49) = 1.44$, $p = .156$. Style consistency did not depend on actual release date.

The choice of reflective/complex decreased with actual release date, $r(54) = -.56$, $p < .0001$, and energetic/rhythmic increased with actual release date, $r(54) = .31$, $p = .02$. The correlations with the other two styles were not significant.

Comparing Results for Experiments 1 and 2

INTERIM SUMMARY

The most noticeable difference of reducing the duration of the short clips from 400 to 300 ms was the marked decrease in the identification of artist and title. An effect also was found on the accuracy of decade of release, although the judgments for the SC300 clips were still strongly correlated with the actual release. The duration of the excerpts had rather little effect on the consistency of emotional content or style judgments.

Many of the same results were found in Experiments 1 and 2. Participants' ratings of the accuracy of their artist and title identifications were bimodal. Their confidence ratings most often corresponded with the actual accuracy of their artist and title identifications. When only the artist or title was identified, it was most often the artist. For both experiments, the judged and actual release date correlated very significantly even when the clip was not identified.

Across experiments, the emotional content judgments of the short clips were quite consistent with those of the long clips. In addition, the level of consistency was similar for SC400 "1", SC300, and SC400 "6" trials; consistency on SC300 trials was intermediate and only marginally higher than SC400 "1" trials and not significantly lower than SC400 "6" trials. Both experiments found intuitively correct changes in judgments of emotional content over the decades; judgments of sad and tenderness decreased with actual decade of release and judgments of anger increased.

The parallel results were found for the style judgments. Both experiments found quite consistent correlations between the style judgments of the short clips and the long clips. The level of consistency was similar for SC400 "1," SC300, and SC400 "6" trials; consistency on SC300 was intermediate and was not significantly different from either SC400 "1" trials or SC400 "6" trials. Both experiments found intuitively correct changes in style judgments over the decades; reflective/complex styles decreased with actual decade of release and energetic/rhythmic styles increased.

LONG CLIP RECOGNITION JUDGMENTS

In both experiments, listeners gave ratings of whether they recognized the long (15 s) clips (LC). These were highly consistent across Experiments 1 and 2, $r(54) = .71$, $p < .0001$, so in the following analyses the two sets of LC recognition judgments were combined. As can be seen in Figure 5a, the vast majority of responses were "6," meaning they could identify both artist and title. LC recognition correlated positively, although weakly, with decade of release, $r(54) = .27$, $p = .048$), as shown in Figure 6a.

LC recognition judgments were tested to see if they correlated with SC300 recognition, SC400 recognition, SC300 emotion consistency, SC400 emotion consistency, SC300 style consistency, and SC400 style consistency. None of these six measures correlated with LC recognition judgments (all $p > .07$, with the only marginally significant correlation being the correlation between SC400 recognition and LC recognition, as might be expected). The low correlations may not be surprising given the limited range (mostly "6") of the LC recognition responses.

LONG CLIP LIKING JUDGMENTS

In both experiments, listeners gave ratings of whether they liked the long (15 s) clips (LC). These were highly

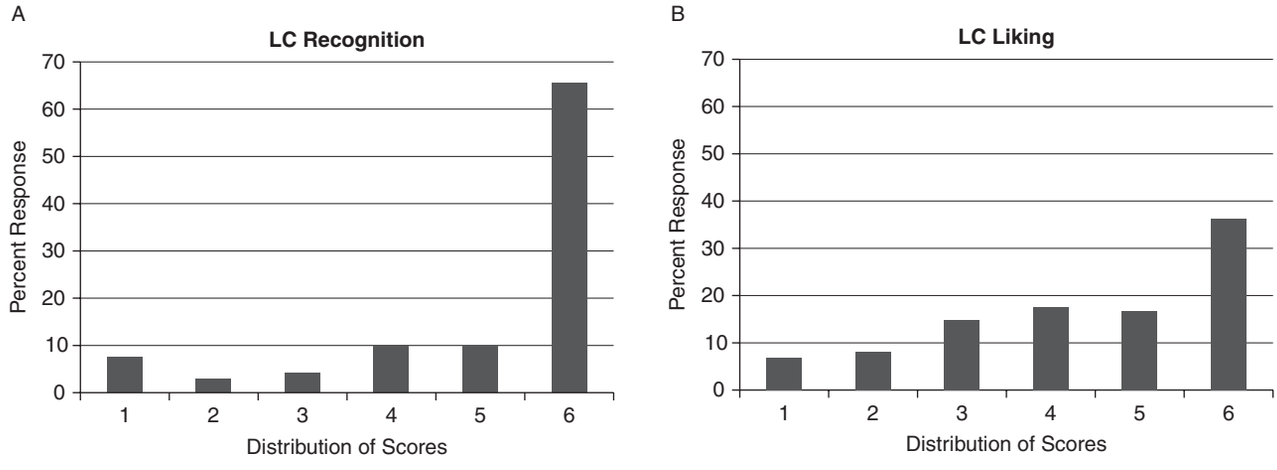


FIGURE 5. (A) Distribution of long clip (LC) recognition scores. (B) Distribution of long clip (LC) liking scores.

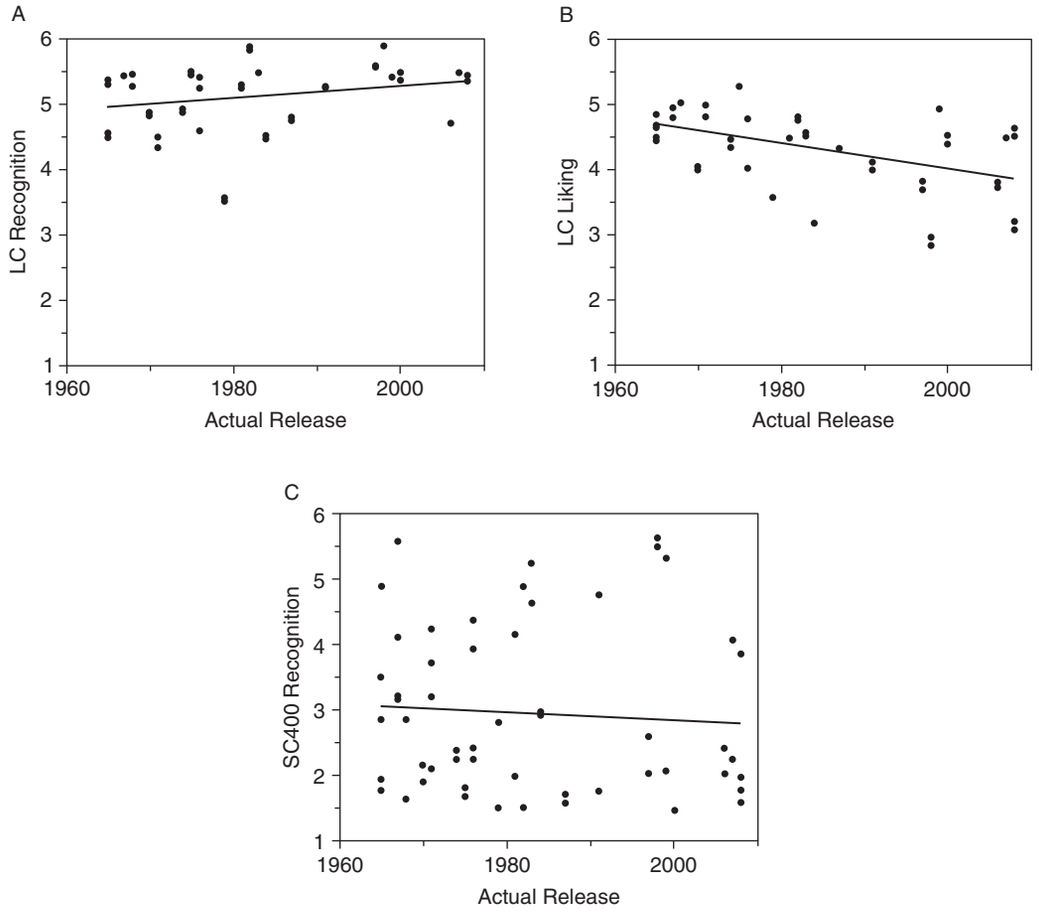


FIGURE 6. (A) Long clip recognition ratings as a function of actual release date. (B) Long clip liking ratings as a function of actual release date. (C) SC400 identification as a function of actual release date.

consistent across experiments, $r(54) = .72$, $p < .0001$, so in the following analyses the two sets of LC liking judgments were combined. The distribution of responses is shown in Figure 5b. LC liking correlated quite strongly and *negatively* with decade of release, $r(54) = -.57$, $p < .0001$, as shown in Figure 6b, so that older songs were quite consistently preferred to more recent songs.

LC liking judgments were tested to see if they correlated with SC300 recognition, SC400 recognition, SC300 emotion consistency, SC400 emotion consistency, SC300 style consistency, and SC400 style consistency. Of these only two were significant and *negatively* related to LC liking: SC300 emotion consistency, $r(54) = -.40$, $p = .002$, and SC400 emotion consistency, $r(54) = -.38$, $p = .004$. This means that songs with more ambiguous emotions were actually preferred. It might be noted that the same negative trend was found for SC300 and SC400 style consistency, $r(54) = -.23$, $p = .088$; $r(54) = -.22$, $p = .097$, respectively).

LC recognition and liking judgments were not correlated with one another, $r(54) = .21$, $p = .132$. As can be seen in Table 1, some of the most recognized songs (for example, *Baby One More Time*) received low rankings of liking.

DECADE OF RELEASE

Decade of release was not correlated with SC300 recognition, SC400 recognition, SC300 emotion consistency, SC400 emotion consistency, SC300 style consistency, or SC400 style consistency. Examining Table 1, which orders the clips in terms of SC400 recognition, suggests a preponderance of songs of the era 1965 through 1975 at the top of the list, specifically, *Respect, Help, Purple Haze,* and *Mr. Tambourine Man.* To investigate this further, the songs were divided into the top third, middle third, and bottom third in terms of SC400 Recognition. Of the songs released from 1965 through 1975, seven were in the top third, nine were in the middle third, and six were in the

bottom third. Thus, the songs released in during 1965–1975 had no special status in terms of being recognized from the short clips. Moreover, Figure 6c shows SC400 recognition as a function of release date, and it shows clearly that there is no relationship between recognizing the short clips and its release date.

General Discussion

The present study points to a musical memory that is very detailed. Listeners were often able to name artist and title from very short clips. This means that they could match the clip with what must be an almost exact memory trace of the sound of the song. This memory trace must also be linked to knowledge of artist and title to allow for correct song identification. At the rate of more than 25% identification for the 400 ms clips, each song would contain an average of over 100 identifiable clips.[1] This estimate might depend somewhat on where the clip was located in the song. However, the present study found only a slight advantage for clips taken from the chorus. Naming accuracy depended somewhat on whether the clip contained a word or part-word of the title but only for the 400 ms clips. The effect of title word replicates Hyman and Rubin's (1990; Rubin, 1995) result on memory for Beatles lyrics. A factor to consider in understanding the effect of title word in the present experiments is that typical syllable durations are in the range of 100 ms (Greenberg, 1996; as cited in Patel, 2007), so a 400 ms clip might contain multiple syllables. However, this factor played a relatively minor role in these experiments; relatively few clips contained title words or part-words. It was not a significant factor for the shorter 300 ms clips that might also contain multiple syllables, and purely instrumental clips (such as *Purple Haze*) also were recognized.

As for the contents of musical memory, most listeners who identified the artist also identified the title, and visa versa. There were very few trials with only artist or title correct. Thus, these two pieces of musical information seem to be tightly linked in memory. This is also apparent in the study by Schulkind et al. (1999), who used a combined recall measure of artist, title, and cued-recall of lyrics. Both younger and older adults scored in the 80% range for music of "their era," suggesting these attributes are recalled together. Interestingly, they found that the co-occurrence of correct artist and title information was higher in the younger subjects, possibly owing to differences between the two groups in the delivery of popular music.

The study by Schellenberg, Iverson, and McKinnon (1999) should be mentioned in this context. They used 100 and 200 ms clips of popular songs, some of which were played backward or low-pass filtered. They interpreted their data to mean that the original clips were recognized although the processed clips were not. Three important differences from the present study should be noted. Only five songs were used in their experiment, long (20 s) clips were played before the short clips, and the answer sheet listed the five song titles and artists, which means that the responses were not independent (although they displayed the number of participants performing "above chance" based on the criterion of more than one correct). Given these differences, their study does not bear on the question of whether very short clips of music can be recognized spontaneously in the absence of such explicit cueing.

Listeners in the present experiments were also very clear as to whether they did or did not recognize the song. When responding "6" ("can name both artist and title") they were correct in both judgments most all the time, and when they named both artist and title correctly they almost always gave a confidence rating of "6." Thus, listeners have very accurate metacognitive knowledge of their ability to recognize songs. When they knew it, they were confident that they knew it. This supports the idea that recall of artist and title is explicit and conscious.

The results here suggest that artist and title are closely associated in memory with release date. When both were named correctly, judged and actual release dates correlated almost perfectly. This may be because they are linked directly in memory, or release date can be inferred from artist and/or title. (Other directions of inference are possible. For example, knowing the release decade might help determine from which "cover" the clip was excerpted.) It is difficult to know what other kinds music information are associated with these three, but they might include rather specific information about the album, the biography of the artist, other songs, as well as reminders of specific events and more general aspects of the time period of the song (Bartlett & Snelus, 1980; Schulkind et al., 1999). One hint regarding the kind of detail stored in memory is the agreement between listeners in the punctuation of the titles (e.g., *Don't Stop Believin', Hey Ya!)*. Another indication that music memories are very specific is Levitin's (1994, 1996) well-known finding that listeners can reproduce the starting pitch and tempo of very familiar recorded songs.

[1]The clips were also analyzed by Shazam (c), the mobile music discovery provider. None of the short clips (400 ms) were identified, but all the long clips (15 s) were with artist, title, release date, and album information.

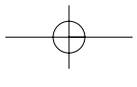
However, even when the song was not identified in the present experiments, there was a significant correlation between judged and actual release dates. This is unlike the results of the study by Bartlett and Snelus (1980) who found temporal judgments were poor unless the song was identified. A likely possibility is that the difference is due to their transcribing and playing their songs on the piano, thus eliminating various cues that might contribute to judgments of release dates such as instrumentations, timbre, vocal qualities, or even recording techniques (e.g., older recordings have larger dynamic ranges). The impression of decade found here might also be related to the changes in emotion and style associated with release date. Older songs were rated as sadder and tenderer; more recent songs were judged to be angrier. Also, older songs were judged to be more reflective/complex (blues, jazz, classical, folk) and more recent songs were judged to be more energetic/rhythmic (rap, hip-hop, soul/funk, electronica). These results suggest that musical memory contains impressions about the sound of songs from different decades, although it is doubtful that participants could describe the basis for their judgments.

In a similar way, emotional content and style did not depend on correct recognition. The emotion judgments made for the short clips were compared to those made for the long (15 s) clips. For both the 300 and 400 ms clips, there was a high level of agreement with the judgments for the long clips. A measure of emotion consistency was used that correlates the emotion "profile" (relative rankings of the five different emotions) for the short and long clips. These profiles were strongly correlated independently of recognition. Consider the contrast between the 300 ms clip trials with a very low rate of recognition and the 400 ms clip trials on which participants named both artist and title correctly. Despite the contrast, the emotion consistency did not differ between these trials. Even for the SC400 clips the difference between the "1" ("don't recognize at all") and "6" ("can name both artist and title") trials was surprisingly small.

A result of the Peretz et al. (1998) study described earlier is of special relevance here. They found that the amusic I. R., who had very poor recognition of songs, was nonetheless comparable to normal controls on judging whether songs were happy or sad. Thus, recognizing the song was not a prerequisite of making the emotion judgments, which parallels the present results. Their study examined what musical cues might carry information about emotional content. Timbre seemed not to be important because performance was similar whether the excerpts were taken from a recording or whether they were transcribed and synthesized with a piano sound. However, tempo (fast versus slow) and to some extent mode (major versus minor) did seem to influence the emotion judgments because other experiments explicitly manipulated both tempo and mode and these affected the emotion judgments. They also considered what information about tempo and mode might be contained in .5 s segments. Density of tones—reflecting tempo—reliably distinguished between the happy and sad excerpts. Mode was less clear, although the major key of the happy segments was apparent in music-theoretic analyses and judgments by music theorists. (That minor mode is less stable cognitively has been found in a variety of measures; see for example, Krumhansl, 1990).

The Bigand et al. (2005) study, which found similar emotion dimensions for short and long excerpts, is also useful for its detailed analysis of the first second of their musical segments. The number of different features that can be identified in these short segments is impressive. For example, the beginning of Beethoven's *Piano Sonata*, Opus 32, second movement begins with an "energetic melodic line of seven arpeggiated notes of a major chord in a repeated compound rhythm announcing a rapid metrical beat" (p. 1138). Schumann's *Traümerei* begins with a "single piano note in the middle high register, with a very soft attack and with full resonance predicting the start of a very slow movement" (p. 1138), an example of how much information can be contained in a single tone. More generally, Bigand et al. made two central points. The first is that a great deal of information can be contained in a musical segment as short as 1 s and this can induce emotions similar to those for much longer segments. The second is that the emotional cues in musical compositions come in interacting combinations of a variety of different musical attributes. Both these studies used classical music, so that it is likely that different cues may signal emotional content in popular music.

Style judgments were treated similarly by creating a style "profile" (relative rankings of the four different styles). Direct comparisons with Gjerdingen and Perrott's (2008) study cannot be made because of the differences in the methods. Here, well-known songs were used and only four general style categories. Both would boost style consistency. Given this, the present results are not impressive compared with their participants' abilities to make fine distinction between ten genre categories. However, as far as style classification is concerned, the present study extends theirs in three ways. First, it showed that judgments of style were quite consistent across individuals; they used self-agreement. Second, it provided information about how style judgments relate to other musical attributes that listeners can extract from short segments, including artist, title, emotional content, and release date. Third, and of most interest, was the finding that style



consistency was high even when the short clips were not identified. Their study did not report whether participants recognized the particular songs.

In sum, a pattern of relationships was found between artist and title identification, confidence ratings, and judgments of decade of release, emotional content, and style. This pattern bears on the organization of musical memory. Artist and title were almost always identified together and when they were, participants were confident in their identifications. With correct identifications came remarkably accurate judgments of release decade. This suggests close associations in memory between artist, title, and decade of release, either by direct links or inference. Even without correct identification, impressions of content and style of the short clips were quite consistent with those for the long clips; in fact, almost as consistent as when the clips were correctly identified by both artist and title. Also, without identification, decade of release judgments were still quite consistent. These results suggest two routes to decade of release: one that relies on the kinds of intuitions used for judging emotion and style, and one that is associated with explicit recall of artist and title.

Finally, let us consider an unexpected finding of the study: the preference, which was quite strong in both experiments, for older songs. Holbrook and Schindler (1989) found a clear inverted U-shaped curve of preference as a function of the age of participant when the song was released, with a peak in early adulthood. It has been speculated that this reflects a plasticity period in adolescence and early adulthood. This would predict that these college-aged participants would be on the ascending part of the curve with a preference for more recent songs. Schulkind et al. (1999) found the predicted function for young adults in measures of familiarity, the combined recall described above, recollections of general memories, and emotionality. Bartlett and Snelus (1980) did not find this pattern but they did find that younger adults did not recognize the songs that were released before they were born or were young children. The experiments here were not designed to systematically study preferences as a function of release date, so the findings here are only suggestive.

Taken together, the results suggest that the preference for the older songs was a consequence, at least in part, of hearing the songs growing up as children. An alternative possibility is that the older songs may be preferred because being on the list of songs "of all time" means that they have stood the test of time and are intrinsically better in some way. They might be richer musically, more distinctive, listened to more often, or otherwise special, so that they are more firmly established in memory. This would predict greater recognition of the older songs. However, decade of release had no effect on recognition of the short clips. The apparent predominance of songs from the decade 1965–1974 is illusory because many of the songs from the same era were not recognized. In addition, or instead, the experiment used a number of songs that to a 60's ear sound "iconic" in the sense that they are emblematic of the cultural conditions of that time. These songs also may have the same iconic status for the younger participants (or one that is modified by the present cultural conditions). If so, then one might expect greater consistency of emotion and style judgments because of their strong contextual associations. However, decade of release had no effect on the consistency of emotional content or style and, if anything, there were trends in the opposite direction with more consistency for the more recent songs.

Instead, the results suggest that the preference for the older songs is a function of the exposure to the older songs going back to the participants' childhoods. It would not be surprising that childhood experiences of music would have an effect at a time of rapid learning in other domains, particularly language. The participants often remarked with apparent fondness that they had learned the older songs from their parents at home. Almost 75% of the participants reported rock as being one of the most frequent styles they heard when they were growing up. On average, the age of fathers was almost 35 years older than the age of the participants, and mothers were about 32 years older. In addition, the majority of participants had one or more older siblings. This means that the participants would be hearing a great deal of rock music at home released at least as far back as the mid-1970's. If strong and stable musical preferences begin to develop during childhood this early exposure would lead to the observed preferences.

Other evidence suggested that the preference for the older songs is not because they are predominant in the listeners' current listening repertoire. At present, the participants seemed to be following the listening pattern typical of adolescents and young adults. They said the most important musical influences currently came from peers, followed by music on the radio/web/etc., with only relatively minor influences of the music of their parents. Their style choices also suggested that they are listening primarily to music of their generation. They reported listening most often to intense/aggressive music (alternative, rock, heavy metal) more prevalent in recent music. The next most frequent style was reflective/complex music (classical, jazz, blues, folk), but this should be viewed in light of the fact that many of the participants were heavily involved in studying and performing classical music. Thus, it appears that the observed preference for the older songs was not simply that they were included

in their current listening repertoire, but that they became familiar with them at home when they were younger and this establishes a strong preference for the older songs that persists into adulthood. The pattern found here differs from earlier experiments. The ready availability of recorded music during the participants' lifetimes seems a factor likely contributing to the change. In sum, the present results suggest that the notion of a plasticity period may not apply to current generations in the same way as it did for earlier generations, although more systematic study of this question is needed.

Finally, we can consider how the present findings relate to studies in other domains. The results invite comparison with those for visual memory given that both visual and musical memories appear to have a very large, possibly limitless capacity. However, a number of methodological differences prevent making direct comparisons. One important difference is that the popular songs were highly familiar to the participants before the experiment, whereas the pictures used in the visual memory studies were novel (although the objects in them would generally be very familiar). These considerations might make the visual memory results seem more impressive. On the other hand, the task here was one of recall rather than forced-choice old/new recognition. The required retrieval of title and artist is more demanding, seemingly making the music task harder.

However, there is one study using music following the same methodology as the visual memory experiments: the presentation of a large number of items followed by a forced-choice old/new recognition task. One of Standing's (1973) experiments included 200 music segments (12 s long) randomly selected from classical, jazz, and popular pieces. Music performance (83%) was lower than for pictures (94%), but at about the same level as words (visual presentation 84%, auditory presentation 86%). Performance on the music stimuli is quite impressive for a number of reasons. His participants were all nonmusicians, and they would likely be unfamiliar with most of the music used in the experiment. In addition, the 12 s segments of music contain a great deal of complex information extended over time. Finally, presenting the entire set of interrupted segments of music took 50 mins which would be taxing and tedious; indeed, participants complained about the familiarization phase. Despite these difficulties, this study points to a capacity to learn and quite accurately recognize a large number of unfamiliar musical segments.

Can the present study be used to roughly estimate of the capacity of musical memory? The songs were identified (both artist and title) from the short (400 ms) clips more than 25% time. They are representative of a much larger set of songs that appears on lists of the top songs "of all time" or that are currently popular. Participants reported listening to popular music about 20–21 hours per week for a total of more than 66,000 minutes per year. Assuming the songs are 3 minutes long, they would hear in the range of 22,000 songs per year. This would obviously include many repetitions, but one might expect that each year a considerable number of songs would become as familiar as those used in the experiment. Song identification did not depend on release date, suggesting that participants were familiar with songs going back at least to the 1960's, further increasing the number of familiar songs. Moreover, the listening repertoires of individuals would only partially overlap with those of others, so their own personal repertoire would be considerably larger than that shared by the group. Lastly, the estimate does not consider other musical styles. This rough calculations points to a musical memory that has an extremely large capacity.[2]

Finally, the results show that listeners can glean a great deal of information from very "thin slices" of music, including artist, title, release date, emotional content, and style. Other research has shown that short exposures ("thin slices") are sufficient for judging a wide range of social characteristics from brief exposures. However, the exposures in those experiments were considerably longer than the music clips. The durations in Ambady and Rosenthal's (1002) metaanalysis ranged from 3.5 to 5 s with more than two-thirds of them having durations of a minute or more. It might be argued that the participants in the studies they reviewed were making complex social judgments (such as, status of patients, anxiety, and gender of person being addressed). On the other hand, recall of artist and title requires very specific information that is either correct or incorrect. More comparable are the correlations between the social judgments and objective measures (averaging .39) and the correlations between the emotional content and style judgments for the short and long clips (many of which were higher than .80–although the differences in degrees of freedom should be noted). To understand this difference it seems intuitive that information in social signals, such as gestures and facial expressions, is conveyed more slowly than in music. In contrast, information in speech is generally conveyed faster than in music–100 for a typical syllable compared to typical shortest duration of tones in music in the range 200–300 (Fraisse, 1982). So it would be interesting to

[2]One participant reported having a year's worth of music on his computer; he estimated that he knew 60% of them as well as those used in the experiment. This gives a total of over 100,000 different familiar songs and 10,000,000 identifiable 400 ms clips.

compare music and speech to determine what properties are conveyed in very short segments.

The other point of contact with the "thin slice" literature is that some of the music judgments could be made based on intuitions rather than explicit criterion. Even when artist and title were not identified, listeners made quite consistent judgments of emotional content, style and, to some extent, release decade based on intuitions rather than explicit knowledge. Similarly, in the case of social judgments it is assumed that the participants were largely unaware of the basis for their judgments. Some progress has been made on the objective cues that convey social information. For example, Kraus and Keltner (2009) were able to trace ratings of socioeconomic status to behaviors of disengagement (for example, doodling, self-grooming) and engagement (head nods, laughter). Pentland (2008) described a device (a "sociometer") for analyzing subtle social cues that may contribute to very concrete and consequential behavioral outcomes (exchanging phone numbers in speed dating, getting a job, making contacts at a professional meeting). In the present case, ongoing analyses are being conducted to try to identify the acoustic cues that allow intuitive judgments about music to be made.

Author Note

The comments of two reviewers on an earlier version of this manuscript are gratefully acknowledged.

Correspondence concerning this article should be addressed to Carol L. Krumhansl, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853. E-MAIL: clk4@cornell.edu

References

AMBADY, N., BERNIERI, N., & RICHESON, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. *Advances in Experimental Social Psychology*, *32*, 201–257.

AMBADY, N., & ROSENTHAL, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256–74.

BARTLETT, J. C., & SNELUS, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology, 93,* 551–560.

BIGAND, E., VIEILLARD, S., MADURELL, F., MAROZEAU, J., & DACQUET, A. (2005). 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.

BRADY, T. F., KONKLE, T., ALVAREZ, G. A., & OLIVA, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*, 14325–14329.

BRUNSWICK, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.

FRAISSE, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (1st ed.). New York: Academic Press.

GJERDINGEN, R. O., & PERROTT, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research, 37*, 93–100.

GLADWELL, M. (2005). *Blink: The power of thinking without thinking.* New York: Back Bay Books.

HEVNER, K. (1935a). The affective character of the major and minor modes in music. *American Journal of Psychology, 47*, 103–18.

HEVNER, K. (1935b). Expression in music: A discussion of experimental studies and theories. *Psychological Review, 42*, 186–204.

HOLBROOK. M. B., & SCHINDLER, R. M. (1989). Some exploratory findings on the development of musical tastes. *Journal of Consumer Research, 16*, 119–24.

HYMAN, I. E., JR., & RUBIN, D. C. (1990). Memorabeatlia: A naturalistic study of long-term memory. *Memory and Cognition, 18*, 205–214.

JOHN, O. P., & SRIVASTAVA, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 102–138). New York: Guilford Press.

JUSLIN, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1797–1813.

JUSLIN, P. N., & LAUKKA, P. (2003). Communication of emotions in vocal expressions and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770–814.

KRAUS, M. W., & KELTNER, D. (2009). Signs of socioeconomic status: A thin-slicing approach. *Psychological Science, 20*, 99–106.

KRUMHANSL, C. L. (1990*). Cognitive foundations of musical pitch*. New York: Oxford.

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

LEVITIN, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception and Psychophysics, 56*, 414–423.

LEVITIN, D. J. (1996). Additional evidence that auditory memory is absolute. *Perception and Psychophysics, 58*, 927–935.

NICKERSON, R. S. (1965). Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology, 19*, 155–160.

NORTH, A. C., & HARGREAVES, D. J. (1999). Music and adolescent identity. *Music Education Research*, *1*, 75–92.

PATEL, A. (2007). *Music, language, and the brain*. New York: Oxford University Press.

PENTLAND, A. (2008). *Honest signals: How they shape our world.* Cambridge, MA: MIT Press.

PERETZ, I., GAGNON, L., & BOUCHARD, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition, 68*, 111–141.

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

RENTFROW, P. J., & GOSLING, S. D. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science, 17*, 236–242.

RUBIN, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. New York: Oxford University Press.

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.

SCHULKIND, M. D., HENNIS, L. K., & RUBIN, D. C. (1999). Music, emotion, and autobiographical memory: They're playing our song. *Memory and Cognition, 27*, 948–955.

SHEPARD, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior, 6*, 156–163.

STANDING, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25,* 207–222.

STANDING, L., CANEZIO, J., & HABER, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2,500 visual stimuli. *Psychonomic Science, 19*, 73–74.

WALLIN, N. L., MERKER, B., & BROWN, S. (Eds.) *The origins of music*. Cambridge, MA: MIT Press.