

Identifying songs from their piano-driven opening chords

Ivan Jimenez¹ Tuire Kuusi¹ Isabella Czedik-Eysenberg² & Christoph Reuter²

¹ Sibelius Academy, University of the Arts Helsinki, Finland

² Institute for Systematic Musicology, University of Vienna, Austria

Author Note

This paper is the *Final Draft* of the following article:

Jimenez, I., Kuusi, T., Czedik-Eysenberg, I., & Reuter, C. (2021). Identifying songs from their piano-driven opening chords. *Musicae Scientiae*. DOI: 10.1177/10298649211003631.

Correspondence concerning this article should be addressed to Ivan Jimenez,
Ivan Jimenez, Sibelius Academy, University of the Arts Helsinki, Helsinki, P.O. Box 30,
FI-00097 UNIARTS, Finland.

Email: ivan.jimenez.rodriguez@uniarts.fi

Abstract

It is currently believed that timbre plays a primary role in the identification of songs from very brief excerpts of music. However, its specific contribution, and those of other characteristics of the music, remain unclear. The aim of the present study was to investigate the contributions of timbre and chord type, voicing, and duration to participants' ability to identify songs from their opening chords, played on the piano alone or on the piano and one other instrument. Ninety-three participants were asked to identify 20 songs from their opening chords. They were also asked to estimate the similarity between pairs of chords. We evaluated the contribution of ten characteristics of chords to song identification rates, including brightness, pitch register, and duration, because of their high perceptual salience, and others chosen on the basis of theoretical predictions relating to auditory long-term memory. Song identification rates were associated with the chords' brightness, familiarity and, to a lesser extent, pitch register, but not participants' musical background. The results of the study suggest that listeners with and without musical training can identify songs from their opening chords, cued by both their timbral and non-timbral characteristics.

Keywords: chord type, chord voicing, music memory, timbre, pitch register, popular music

Identifying Songs from their Piano-Driven Opening Chords

Stimuli consisting of very brief excerpts of commercial recordings have been used for various purposes, including the investigation of genre identification (Gjerdingen & Perrott, 2008; Mace et al., 2012), familiarity (Bigand et al., 2009; Filipic et al., 2010), preference (Belfi et al., 2018), emotion (Bachorik et al., 2009; Bigand et al., 2005; Peretz et al., 1998), and timbre (Layman & Dowling, 2018; Siedenburg & Müllensiefen, 2017). Since brief stimuli often contain very little melodic and rhythmic information, researchers have interpreted the results as indicating that musical memories can include very detailed information about timbre (McAdams & Siedenburg, 2019; Wallmark et al., 2018). Perhaps one of the most compelling pieces of evidence for the timbral specificity of musical memories is provided by listeners' ability to identify songs from excerpts as short as 100 ms (Schellenberg et al., 1999). A recent study of listeners' ability to identify songs from short chord progressions suggests that auditory long-term memory for song also includes characteristics of individual chords such as their type and the ways in which their pitches are voiced (Kuusi et al., 2021). It is therefore possible that participants in previous experiments (e.g., Krumhansl, 2010; McKellar & Cohen, 2015) made use of such characteristics, as well as timbre, when identifying songs from very brief excerpts. It may be difficult to separate timbre from chord type and voicing since they are all based on the same acoustic raw material, in the form of a bundle of frequencies (Hasegawa, 2019; London, 2011). This difficulty has posed major challenges to the development of accurate automatic chord recognition systems (Cho & Bello, 2013; Mauch & Dixon, 2010; Ueda et al., 2010). The specific contributions of timbral and other characteristics of individual chords to listeners' ability to identify songs from brief excerpts remains unclear.

The identification of songs from brief excerpts

To date, only three studies of song identification from brief excerpts have been conducted. Schellenberg et al. (1999) used a closed-set identification task in which the participants were given a list of the songs they would be asked to identify before they carried out the task. The researchers used stimuli taken from five songs and a pre-task familiarization phase. Song identification was above chance level even for stimuli lasting only 100 ms but fell below chance level when the stimuli were played backwards, or when some of their upper frequencies were filtered out. These results suggest that aspects of timbre both dependent on and independent of time are important for identifying songs. The excerpt that was identified most often did not include any vocals, which suggests that listeners do not have to recognise the voice of a singer or hear lyrics to be able to identify a song.

Krumhansl (2010) used an open-set identification task in which participants were not given any information in advance about the songs to be used as experimental stimuli. Participants were able to identify both artist and song title from 25% of 400-ms excerpts and from 11% of 300-ms excerpts. Additionally, they identified songs from 400-ms excerpts that included a word or part of a word from the title more often than those that did not. McKellar and Cohen (2015) tested participants using 400-ms excerpts only, and found song identification rates comparable to Krumhansl's. Unlike Krumhansl, they also found a significant effect of year of release such that more recent songs were easier to identify.

In sum, there is some evidence that time-dependent and time-independent timbral characteristics of brief excerpts (Schellenberg et al., 1999), their duration, the presence of lyrics, melodic information (Krumhansl, 2010), and year of release (McKellar & Cohen, 2015) can affect song identification. Research on this topic is scarce, however, and has made use of two very different approaches (i.e., closed-set vs. open-set identification tasks). Accordingly, we aimed to improve the understanding of song identification from brief

excerpts in the present study by investigating the effects of chord type, voicing, and register, duration, and, to some extent, timbre. In what follows, we provide some background to specific aspects of the study.

Closed- vs. open-set identification

Helping participants to identify songs by giving them a list of pieces during the experimental task (closed-set identification) not only tends to inflate general performance (Schellenberg et al., 1999) but can also artificially boost the effect of some experimental variables. While recognition is generally easier than recall, closed-set identification tasks also allow specific mental representations to be activated in listeners before they hear the stimuli. These representations are then compared with the stimuli that are heard; listeners thus use a predominantly top-down strategy that greatly facilitates identification (Hébert & Peretz, 1997). A small closed set may also allow participants to identify some songs based on genre, since timbral characteristics can be sufficient for listeners to differentiate between genres (Mace et al., 2012). Support for the contribution of timbral cues to song identification is provided by the fact that participants' performance in Schellenberg et al.'s (1999) study fell below chance level once the excerpts were played backwards and low-pass filtered. By contrast, the use of an open-set identification task greatly reduces the potential role of top-down strategies and identification via genre. It is difficult, however, for researchers using open-set approaches to control for participants' familiarity with the songs presented in the task. In the present study we administered an open-set identification task, and we will describe in the Methods section how we controlled for participants' familiarity with the songs used in our experiment.

Opening chords of the songs

Even though the use of voice(s) can be an important aspect of overall timbre in popular music, and play an important role in auditory memory for songs (see e.g., Weiss et

al., 2019), we decided to use opening-chord stimuli without vocals. This was because we wanted to reduce the number of potential variables affecting song identification, and thereby make it easier for us to interpret our results. We decided to limit the stimuli to block chords, that is, excluding broken chords and arpeggios. Further, we decided to use only *piano-driven* chords, that is, chords played on the piano or on the piano and another instrument. In the latter case we used stimuli meeting four criteria: (1) only one other instrument was heard in each chord, besides the piano (see Stimuli); (2) the other instrument doubled the piano rhythmically; (3) the other instrument produced pitches either higher or lower than those produced on the piano, so that the piano could be heard clearly in its own register; (4) the other instrument was never played louder than the piano (for more details about stimuli, please see Appendix A).

These decisions allowed us to reduce the potential effects on song identification of melodic and rhythmic factors, and instrumental categories and their stylistic associations. Additionally, the fact that all the excerpts were taken from the very beginning of the song meant that the position of the excerpts within the song did not have to be considered as a potential contributor to identification rates.

Duration of excerpts

Studies of rapid musical recognition categorize stimuli either according to the number of notes presented (Dalla Bella et al., 2003; Schulkind, 2000, 2002, 2004a, 2004b; Schulkind & Davis, 2013; Schulkind et al., 2003), or their duration in milliseconds. The use of same-duration stimuli often produces fragments of music that rarely correspond to full musical units. We decided to use the first complete block chord of each song as stimuli, since this allowed us to focus on song identification from stimuli that are, despite their brevity, complete musical units.

Effect of chord type and voicing

There is evidence that songs can be identified from two chords even when the chords are not taken from the original audio data but composed using MIDI tones in such a way that timbral, melodic, and rhythmic cues are considerably downgraded (Kuusi et al., 2021).

Research on song identification from two chords shows that chord type and voicing are encoded in long-term memory for songs. Although Kuusi et al.'s results suggest that listeners' mental representations of chord type and voicing are at least partially independent of their mental representations of timbre, it might be more difficult to perceive chord type and voicing as separate from timbre in single block chords than in longer passages of music, since the rhythmic and melodic information contained in single block chords do not contribute to auditory stream segregation or prevent the fusion of pitches into a single vertical gestalt.

To date, there has not been any research on the effects of chord type and voicing on the perception of timbre, although there is evidence that it is harder for both non-musicians and musicians to determine whether two chords are harmonically the same if they are played on different instruments (Beal, 1985; Cho et al., 1991). Additionally, the spectral features of timbre have been shown to affect the perception of chords. For instance, the manipulation of the degree of distortion in an electric guitar can affect both spectral similarity and the neural discrimination of different chord types (Virtala et al., 2018). These results are consistent with the practice of spectralist composers of the 1970s for whom timbre, chord type, and voicing were considered part of the same continuum (Grisey & Fineberg, 2000; Pressnitzer & McAdams, 2000; Rose, 1996).

Considering that timbre can be difficult to perceive as distinct from chord type and voicing in very brief excerpts of music, and that song identification is possible from just two chords even when most extra-harmonic cues are downgraded, we anticipated that chord type and voicing could play a role in song identification from opening chords. To investigate this

possibility, we used a range of different types of chords (e.g., major, minor, dominant seventh, major seventh, etc.) and voicings (e.g., various pitch spans, pitch registers) when selecting the stimuli for our song-identification tasks.

Methods

Participants

A total of 131 participants (53 male, 78 female; $M = 36.35$, $SD = 12.12$; for more information about the participants, see Results) completed the experiment online. Participants were recruited using Amazon Mechanical Turk (MTurk), a crowdsourcing platform that provides access to more than 100,000 potential participants (Difallah et al., 2018). Armitage and Eerola (2020) have shown that the results of music cognition experiments carried out in standard lab settings are comparable to those from online experiments that recruit participants using services similar to MTurk. The software PsyToolkit was used to collect data (Stoet, 2010, 2017). The invitation to take part in the experiment was addressed to people who were “often good at naming pop songs.” Prospective participants were given a general description of the experiment and told that headphones or earphones were required. The experiment website was accessed a total of 838 times between April 5 and April 12, 2019. The pre-test using headphones or earphones (referred to from now on as “headphones/earphones”) was completed by 52% of prospective participants, of whom 30% also completed the entire experiment. This completion rate (30%) is relatively low compared to completion rates for other online experiments (Bosnjak & Tuten, 2003; O’Neil & Penrod, 2001; O’Neil et al., 2003; Tuten et al., 2004), and may suggest that the experimental tasks were somewhat demanding.

In order to identify participants who were very familiar with the songs used in the experiment, we took two steps. First, we discarded 20 participants who reported having never

heard 40% or more of the test songs. Second, we discarded 13 participants who could not name, on the basis of listening to commercial recordings, at least 60% of the songs they reported having heard five or more times in their lives. Finally, we discarded five participants whose data were partially lost due to a glitch in the system. The total number of participants whose responses were included in our main analysis was 93 (30 male, 63 female; $M = 37.08$, $SD = 11.64$), indicating that there were more than twice as many female as male participants. Even though female participants have been shown to use adjective-scale evaluations of harmonic intervals and chords slightly differently from male participants (Costa et al., 2000; Lahdelma & Eerola, 2016), there is no evidence in the literature, to our knowledge, of a gender effect on song identification from brief excerpts. The majority of the participants in our study also had some experience of playing musical instruments (for details, see Table 1).

Table 1

Participants' experience playing and practicing musical instruments

Experience	Ps	%Ps
5 years or more	33	35.5%
Less than 5 years	30	32.3%
Had never played an instrument	30	32.3%
Total	93	100%

Stimuli: Song-identification task

Songs were selected for the song-identification tasks based on two separate online pilots using well-known commercial recordings of the songs that started with a block chord played on the piano alone or on the piano with another instrument (e.g., bass guitar, synth, bass drum). We chose songs that had had been heard by more than 100,000 listeners to Last.fm. Based on these two pilots, we selected 20 songs for the main experiment (see Appendices A and B). All stimuli were extracted from commercial CDs as standard WAV

files. In order to prevent clicks, a 10-ms amplitude ramp was added at the end of each stimulus. The duration of the longer excerpts was 15 seconds. All the longer excerpts included vocals, but no excerpt contained any part of a song's title. We equalized the loudness of all the stimuli used in the song-identification tasks subjectively (although we used an automatic procedure for equalizing the loudness of stimuli in the chord similarity estimation task, as described below), because we believed this would be sufficient to ensure that participants heard the stimuli clearly without being distracted or experiencing negative emotional responses as the result of unexpected changes in loudness.

Stimuli: Similarity-estimation task

The purpose of the similarity-estimation task was to single out features that can make a chord sound particularly distinctive and may thus affect song identification from single chords. Although we wanted the chords used in the similarity-estimation task to be as similar as possible to the stimuli used in the song-identification task, so as to increase the chances that the singled-out features would also be relevant to song identification, we did not use exactly the same stimuli. The main differences between the stimuli for the two tasks were as follows. First, we used only 12 chords in the similarity-estimation task. Second, we chose only chords with the root C (e.g., C major, C minor), because the immediate succession of different-root chords in similarity-estimation tasks necessarily involves root motion (i.e., an interval between the pitches representing the roots of different chords), which was irrelevant to our investigation. Third, we modified some of the chords to increase the heterogeneity of the set by adding bass-guitar or piano tones to some of the originals, using sample instruments from Logic Pro X. In this way we created a set of 12 chords including five types of chord (Maj, min, Maj7, min7, Dom7), and one without a third, in that it contained the P5 interval; the set was thus heterogeneous in terms of timbre, register, and number of pitches. We also changed the duration of the original chords. Stimuli were trimmed so that each chord

had a different duration (400, 420, 440, 470, 500, 540, 580, 620, 660, 700, 750, and 800 ms). This set of durations falls within the range of the chord durations used in the song-identification task and is above the threshold of the just-noticeable difference observed in temporal discrimination tasks (Halpern & Darwin, 1982; Rammsayer & Altenmüller, 2006).

Further, we removed excessive hiss by implementing the multiband de-noising function of Amadeus Pro, and we equalized the loudness of the audio files by implementing a Matlab script written for this purpose.¹ We also converted the stimuli to mono files by using Amadeus Pro, after which we inspected them to verify that there was no noticeable distortion generated by phase issues.

Our treatment of loudness in the similarity-estimation task stimuli deserves some further clarification. The perception of loudness is a complex phenomenon, and it is argued that perceived loudness can never really be equal for every listener (Hajda et al., 1997). In the case of musical instruments, a distinction can often be made between loudness and listeners' perception of the physical force employed by a performer to produce a sound event (e.g., keystroke force) (Fabiani & Friberg, 2011). Loudness equalization does not alter the acoustic cues of keystroke force such as brightness (Palmer & Brown, 1989) and attack time (Askenfelt & Jansson, 1993). This is likely to be the case not only for stimuli that feature acoustic pianos, but also for those stimuli featuring electric pianos and other electronic instruments that have been designed to imitate some of the properties of acoustic instruments. Accordingly, we did not expect loudness equalization in this study to neutralize the effect of perceived keystroke force, but to prevent large differences in the general loudness of the audio files from overemphasizing the perceptual salience of keystroke force.

¹Loudness was calculated using the function `Loudness_ANSI_S34_2007` (ANSI, 2007), which is based on the Moore and Glasberg model for steady sounds from the Matlab Genesis loudness toolbox. For the toolbox, see http://alpha.tmit.bme.hu/speech/docs/education/kognitiv_announcement_MATLAB_loudness_toolbox.txt. Unfortunately, this information is no longer available online.

Given a set of 12 chords, 132 trials would have been required to estimate the similarity between all possible pairs. Since the main focus of the study was song identification, not similarity estimation, we reduced the number of trials to 66 by presenting each pair three times in immediate succession (e.g., A-B-A-B-A-B) rather than each pair in both orders (e.g., A-B and B-A). The stimuli-onset asynchrony between chords in each trial was kept constant at 900 ms to avoid potentially distracting pulse irregularities. Because the chords had different durations, however, the silences between chords varied between 100 and 500 ms. This was considered acceptable, since previous research has found no significant decay of auditory sensory memory when the interval between block chords falls in that range (Tekman & Bharucha, 1992; Virtala et al., 2014).

Stimuli: Chord variables

The chords used in the similarity-estimation task were analyzed in terms of timbral characteristics extracted from the signal, as well as in terms of chord-type features, chord-voicing characteristics, and one song-related variable. Table 2 lists these variables and indicates those that were subsequently included in the analysis of the data from the song-identification task. The process for selecting these ten variables will be explained in the Results. The ten variables are described in Appendix C.

Table 2

Chord variables

Timbral characteristics	Chord-type features
Brightness**	Number of pitch classes
Envelope flatness	Chordal third (M3, m3, no 3 rd)
Attack time****	Harmonic similarity****
Harmonic/percussive energy ratio	Chord-type frequency of occurrence****
Inharmonicity	Musicians' and non-musicians' consonance ratings
MFCC 1-13	
Roughness	Chord-voicing characteristics
Spectral centroid**	Number of notes
Spectral entropy	Number of notes below middle c
Spectral flatness	Number of notes that are middle c or higher
Spectral kurtosis	General register*

Spectral rolloff**	Highest note*
Spectral flux	Lowest note
Zero crossing rate	Pitch span
	Duration***
Song related variable	Instrument
Year of release****	

Note. * Dimension 1 in chord-similarity estimates

** Dimension 2 in chord-similarity estimates

*** Dimension 3 in chord-similarity estimates

**** Other variables considered for the main analysis

Procedure

The project was approved by the Research Ethics Committee of the University of the Arts Helsinki. Participants started the experimental session by reading a brief description of the tasks involved in the study. After that, they took a headphones/earphones pre-test that was a shortened version of a test designed by Wycisk et al. (2018) and took three minutes to complete, on average. We wanted to evaluate the participants' hearing of low sinusoidal tones ranging from 20 Hz to 200 Hz, since the capacity of headphones/earphones to reproduce low frequencies is one indicator of their general sound quality (Arora et al., 2006; Wersényi, 2010; Breebaart, 2017). Additionally, we asked about the noise level in the acoustic environment where the participants did the experiment. When they had carried out the pre-test, participants were asked a series of questions about their background, for example their experience of playing musical instruments in general and harmonic instruments in particular, as well as their experience of playing popular music (*participant variables*; see Table 3 and Appendix D).

After answering the questions, participants completed the first part of the main task, which consisted of identifying 20 songs from their opening chords. They were asked to identify the artist or band and the song title, and they could give comments. They were told that if they could not remember the name of the artist or band, or the title of the song, they could write down some of the lyrics or provide any other piece of information that would allow us to identify the song they were thinking of. Although participants were given

unlimited time to write their responses, they were allowed to listen to each chord only once, and were asked not to consult the internet, apps, playlists, or personal music collections during the experiment. The order of the songs was randomized for each participant.

The participants then completed the second part of the main task: estimating the similarities between pairs of chords. The order of presentation of the pairs was randomized. Participants were asked to use a seven-point scale ranging from 1 (*very dissimilar*) to 7 (*very similar*). Participants were given the following instructions:

You will hear two short excerpts alternating in an a-b-a-b-a-b pattern. By clicking on any of the points of the scale below, please indicate how similar or dissimilar the two short excerpts in the clip are from each other. We are interested in your general impression of similarity. We want you to rely on your intuition as much as possible. Listen to the entire clip but try to respond quickly.

Four practice trials were included, so that the participants could get used to the task and calibrate their use of the numerical scale representing the full range of levels of similarity or dissimilarity between the pairs of stimuli used in the task. The four trials differed from each other in terms of both degree and type of similarity/dissimilarity between pairs of stimuli. None of the eight chords used in the practice trials were used in the 66 main similarity-estimation trials.

After completing the similarity-estimation task, participants were asked whether they could “identify major and minor chords just by listening to them.” Participants who responded “yes”, “most of the time”, and “only sometimes” then undertook a short aural chord-type recognition test comprising trials of different degrees of difficulty. On each of these trials, participants were played a piano chord and were asked to select the correct chord type from a list. The results of this test were subsequently included as a participant variable (V10, Table 3).

In the third and final part of the main task, participants were asked to listen to and identify 15-second excerpts from the 20 songs that were used in the first part of the task. As in the previous parts, the order of presentation of the songs was randomized. As well as being asked to provide the name of the artist/band, song title, or other identifying information, participants were asked multiple-choice questions as to how often they had heard the song and, in the case of participants who had reported playing instruments, how often they had played the song. Most participants completed the entire experimental session in less than 40 minutes.

Design of analyses

Here we summarize our analytic procedure. Further details of the analyses will be given in Results. We analyzed the participant variables using Principal Component Analysis (PCA), after which we used the components in a linear regression to see if they affected song identification from opening chords. We then analyzed the similarity-estimation data using Multidimensional Scaling and extracted the characteristics of chords most likely to affect similarity estimates. These characteristics, together with a further three that were related to single chords, were used as variables in a PCA and regression analysis to investigate factors affecting chord identification. Finally, we used one-way analyses of variance (ANOVA) to compare the characteristics of the presented chords and the characteristics of the opening chords of songs that participants mentioned when providing incorrect identifications.

Results

Analysis of items using piano, and piano and some other instrument

First, we analyzed the percentages of songs correctly identified from their opening chords. The songs were those that each participant had identified correctly from hearing commercial recordings of the songs, and the percentages were calculated for each participant

(1198 cases, 64% of all trials). Additionally, in order to avoid any noise introduced by top-down influences, we did not deem a song to have been correctly identified if the participant had named it at the beginning of the experiment as one that starts with piano (15 cases, 0.8% of total trials) or if they named it as a song for which an opening chord was not presented (125 cases, 6.7% of total trials); instead, these were analyzed separately (see Analysis of other chords). We calculated identification rates for the ten songs using piano timbre only, and the ten songs using the timbre of piano and another instrument, separately. Average identification rates were similar for the two groups of songs (27% and 29%, respectively), suggesting that identification was not affected by the addition of an instrument to the piano timbre. Accordingly, these songs were considered together in subsequent analyses.

Analysis of participant variables

We conducted an exploratory factor analysis (PCA, varimax rotation) on the participant variables (see Table 3). These included the sound quality of headphones/earphones and potential interference from environmental noise (V1); musical training and playing chords (V3–V10); and the familiarity of the songs used in the experiment (V11–V14). For this set of 14 variables, the KMO test (.719) and Bartlett's test ($p < .001$) indicated that a PCA could be conducted.

The analysis revealed a five-factor structure explaining 71.32% of the variance. The structure was understandable, and easy to interpret as follows (see the highlighted numbers in the Varimax-rotated matrix in Table 3). Component 1 consisted of variables related to the participants' musical training. Components 2 and 3 were related to the participants' familiarity with the songs used in the test, as the result of either having played the chords of the songs (C2) or having listened to the songs (C3). The last two components were related to the age of the participant when they started taking lessons on an instrument (C4), and the

sound quality of the headphones/earphones they used during the experiment and potential interference from environmental noise (C5).

Table 3

Varimax-rotated component matrix with 14 participant variables. The highest loadings on each component are in bold print.

Variables	Rotated Component Matrix				
	Component				
	C1	C2	C3	C4	C5
V1 Headphone minimum frequency	.093	.005	.051	-.038	.940
V2 Age	.037	-.383	.541	.402	-.211
V3 Main instrument years	.799	-.162	.079	.199	.094
V4 Playing by ear	.768	.289	.097	-.016	.103
V5 Playing piano	.701	.232	.240	-.245	.060
V6 Playing guitar	.545	.463	.030	.375	-.087
V7 Playing popular music on piano	.689	.457	.138	-.290	-.010
V8 Age at first instrument lessons	-.211	.024	.004	.767	-.008
V9 Ability to name chords	.713	.141	-.213	-.169	-.097
V10 Score for identifying types of chords	.749	.059	-.236	-.255	.023
V11 Total songs heard of 20	.081	.254	.535	-.373	-.355
V12 Average times songs were heard	-.044	.081	.863	.012	.191
V13 % of songs played	.237	.887	-.012	-.001	-.046
V14 Average times songs were played	.100	.840	.071	-.037	.021

We conducted a linear regression analysis (all variables entered at the same time) to find out if the five components predicted the participants' ability to identify the 20 songs from their opening chords. The dependent variable was the percentage of correct identifications from the opening chord (the average identification rates can be found in Appendix B). There was no multicollinearity between the predictor variables, since orthogonal rotation (varimax) was used in the PCA analysis. The residuals were normally distributed and linear, and the Durbin-Watson statistic (2.308; see Table 4) was acceptable. The analysis revealed, however, that the participant variables did not predict the percentages of songs correctly identified from their opening chords.

Table 4
Model summary

Model Summary										
					Change Statistics					
Model	<i>R</i>	<i>R</i> ²	Adjusted <i>R</i> ²	Std. Error of the Estimate	<i>R</i> ² Change	<i>F</i> Change	<i>df</i> 1	<i>df</i> 2	<i>p</i>	Durbin-Watson
1	.185	.034	-.054	21.604%	.034	.390	5	55	.854	2.308

NOTE. Predictors: (Constant), C1, C2, C3, C4, C5

Analysis of the similarity data (66 chord pairs)

This being the case, we conducted a second set of analyses to study the role of the chord variables in song identification. We began by analyzing the variables affecting similarity estimation, and then grouped those variables using PCA and regression analysis to find out if these variables predicted song identification.

The averaged similarity estimates were analyzed by Multidimensional Scaling. The averages were on the same measurement scale, that is, the input was matrix conditional, the level of measurement was supposed to be ratio, and the analysis was metric. The MDS (*Alscal*) algorithm was used. We chose a three-dimensional solution since goodness-of-fit measures (RSQ = .98810 and stress = .11956) were good and the dimensions were interpretable.

We compared the chord variables listed in Table 2 with the coordinates of each chord on the dimensions, and the highest statistically significant correlations were used for interpreting the dimensions. Dimension 1 was correlated with General register, $r(10) = -.805$, $p = .002$, and Highest note, $r(10) = -.704$, $p = .011$, and was labelled *Pitch register* with high chords at the negative end and the low chords at the positive end of the dimension. Dimension 2 was correlated with Spectral Centroid, $r(10) = .751$, $p = .005$, Spectral Rolloff, $r(10) = .750$, $p = .005$, and brightness, $r(10) = .717$, $p = .009$, and the dimension was labelled *Timbral brightness*. Dimension 3 correlated with *duration*, $r(10) = .777$, $p = .003$, and was

labelled accordingly. The three-dimensional configuration can be seen in Figure 1, and the chords at the ends of Dimensions 1 and 2 are shown in Figure 2.

Figure 1

Three-dimensional solution of the similarity-estimation task

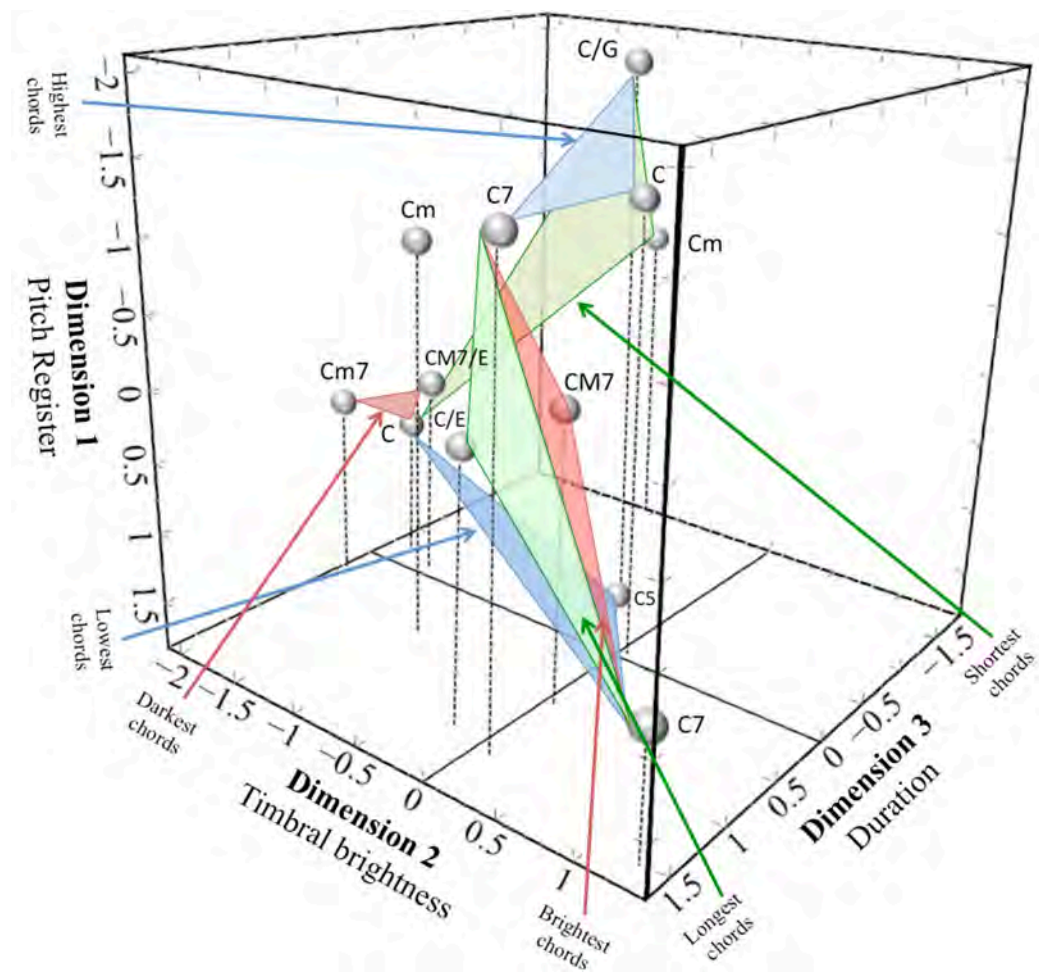
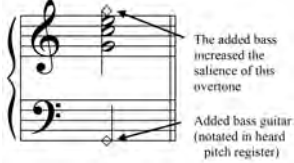


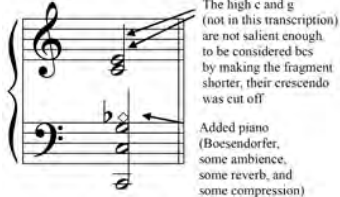

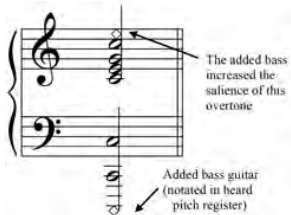
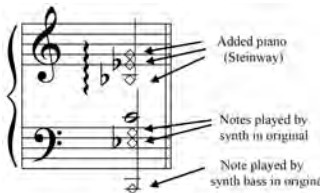
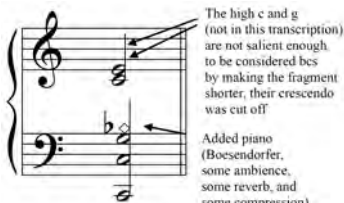
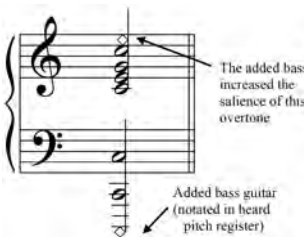
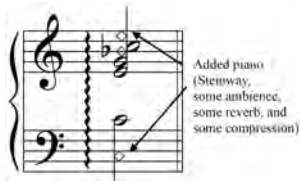

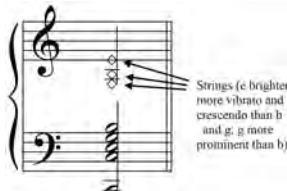


Figure 2*Chords at the ends of Dimensions 1 and 2*

Dim 1	Highest chords (negative end)		Lowest chords (positive end)
-2.0267		1.7671	
-1.3667		0.8640	
-1.2652		0.7894	
Dim 2	Darkest chords (negative end)		Brightest chords (positive end)
-2.0474		1.3594	
-1.9078		0.0797	
-1.4058		0.0245	

Analysis of the identification of 20 songs from their opening chords

The analysis of the dimensions affecting similarity estimation was used as a starting point when we analyzed the variables affecting the identification of songs. It takes only a short time to estimate the similarity of two chords played in quick succession, so it makes sense that variables related to surface-level characteristics of chords such as their register, timbral brightness, and duration might be the most salient and important. Although these characteristics may also play a role in song identification, other more abstract features may also contribute to the activation of auditory long-term memories. For this reason, we added two variables related to the harmonic characteristics of the chords (Chord-type frequency of occurrence and Harmonic similarity score), and one variable that was related to the songs used in the experiment (Year of release). One more variable, Attack time, was also included in the analysis, due to its relationship with timbral brightness. Descriptions of these variables and the rationale for including them in this analysis can be found in Appendix C.

We continued the analysis by running a PCA. For the set of ten variables, the KMO test (.596) was acceptable, and Bartlett's test ($p < .001$) indicated that a PCA could be conducted. This revealed four components that were used in a stepwise regression analysis to see if they predicted correct identification of songs from their opening chords. The components were as follows (see also Appendix C):

(1) *Brightness* consisted of four variables, three of which directly measured spectral brightness: Brightness, Spectral centroid, and Spectral rolloff (see Appendix C). The fourth, Attack time, was related to brightness, because both sound characteristics are similarly affected by keystroke force.

(2) *Familiarity* consisted of the variables Year of release, Chord-type frequency of occurrence, and Duration. The year of release may relate to participants' familiarity with the song (e.g., Krumhansl & Zupnick, 2013) and to the age of the participant. The Chord-type

frequency of occurrence was meant to represent the relative amount of exposure that participants had to different chord types and can be understood as an approximate measure of chord-type familiarity (for further information, see Appendix C). Finally, Duration had a clear relationship with Chord-type frequency of occurrence. The nature of this relationship, as well as other aspects of the variables constituting Component 2, will be discussed later in this section.

(3) *Pitch register* consisted of the variables General register and Highest note.

(4) *Harmonic similarity* consisted of only one variable.

Since the component matrix was varimax rotated, there was no multicollinearity between the components. Additionally, the residuals were normally distributed and linear, and the Durbin-Watson statistic (1.780; see Table 5) was acceptable. The analysis revealed that, taken together, the four components explained 43.6% of the variance, although only Components 1 and 2 had statistically significant effects, while the contribution of Component 3 was marginally significant (see Table 5).

Table 5

Model summary of the regression analysis

Model Summary										
Model	<i>R</i>	<i>R</i> ²	Adjusted <i>R</i> ²	Std. Error of the Estimate	<i>R</i> ² Change	Change Statistics				Durbin-Watson
						<i>F</i> Change	<i>df</i> 1	<i>df</i> 2	<i>p</i>	
1	.466	.217	.173	15.0%	.217	4.897	1	18	.038	
2	.666	.444	.378	13.0%	.227	6.929	1	17	.017	
3	.736	.542	.456	12.2%	.098	3.437	1	16	.082	
4	.745	.555	.436	12.4%	.013	.425	1	15	.524	1.780

Note. Model 1: (Constant), Brightness

Model 2: (Constant), Brightness, Familiarity

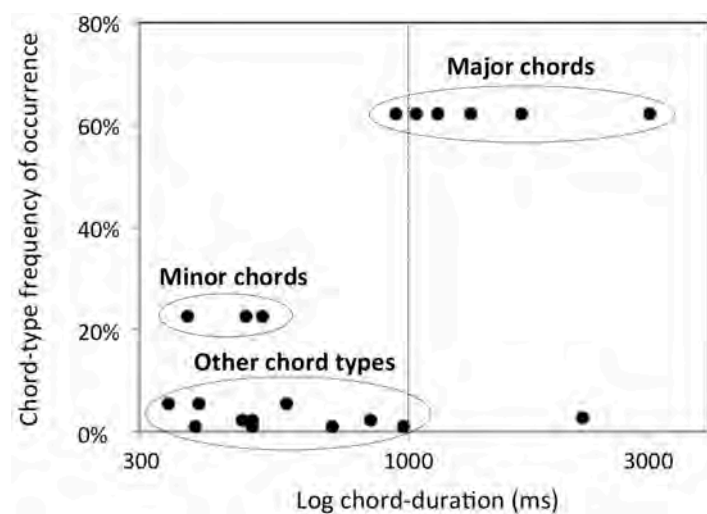
Model 3: (Constant), Brightness, Familiarity, Pitch register

Model 4: (Constant), Brightness, Familiarity, Pitch register, Harmonic similarity

The coefficients (no table included) showed positive effects for *Brightness* and *Pitch register*, indicating that higher scores of Brightness and Pitch register were related to higher percentages of songs correctly identified. The coefficient for *Familiarity* showed a negative effect and can be explained in terms of two complementary tendencies. First, higher identification rates were linked with more recent year of release (negatively loaded within the component). The relationship between year of release, participants' age, and familiarity with specific songs will be considered in the Discussion. Second, higher identification rates were linked to lower chord-type frequency of occurrence (positively loaded within the component) and thus potentially to higher chord-type distinctiveness. We also noticed that higher identification rates were linked to shorter duration, a relationship that may be subordinated to the strong connection between duration and chord-type frequency of occurrence in our chord set. Figure 3 shows that, generally, major chords are used more often than minor chords, and other chord types less often. It also shows that major chords generally had longer durations than other chords.

Figure 3

Distribution of chord-type frequency of occurrence and chord duration



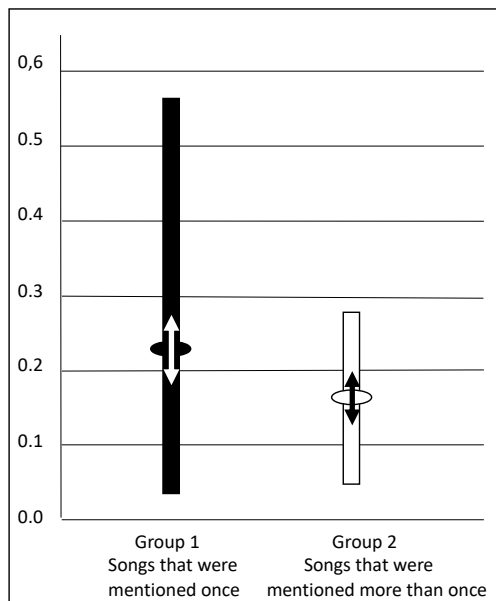
Analysis of chords from other songs

When the participants were asked to name the song from the opening chord, they sometimes named songs other than the ones from which the chords were taken. The total number of these responses was 215 (about two per participant, on average). Some of them corresponded to songs the participant had mentioned in earlier phases of the session, and so were excluded. The resulting subset of other-chord responses included 124 songs.

We calculated the average difference between each pair constituted by a test stimulus and the opening chord of the song the participant mentioned using the ten variables that were used in the previous analyses. On a scale from 0 to 1, the average differences varied between .046 and .560, indicating that some responses were very similar to the test stimuli with regard to the ten variables. We further divided the responses into two categories: songs that were named only once ($n = 113$), and songs that were named more than once ($n = 11$). When we compared the two groups, we noticed that the differences between the test chord and the response were generally smaller for the chords that were named more than once (see Figure 4). This is an understandable result and indicates that there were measurable similarities between the test chords and the alternative responses. Since the group with chords mentioned more than once was small, the difference between the groups was only marginally significant (see Table 6).

Figure 4

The range (bars), the average (ellipse), and the standard deviation (arrow) of the average differences between test chords and the opening chord of the response for the two groups of songs

**Table 6**

Statistics for the ANOVA analysis between songs that were mentioned once ($n = 113$) and songs that were mentioned more than once ($n = 11$)

	ANOVA				
	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>
Between Groups	.033	1	.033	3.536	.062
Within Groups	1.123	122	.009		
Total	1.155	123			

Discussion

The present study investigated how the timbral and non-timbral characteristics of brief excerpts of music can affect both similarity estimation and the identification of songs from those excerpts. Stimuli consisted exclusively of the first complete piano or piano-driven block chords with which popular songs begin. Using this type of stimulus reduced the number of variables by avoiding the potential effect of vocals, the position of the excerpt within the song, melodic and rhythmic patterns, and by reducing the effect of instrumentation.

Some of the variables investigated in our analysis of the song-identification data were chosen on the basis of their perceptual salience in a chord similarity-estimation task. However, these two experimental tasks tend to rely on different memory systems. Whereas long-term memory is critical for the identification of well-known songs, and even though long-term familiarity can also have some effect on similarity ratings (Lou, 2017), our similarity-estimation task was likely to rely more on sensory auditory memory. Sensory auditory memory is believed to be the memory system that enables listeners to compare just-heard sounds (Massaro, 1972; Crowder, 1982). The fact that brightness had a strong effect on both song identification and chord-similarity ratings, but pitch register - an aspect of chord voicing - had less effect on identification than similarities, may suggest that memory for timbre lasts longer than memory for pitch register. However, more studies of the relationship between long-term memory and pitch register and brightness are needed to clarify the connection between timbre and pitch register.

We found no evidence that participant variables affected song identification from opening chords. This was consistent with previous research that has found no effect of musical training on ability to identify the genre of very brief musical excerpts (Mace et al., 2012). However, it should also be noted that we recruited participants who identified themselves as being “often good at naming pop songs,” and participants who do not fall into that category might be less able to identify songs from opening chords.

The analysis of the effect of chord variables on song identification showed that songs were identified more often when chords were bright. Additionally, the familiarity of both song and chord-type had an effect on identification. The results also revealed a marginally significant effect of pitch register on song identification. These results are generally consistent with those of previous research (Krumhansl, 2010; McKellar & Cohen, 2015, Schellenberg et al., 1999), since they provide additional evidence that timbral and non-timbral characteristics

of the stimuli, as well as the year of release of the song, can affect rapid song identification. We discuss the effect of brightness and familiarity in more detail below.

Brightness

Songs in our experiment may have been identified more often from brighter chords because brightness has a tendency to attract listeners' attention. Estimates of similarity showed that brightness was one of the most perceptually salient aspects of our stimuli. This result is consistent with previous research on timbre (Caclin et al., 2005; Grey & Gordon, 1978; Iverson & Krumhansl, 1993; Lakatos, 2000; McAdams et al., 1995). Since listeners are more likely to pay attention to brighter sound events due to their perceptual salience (Huang & Elhilali, 2017), brighter sound events are also likely to be better encoded in auditory long-term memory and listened to more attentively than less bright sound events during identification tasks. This memory advantage for brighter sounds may also help explain why brighter songs tend to be more commercially successful (Interiano et al., 2018). However, the relationship between commercial success and brightness, as well as the effect of brightness in our study, could also be simply the byproduct of the better reproducibility of higher frequencies. Ambient noise (Anzenbacher et al., 2013) and the size of loudspeakers (Arora et al., 2007; Arora et al., 2006; Larsen & Aarts, 2002) may often prevent listeners from hearing some of the lower frequencies in a recording. Although our headphones/earphones test controlled for some of the potential effects of headphones/earphones quality and ambient noise, we had no control over the conditions under which participants had heard the target song throughout their lives. However, this theory would also predict that it is easier to identify songs from higher-pitched excerpts, but this was not supported by our results.

The effect of brightness and attack time are related to each other via perceived keystroke force. In acoustic pianos and electronic instruments modeled on acoustic pianos, performers control loudness by controlling keystroke force (Askenfelt & Jansson, 1992;

Kinoshita et al., 2007; Oku & Furuya, 2017). In these types of instrument, an increase in the vertical force on a key produces tones that are not only louder, but that are also brighter (Palmer & Brown, 1989) and have a shorter attack time (Askenfelt & Jansson, 1993). Loudness was equalized in the present study, but differences in terms of brightness and rise attack may have led participants to perceive the chords as having been produced by different amounts of keystroke force (Fabiani & Friberg, 2011). It is possible that the effects of brightness and attack time could be a byproduct of the effect of keystroke force on the encoding and retrieval of piano and piano-driven block chords in auditory long-term memory.

It could also be argued that chords that are perceived as having been produced using force are also likely to attract listeners' attention, and therefore to be encoded in auditory long-term memory and attended to during identification tasks. To our knowledge, there are no studies of the attention-grabbing properties of keystroke force, but there is some evidence that perceived higher energy levels tend to correlate with the perceived familiarity of short musical excerpts (Filipic et al., 2010) and the popularity of songs (Interiano et al., 2018; North et al., 2019). It is also possible to explain the effect of high energy levels on familiarity ratings and popularity by the direct connection between high energy levels and perceived emotional arousal in auditory stimuli (Nordström & Laukka, 2019), in addition to the well-documented effects of perceived emotional arousal on the encoding (Christianson & Loftus, 1991; Talmi, 2013) and retention of memories (Sharot & Phelps, 2004; Yonelinas & Ritchey, 2015).

Familiarity

As stated, familiarity seemed to influence song identification in two different ways. First, we found an effect for Year of release, which may have been a consequence of more recent songs being fresher in participants' memories. Since most of our participants were born after 1980, the memory advantage for more recent songs was likely a combination of a recency effect (Spivack et al., 2019) and the fact that songs released when individuals were

teenagers or early adults tend to be recognized more easily (Platz et al., 2015; Rathbone et al., 2017). Although familiarity with specific songs might be affected by participants' age, we found no effect of age when analyzing the effect of participant variables.

Second, we found an effect for chord-type frequency of occurrence, which may have been the result of the relative amount of exposure that the participants had to different types of chords. Greater familiarity with a chord type had a negative effect on identification, which can be explained by the number of songs known by the participants that start with a certain type of chord: the more songs that compete for memory activation, the more difficult the identification becomes.

Chord duration was one variable constituting Component 2 (*Familiarity*). In our study, songs were easier to identify from shorter chords. This finding is counterintuitive and contradicts the results of Schellenberg et al. (1999), who showed that song identification is easier from 200-ms than 100-ms excerpts, and with Krumhansl (2010), who showed it is easier from 400-ms than 300-ms excerpts. As stated, our stimuli were always composed using the complete first piano or piano-driven block chord, but no more, while the excerpts from Krumhansl (2010) could include parts of the second chord and additional melodic and rhythmic information as well. Although those differences could have explained the absence of an effect of duration on our experiment, they do not explain a negative effect. It is likely that this negative direction has to do with the relationship between chord-type frequency of occurrence and duration in our set of stimuli. As was shown in the results, major chords were longer than the non-major chords. Considering that our stimuli were taken from real music, it is possible that there is a general tendency in mainstream popular music for major chords to be longer than other chords. The connection between duration and the pitch content of block chords and their role in song identification is a topic that deserves further study.

Conclusions

Our study showed that timbral (e.g., instrumentation and brightness) and non-timbral features (e.g., chord type and voicing) can affect both similarity estimation and song identification. Although our study was not designed to compare auditory long-term memory to auditory sensory memory, results from our two experiments suggest that these two memory systems may differ in terms of the importance they give to attributes such as pitch register and chord-type frequency of occurrence, at least in the context of single piano and piano-driven block chords.

We found strong evidence for the effect of brightness and two familiarity variables (Year of release and Chord-type frequency of occurrence) on song identification from the opening chord. The effect of other variables was less clear. The effect of brightness may relate to perceptual salience, issues of sound reproduction, and/or the frequency of occurrence of songs that start with timbrally bright block chords. Our study highlights the potential connections between brightness, keystroke force, perceived energy level, and perceived emotional arousal, and suggests that their shared effect on long-term memory may be explained by their tendency easily to grab listeners' attention.

This study was the first to demonstrate that songs can be identified from their opening chord, and to find some evidence that brightness, and variables related to chord type can play a role in the identification of songs from brief excerpts of music. Future studies could investigate whether those variables can also play a role in the identification of songs from brief excerpts taken from other parts of a song. Such studies will be needed to deepen our understanding of long-term memory for timbral and non-timbral features, and how such memories may be activated and influence listeners' experiences.

References

- Alluri, V., & Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Perception*, 27(3), 223-242.
- ANSI. (2007). ANSI S3. 4-2007: American National Standard: Procedure for the computation of loudness of steady sounds. *Acoustical Society of America*.
- Anzenbacher, C., Reuter, C., & Oehler, M. (2013). Sound quality vs. sound identity. The perceptibility of audio logos under everyday conditions of transmission and reception. *Audio Branding Academy Yearbook 2013*, 14, 127–135.
- Armitage, J., & Eerola, T. (2020). Reaction time data in music cognition: Comparison of pilot data from lab, crowdsourced, and convenience Web samples. *Frontiers in Psychology* 10. doi: 10.3389/fpsyg.2019.02883
- Arora, M., Chung, C., Jang, S., & Moon, H. G. (2007). Enhanced bass reinforcement algorithm for small-sized transducer. Paper presented at the 122th Audio Engineering Society Convention (Vienna, Austria, May 5–8, 2007).
- Arora, M., Jang, S., & Moon, H. (2006). Low complexity virtual bass enhancement algorithm for portable multimedia device. Paper presented at the 29th Audio Engineering Society Conference (Seoul, Korea, September 2–4, 2006).
- Arthurs, Y., Beeston, A. V., & Timmers, R. (2018). Perception of isolated chords: Examining frequency of occurrence, instrumental timbre, acoustic descriptors and musical training. *Psychology of Music*, 46(5), 662–681.
- Askenfelt, A., & Jansson, E. V. (1992). On vibration sensation and finger touch in stringed instrument playing. *Music Perception: An Interdisciplinary Journal*, 9(3), 311–349.
- Askenfelt, A., & Jansson, E. V. (1993). From touch to string vibrations. III: String motion and spectra. *The Journal of the Acoustical Society of America*, 93(4), 2181–2196.
- Bachorik, J. P., Bangert, M., Loui, P., Larke, K., Berger, J., Rowe, R., & Schlaug, G. (2009). Emotion in motion: Investigating the time-course of emotional judgments of musical stimuli. *Music Perception: An Interdisciplinary Journal*, 26(4), 355–364.

- Beal, A. L. (1985). The skill of recognizing musical structures. *Memory & Cognition*, 13(5), 405–412.
- Belfi, A. M., Kasdan, A., Rowland, J., Vessel, E. A., Starr, G. G., & Poeppel, D. (2018). Rapid timing of musical aesthetic judgments. *Journal of Experimental Psychology: General*, 147(10), 1531–1543.
- 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 & Emotion*, 19(8), 1113–1139.
- Bigand, E., Gérard, Y., & Molin, P. (2009). The contribution of local features to familiarity judgments in music. *Annals of the New York Academy of Sciences*, 1169(1), 234–244.
- Bosnjak, M., & Tuten, T. L. (2003). Prepaid and promised incentives in Web surveys: An experiment. *Social Science Computer Review*, 21, 208–217.
- Bowling, D. L., Purves, D., & Gill, K. Z. (2018). Vocal similarity predicts the relative attraction of musical chords. *Proceedings of the National Academy of Sciences*, 115(1), 216–221.
- Breebaart, J. (2017). No correlation between headphone frequency response and retail price. *The Journal of the Acoustical Society of America*, 141(6), EL526–EL530.
- Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *The Journal of the Acoustical Society of America*, 118(1), 471–482.
- Cho, J. L., Hall, M. D., & Pastore, R. E. (1991). Normalization process in the human auditory system. *The Journal of the Acoustical Society of America*, 89(4B), 1988–1988.
- Cho, T., & Bello, J. P. (2013). On the relative importance of individual components of chord recognition systems. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(2), 477–492.
- Christianson, S. Å., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition & Emotion*, 5(2), 81–108.
- Costa, M., Ricci Bitti, P. E., & Bonfiglioli, L. (2000). Psychological connotations of harmonic musical intervals. *Psychology of Music*, 28(1), 4–22.

- Crowder, R. G. (1982). A common basis for auditory sensory storage in perception and immediate memory. *Perception & Psychophysics*, 31, 477–483. doi: 10.3758/BF03204857
- Dalla Bella, S., Peretz, I., & Aronoff, N. (2003). Time course of melody recognition: A gating paradigm study. *Perception & Psychophysics*, 65(7), 1019-1028.
- Di Giorgi, B., Zanoni, M., Sarti, A., & Tubaro, S. (2013). Automatic chord recognition based on the probabilistic modeling of diatonic modal harmony. In *nDS'13; Proceedings of the 8th International Workshop on Multidimensional Systems* (pp. 1–6). VDE.
- Difallah, D., Filatova, E., & Ipeirotis, P. (2018). Demographics and dynamics of mechanical turk workers. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 135–143. ACM.
- Fabiani, M., & Friberg, A. (2011). Influence of pitch, loudness, and timbre on the perception of instrument dynamics. *The Journal of the Acoustical Society of America*, 130(4), EL193–EL199.
- Filipic, S., Tillmann, B., & Bigand, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review*, 17(3), 335–341.
- Gjerdingen, R. O., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research*, 37(2), 93–100.
- Goto, M., Hashiguchi, H., Nishimura, T., & Oka, R. (2002). RWC music database: Popular, classical and jazz music databases. In *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR)* (pp. 287–288).
- Grey, J. M., & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *The Journal of the Acoustical Society of America*, 63(5), 1493–1500.
- Grisey, G., & Fineberg, J. (2000). Did you say spectral? *Contemporary Music Review*, 19(3), 1–3.
- Hajda, J. M., Kendall, R. A., Carterette, E. C., & Harshberger, M. L. (1997). Methodological issues in timbre research. In I. Deliege & J. A. Sloboda (Eds.), *Perception and cognition of music* (pp. 253–306). Hove: Psychology Press.

- Hasegawa, R. (2019). Timbre as Harmony—Harmony as Timbre. In E. I. Dolan & A. Rehding (Eds.) *The Oxford handbook of timbre (online)*. Oxford University Press. DOI: 10.1093/oxfordhb/9780190637224.013.11
- Halpern, A. R., & Darwin, C. J. (1982). Duration discrimination in a series of rhythmic events. *Perception & Psychophysics*, 31(1), 86–89.
- Hébert, S., & Peretz, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory & Cognition*, 25, 518–533.
- Huang, N., & Elhilali, M. (2017). Auditory salience using natural soundscapes. *The Journal of the Acoustical Society of America*, 141(3), 2163–2176.
- Interiano, M., Kazemi, K., Wang, L., Yang, J., Yu, Z., & Komarova, N. L. (2018). Musical trends and predictability of success in contemporary songs in and out of the top charts. *Royal Society Open Science*, 5(5), 171–274.
- Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre. *The Journal of the Acoustical Society of America*, 94(5), 2595–2603.
- 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(6), 1797–1813.
- Kinoshita, H., Furuya, S., Aoki, T., & Altenmüller, E. (2007). Loudness control in pianists as exemplified in keystroke force measurements on different touches. *The Journal of the Acoustical Society of America*, 121(5), 2959–2969.
- Kolchinsky, A., Dhande, N., Park, K., & Ahn, Y. Y. (2017). The minor fall, the major lift: inferring emotional valence of musical chords through lyrics. *Royal Society Open Science*, 4: 170952.
- Krumhansl, C. L. (2010). Plink: “Thin slices” of music. *Music Perception: An Interdisciplinary Journal*, 27(5), 337–354.
- Krumhansl, C. L., & Zupnick J. A. (2013). Cascading reminiscence bumps in popular music. *Psychological Science*, 24(10), 2057–2068. <https://doi.org/10.1177/0956797613486486>
PMID: 24006129

- Kuusi, T., Jimenez, I., & Schulkind, M. (2021) Revisiting the effect of listener and musical factors on the identification of music from chord progressions. In L. Suurpää & J. Ojala (Eds.) *Musical performance in context: A Festschrift in celebration of doctoral education at the Sibelius Academy*. DocMus Research Publications 17. Helsinki: Sibelius Academy.
- Labuschagne, I. B., & Hanekom, J. J. (2013). Preparation of stimuli for timbre perception studies. *The Journal of the Acoustical Society of America*, 134(3), 2256–2267.
- Lahdelma, I., & Eerola, T. (2016). Single chords convey distinct emotional qualities to both naïve and expert listeners. *Psychology of Music*, 44(1), 37-54.
- Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception & Psychophysics*, 62(7), 1426–1439.
- Larsen, E., & Aarts, R. M. (2002). Reproducing low-pitched signals through small loudspeakers. *Journal of the Audio Engineering Society*, 50(3), 147–164
- Lartillot, O., Toivainen, P., & Eerola, T. (2008). A matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme & R. Decker (Eds.), *Data analysis, machine learning and applications* (pp. 261-268). Berlin, Heidelberg: Springer.
- Layman, S. L., & Dowling, W. J. (2018). Did you hear the vocalist? Differences in processing between short segments of familiar and unfamiliar music. *Music Perception: An Interdisciplinary Journal*, 35(5), 607–621.
- London, J. (2011). Musicology. In T. Gracyk & A. Kania (Eds.), *The Routledge companion to philosophy and music* (pp. 495-505). New York, NY: Taylor & Francis Group.
- Lou, L. I. (2017). *Timbre memory, familiarity and dissimilarity* (Doctoral dissertation, McGill University Libraries).
- Mace, S. T., Wagoner, C. L., Teachout, D. J., & Hodges, D. A. (2012). Genre identification of very brief musical excerpts. *Psychology of Music*, 40(1), 112–128.
- Massaro, D. W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, 79, 124–145. doi: 10.1037/h0032264

- Mauch, M., & Dixon, S. (2010). Approximate note transcription for the improved identification of difficult chords. In *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR)* (pp. 135–140).
- McAdams, S., & Siedenburg, K. (2019). Perception and Cognition of Musical Timbre. In P. J. Rentfrow & D. J. Levitin (Eds.), *Foundations of music psychology: Theory and research* (pp. 71–120), Cambridge, MA: MIT Press.
- McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, 58(3), 177–192.
- McKellar, J. L., & Cohen, A. (2015). Identification of thin slices of music by university students in PEI. *Canadian Acoustics*, 43(3).
- Nadar, C. R., Abeßer, J., & Grollmisch, S. (2018). Towards CNN-based acoustic modeling of seventh chords for automatic chord recognition. Paper presented in at the 16th Sound & Music Computing Conference (SMC) (Malaga, Spain, April 2019).
- Nordström, H., & Laukka, P. (2019). The time course of emotion recognition in speech and music. *The Journal of the Acoustical Society of America*, 145(5), 3058–3074.
- North, A. C., Krause, A. E., Sheridan, L. P., & Ritchie, D. (2019). Comparison of popular music in the United States and the United Kingdom: Computerized analysis of 42,714 pieces. *Psychology of Music*. DOI <https://doi.org/10.1177/0305735619830185>
- O'Neil, K. M., & Penrod, S. D. (2001). Methodological variables in Web-based research that may affect results: Sample type, monetary incentives, and personal information. *Behavior Research Methods, Instruments, & Computers*, 33, 226–233.
- O'Neil, K. M., Penrod, S. D., & Bornstein, B. H. (2003). Web-based research: Methodological variables' effects on dropout and sample characteristics. *Behavior Research Methods, Instruments, & Computers*, 35, 217–236.
- Oku, T., & Furuya, S. (2017). Skilful force control in expert pianists. *Experimental Brain Research*, 235(5), 1603–1615.

- Palmer, C., & Brown, J. (1989). Influence of hammer velocity in piano sound. *The Journal of the Acoustical Society of America*, 85(S1), S142–S142.
- Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68(2), 111–141.
- Platz, F., Kopiez, R., Hasselhorn, J., & Wolf, A. (2015). The impact of song-specific age and affective qualities of popular songs on music-evoked autobiographical memories (MEAMs). *Musicae Scientiae*, 19(4), 327–349.
- Pressnitzer, D., & McAdams, S. (2000). Acoustics, psychoacoustics and spectral music. *Contemporary Music Review*, 19(2), 33–59.
- Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception: An Interdisciplinary Journal*, 24(1), 37–48.
- Rathbone, C. J., O’connor, A. R., & Moulin, C. J. (2017). The tracks of my years: Personal significance contributes to the reminiscence bump. *Memory & Cognition*, 45(1), 137–150.
- Rose, F. (1996). Introduction to the pitch organization of French spectral music. *Perspectives of New Music*, 34(2), 6–39.
- Schellenberg, E. G., Iverson, P., & Mckinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review*, 6(4), 641–646.
- Schulkind, M. D. (2000). Perceptual interference decays over short unfilled intervals. *Memory & Cognition*, 28(6), 949–956.
- Schulkind, M. D. (2002). Feature modulation search: A novel memory search model that extends the perceptual interference effect to musical stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(2), 346–352.
- Schulkind, M. D. (2004a). Conceptual and perceptual information both influence melody identification. *Memory & Cognition*, 32(5), 841–851.
- Schulkind, M. D. (2004b). Serial processing in melody identification and the organization of musical semantic memory. *Perception & Psychophysics*, 66(8), 1351–1362.

- Schulkind, M. D., & Davis, S. J. (2013). The cohort model of melody identification: Evaluating primacy and similarity. *Psychology of Music*, 41(4), 422–439.
- Schulkind, M. D., Posner, R. J., & Rubin, D. C. (2003). Musical features that facilitate melody identification: How do you know it's "your" song when they finally play it? *Music Perception: An Interdisciplinary Journal*, 21(2), 217–249.
- Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience*, 4(3), 294–306.
- Siedenburg, K., & Müllensiefen, D. (2017). Modeling timbre similarity of short music clips. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00639>
- Spivack, S., Philibotte, S. J., Spilka, N. H., Passman, I. J., & Wallisch, P. (2019). Who remembers the Beatles? The collective memory for popular music. *PloS ONE*, 14(2), e0210066.
- Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096–1104. doi: 10.3758/BRM.42.4.1096
- Stoet, G. (2017). PsyToolkit: A novel Web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24–31.
- Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science*, 22(6), 430–436.
- Tekman, H. G., & Bharucha, J. J. (1992). Time course of chord priming. *Perception & Psychophysics*, 51(1), 33–39.
- The Beatles. (1993). *The Beatles: Complete scores*. Transcribed by Tetsuya Fujita, Yuji Hagino, Hajime Kubo and Goro Sato. London: Hal Leonard Publishing Corporation.
- Tuten, T. L., Galesic, M., & Bosnjak, M. (2004). Effects of immediate versus delayed notification of prize draw results on response behavior in Web surveys – An experiment. *Social Science Computer Review*, 22 (3), 377–384.
- Tzanetakis, G. & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5), 293–302.

- Ueda, Y., Uchiyama, Y., Nishimoto, T., Ono, N., & Sagayama, S. (2010). HMM-based approach for automatic chord detection using refined acoustic features. In *2010 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 5518–5521). IEEE.
- Virtala, P., Huotilainen, M., Lilja, E., Ojala, J., & Tervaniemi, M. (2018). Distortion and Western music chord processing: An ERP study of musicians and nonmusicians. *Music Perception: An Interdisciplinary Journal*, 35(3), 315–331.
- Virtala, P., Huotilainen, M., Partanen, E., & Tervaniemi, M. (2014). Musicianship facilitates the processing of Western music chords – An ERP and behavioral study. *Neuropsychologia*, 61, 247–258. doi: 10.1016/j.neuropsychologia.2014.06.028
- Wallmark, Z., Iacoboni, M., Deblieck, C., & Kendall, R. A. (2018). Embodied listening and timbre: Perceptual, acoustical, and neural correlates. *Music Perception: An Interdisciplinary Journal*, 35(3), 332–363.
- Weiss, M. W., Schellenberg, E. G., Peng, C., & Trehub, S. E. (2019). Contextual Distinctiveness Affects the Memory Advantage for Vocal Melodies. *Auditory Perception & Cognition*, 2(1), 47–66.
- Wersényi, G. (2010). On the measurement and evaluation of bass enhanced in-ear phones. In *Proceedings of the 20th International Congress on Acoustics, ICA, Sydney, Australia*.
https://www.acoustics.asn.au/conference_proceedings/ICA2010/cdrom-ICA2010/index.htm
- Wycisk, Y., Kopiez R., & Wolf, A. (2018). Control of Headphone and Loudspeaker Characteristics in Online. Paper presented at the ICMPC15/ESCOM10 (Graz, Austria, July 23–28).
- Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: an emotional binding account. *Trends in Cognitive Sciences*, 19(5), 259–267.

Appendices

Appendix A

20 chords used for the song-identification task

Opening chords with piano as the only instrument

John Legend, "All of Me" (2017) 720 ms

Daniel Powter, "Bad Day" (2005) 860 ms

John Lennon, "Imagine" (1971) 474 ms

Bill Withers, "Lean On Me" (1972) 1693 ms

Sara Bareilles, "Love Song" (2007) 570 ms

Two lowest notes softer

Elton John, "Rocket Man" (1972) 408 ms

Wiz Khalifa, "See You Again" (2015) 390 ms

Billy Joel, "She's Always a Woman" (1977) 1162 ms

Hozier, "Take Me to Church" (2014) 375 ms

Coldplay, "The Scientist" (2002) 350 ms

Two lowest notes softer

Opening chords with piano as the main instrument

Sarah McLachlan, "Angel" (1997) 3052 ms

Timbaland, "Apologize" feat. OneRepublic (2006) 532 ms

David Bowie, "Changes" (1971) 2230 ms

Train, "Drops of Jupiter" (2001) 375 ms

Elton John, "Goodbye Yellow Brick Road" (1973) 940 ms

Soft noise likely from gently blowing air through a bamboo flute

Synth

Synth bass

Strings

Two lowest notes softer

Bass guitar

Chicago, "Hard to Say I'm Sorry" (1982) 1050 ms

Lionel Richie, "Hello" (1983) 1008 ms

Mariah Carey, "Hero" (1993) 507 ms

The Police, "King of Pain" (1983) 495 ms

Sam Smith, "Stay with Me" (2014) 720 ms

soft synth

loud synth

Synth strings

Low synth or bass guitar

Bass drum

Appendix B

List of songs used in the main experiment

Song title	Artist/band	ID 15-sec*	ID Chord**
Songs starting with piano as the only instrument			
“All of Me”	John Legend	67%	49%
“Bad Day”	Daniel Powter	65%	35%
“Imagine”	John Lennon	87%	10%
“Lean on Me”	Bill Withers	94%	16%
“Love Song”	Sara Bareilles	47%	70%
“Rocket Man”	Elton John	83%	15%
“See You Again”	Wiz Khalifa (feat. Charlie Puth)	63%	17%
“She's Always a Woman”	Billy Joel	56%	11%
“Take me to Church”	Hozier	57%	22%
“The Scientist”	Coldplay	51%	28%
Songs starting with a piano as the main instrument			
“Angel”	Sarah McLachlan	59%	21%
“Apologize”	Timbaland (feat. OneRepublic)	70%	32%
“Changes”	David Bowie	57%	22%
“Drops of Jupiter”	Train	71%	37%
“Goodbye Yellow Brick Road”	Elton John	63%	6%
“Hard to Say I'm Sorry”	Chicago	49%	24%
“Hello”	Lionel Richie	68%	54%
“Hero”	Mariah Carey	59%	41%
“King of Pain”	The Police	49%	12%
“Stay with Me”	Sam Smith	73%	37%

Note. * Identification from 15-sec excerpts for all 93 participants.

** Identification from opening chord in main experiment. These percentages do not consider all responses. The procedure for calculating these percentages is explained in detail in the Results section.

Appendix C

Chord variables used in the analysis of the data from the identification of songs from chords

Timbral Characteristics

Most of the values for the timbral variables were obtained by using version 1.7.2 of the MIRtoolbox developed by Lartillot et al. (2008). Before the analysis was carried out, all stimuli were shortened to 300 ms. Spectral analyses of the 300-ms versions and the full duration versions were different in terms of brightness, spectral centroid, spectral rolloff, spectral entropy, spectral flatness, envelope flatness, inharmonicity, and various mfccs (mel-frequency cepstral coefficients that describe the distribution of the spectral energy). We

decided to use the spectral analysis of the 300-ms versions for our main analysis for several reasons. First, we wanted to analyze duration as a variable independent of other spectral variables, and analyzing spectral variables from full durations generates values that are influenced by duration. Second, the onset of a piano tone is usually louder than the rest of the tone. As a consequence, listeners are likely to encode that part of the sound in memory better because it is easier to hear. The third reason has to do with the fact that piano performers control the loudness and the brightness of piano tones by increasing the force of their keystrokes. Since performers have relatively little control over the loudness and brightness of piano tones once the key has been struck, the onset of a piano tone can be expected to provide most of the essential information about the loudness and timbral character of the piano tone (i.e., the loudness and timbral character of the rest of the sustained sound of a piano tone can be predicted by hearing its attack). Finally, Krumhansl (2010) and the results of our pilot studies have shown that excerpts as short as 300 ms can provide sufficient timbral information for song identification.

Brightness: This feature is measured as the ratio of the signal energy that is above a certain frequency threshold, as proposed by Juslin (2000). We used the default threshold of 1500 Hz, which has been often adopted when analyzing brightness in excerpts of popular music (e.g., Alluri & Toiviainen, 2010; Wallmark, et al., 2018).

Spectral centroid and spectral rolloff: These features are statistical descriptors of the spectral shape of the sound. Spectral centroid is a measure of the central moments of the spectral distribution. The rolloff threshold is defined as the frequency below which a defined fraction of the total spectral energy is contained. For this fraction, the default value of 85% (as proposed by Tzanetakis & Cook, 2002) was used.

Attack time: This variable, also known as rise time, is an estimate of the length of the attack phase within the signal. The beginning and end of the attack phase of the sound

event were detected using the Derivative method implemented in the MIRtoolbox. Although attack time did not have a clear effect on the chord-similarity estimations, we included *attack time* in the analysis of the data from the song-identification task because attack time has been shown to be one of the most salient attributes of timbre (Labuschagne & Hanekom, 2013), and the strong tendency for attack time to be related to timbral brightness, a perceptually salient feature in participants' similarity estimates, due to both attack time (Askenfelt & Jansson, 1993) and timbral brightness (Palmer & Brown, 1989) being signals of keystroke force.

Chord-type Features

Descriptions of the chord-type features and chord-voicing characteristics were based on transcriptions made by two of the authors and one very experienced musician, who all worked together on the transcription of each chord. Aural assessments were also complemented by automated pitch analysis run in Sonic Visualiser, Anthem Score, and Melodyne.

Harmonic similarity: Similarity between different chord types and the harmonic series as calculated by Bowling et al. (2018). This variable is conceptually related, but not highly correlated, to the “chordal third” value. Harmonic similarity also closely relates to the concept of inharmonicity. However, inharmonicity is influenced by both timbre and pitch as it is measured directly from the audio signal, whereas harmonic similarity only takes into account pitch.

Chord-type frequency of occurrence: Average frequency of occurrence of different chord types (e.g., Major, minor, Major seventh, etc.) in Western pop and rock. These values were calculated by averaging information about thousands of songs from the following sources:

Source	Repository	Songs	Method of chord identification	Numbers of chords
Kolchinsky et al., 2017	Guitar tabs from ultimate-guitar.com	123,837 songs from various styles of popular music	Expert and non-expert human annotation	924,418
Nadar et al., 2018	RWC music database (Goto et al., 2002)	100 songs from the RWC Popular Music database	Expert human annotation	110,331
Nadar et al., 2018	Di Giorgi et al., 2013	26 Robbie Williams songs	Expert human annotation	25,569
Nadar et al., 2018	Isophonics dataset reference annotations	20 Queen songs	Expert human annotation	20,610
Nadar et al., 2018	Isophonics dataset reference annotations	128 Beatles songs	Expert human annotation	86,868
Arthurs et al., 2018	Guitar chord symbols from The Beatles: Complete scores (The Beatles, 1993).	The Beatles' 30 best-selling UK hits.	Expert human annotation	ca 3,000
HookTheory.com (June 1, 2019)	HookTheory.com	12,245 tabs from various styles of popular music	Expert and non-expert human annotation	38,759*

Note. * The number of chords in HookTheory.com was calculated by counting all the occurrences of a specific chord on a tab as being one chord

Chord-voicing Characteristics

General register: percentage of notes represented by C4 or higher.

Highest note: these notes were tabulated using MIDI note numbers.

Duration: tabulated in ms.

Song Information

Year of release: we included this variable in our analysis to detect any effect of the participants' familiarity with the song that was not already controlled for by our method of screening participants and our criterion for categorizing responses as a missed identification, which required that participants identified the song from a 15-sec recording (10-sec excerpts in the pilots).

Appendix D

Questionnaire about musical background

Main instrument

1. What musical instrument do you play best, including voice? _____
Write "NA" if you do not sing and have never played an instrument.
2. How long have you played this instrument? Years ____ If less than a year please indicate months: ____
3. At what age did you first start practicing a musical instrument? For this question, "practicing" refers to at least one practice session a week during a period of two months or longer. If you have never practiced an instrument weekly for at least two months, please write "0." In this question, voice is not considered a musical instrument. ____

Playing by ear

4. Which of these options best describes your experience playing an instrument by ear (i.e., without ever having seen any notation for that piece of music)?
 - a. I have never played an instrument by ear.
 - b. I have sometimes played by ear and my level of playing by ear is beginner.
 - c. I have played by ear for less than 3 years and my level of playing by ear is intermediate.
 - d. I have played by ear for more than 3 years and my level of playing by ear is intermediate.
 - e. I have played by ear for more than 5 years and my level of playing by ear is advanced.

Playing piano

5. Which of these options best describes your experience playing piano?
 - a. I have never played piano.
 - b. I have played some piano and my level of playing is beginner.
 - c. I have played piano for less than 3 years and my level of playing is intermediate.
 - d. I have played piano for more than 3 years and my level of playing is intermediate.
 - e. I have played piano for more than 5 years and my level of playing is advanced.

Playing guitar

6. Which of these options best describes your experience playing guitar?
 - a. I have never played guitar.
 - b. I have played some guitar and my level of playing is beginner.
 - c. I have played guitar for less than 3 years and my level of playing is intermediate.
 - d. I have played guitar for more than 3 years and my level of playing is intermediate.
 - e. I have played guitar for more than 5 years and my level of playing is advanced.

Playing popular music on the piano

7. Which of these options best describes your experience playing pop, rock, or other types of popular music on the piano?
 - a. I have never played that type of music on the piano.
 - b. I have played that type of music on the piano and my level of playing that type of music is beginner.
 - c. I have played that type of music on the piano for less than 3 years and my level of playing that type of music is intermediate.
 - d. I have played that type of music on the piano for more than 3 years and my level of playing that type of music is intermediate.
 - e. I have played that type of music on the piano for more than 5 years and my level of playing that type of music is advanced.