

## CHOOSING THE RIGHT TUNE: A REVIEW OF MUSIC STIMULI USED IN EMOTION RESEARCH

---

LINDSAY A. WARRENBURG  
*The Ohio State University*

**WHEN DESIGNING A NEW STUDY REGARDING HOW** music can portray and elicit emotion, one of the most crucial design decisions involves choosing the best stimuli. Every researcher must find musical samples that are able to capture an emotional state, are appropriate lengths, and have minimal potential for biasing participants. Researchers have often utilized musical excerpts that have previously been used by other scholars, but the appropriate musical choices depend on the specific goals of the study in question and will likely change among various research designs. The intention of this paper is to examine how musical stimuli have been selected in a sample of 306 research articles dating from 1928 through 2018. Analyses are presented regarding the designated emotions, how the stimuli were selected, the durations of the stimuli, whether the stimuli are excerpts from a longer work, and whether the passages have been used in studies about perceived or induced emotion. The results suggest that the literature relies on nine emotional terms, focuses more on perceived emotion than on induced emotion, and contains mostly short musical stimuli. I suggest that some of the inconclusive results from previous reviews may be due to the inconsistent use of emotion terms throughout the music community.

*Received: November 15, 2018, accepted October 17, 2019.*

**Key words:** emotion, music, stimuli, review, database

---

**S**INCE THE 1930s, SCHOLARS HAVE investigated how music is able to portray and elicit emotion in listeners, such as *sadness*, *happiness*, and *fear* (Hevner, 1935, 1936, 1937; Sherman, 1928). Researchers in the areas of developmental psychology, cognitive psychology, and music theory and composition have each offered unique contributions to the field. Indeed, substantial progress has been made in this area of music cognition such that several conjectures for how music is able to convey and evoke emotions in listeners have been thoroughly examined.

---

Because there exist thousands of articles and books published with regard to music and emotion, review articles of these studies are important resources to researchers interested in this area. To my knowledge, there are six major reviews that have been published on the topic of music and emotion:

- (1) Västfjäll (2002) examined how mood can be evoked from music-listening experiences. In particular, he discussed the Musical Mood Induction Procedure (MMIP), a popular technique in the psychology literature around the 1980s and 1990s. The MMIP is a procedure where participants are told to listen to music with the hopes that the music will change their experienced mood state (typically *depressed* and *elated*). Västfjäll examined the methodology and the stimuli used in this procedure in around 40 studies to give a comprehensive picture of music's emotional evocation properties. For example, Västfjäll described which types of genres of music were represented in MMIP studies, noted which pieces of music have been used in multiple studies, and mentioned how some of the researchers selected their stimuli. The descriptions of the stimuli were summaries of the MMIP stimuli writ large; features of individual musical stimuli were not discussed.
- (2) Juslin and Laukka (2003) surveyed literature regarding how musical performance expresses emotion and its comparison to communication of emotions in vocal speech. The authors provided one of the few meta-analyses on music and emotion to date, as they compare findings across 145 studies from the speech and music literatures. The authors found that an evolutionary perspective on emotional speech was consistent with some of the findings of music's expressive power. For example, sad music tends to mirror sad speech, and therefore a single explanatory code can be used to describe how music and speech are expressive of emotion. The authors provided valuable insight into acoustic cues in vocal expression and music performance. Their review found that sad speech and music often contain features

like slow speech rate/tempo, low voice intensity/sound level, and little F0/pitch variability. The features examined by Juslin and Laukka are undeniably important, but their review does not relate features to individual musical stimuli.

- (3) Gabrielsson and Lindström (2010) surveyed how musical performance and composition tend to be associated with emotional expression in music. The authors meticulously chronicled the effects of features like tempo/speed, intensity/loudness, pitch, and timbre/spectrum. They discussed that more typical “musical” features (like mode and melody) often depend on the musical context, rather than acting in isolation. Furthermore, they discussed how more extreme examples of musical features tend to result in emotional affects, whereas milder examples of these musical features are unable to affect the listener’s emotional response. The purpose of this review, like the other studies mentioned above, was to summarize findings across many published papers. Therefore, the chapter does not contain information about musical and contextual features of individual stimuli.
- (4) Eerola and Vuoskoski (2013) reviewed the research approaches and emotional models that have been used to study music and emotion writ large. The authors found that most studies used some variation of a discrete or dimensional model, that many studies rely on Western Art Music samples, and that there have been increasing variants in methodological design, including self-report, biological approaches, developmental approaches, cross-cultural approaches, and music analysis approaches. The section on music stimuli examined the number of musical passages used in each study, the durations of these passages, the genres represented, how the stimuli were selected, and familiarity of the participants with the musical passages. The current study follows up on Eerola and Vuoskoski’s work. In addition to summarizing these features in a large body of work, the current study presents a database where these features are denoted for individual stimuli.
- (5) Schubert (2013) compared different loci of emotions—felt emotion and expressed emotion in 19 studies. He found that expressed emotions tend to have higher ratings of emotional intensity than experienced emotion and that demographic characteristics and musical structure influenced both perceived and felt emotions. Importantly, he called for a more systematic use of emotional terminology in the literature. Schubert described features

of musical stimuli pertinent to the focus of his review, such as the genre of musical passages, the familiarity of participants with the musical selections, and the duration of stimuli. Once again, there is no explicit mention of how these features correspond to individual studies or stimuli, as that is beyond the scope of Schubert’s review.

- (6) Garrido (2014) summarized the distinction between mood and emotion, how researchers utilize excerpts in these two affective states, and the different time frames used for measuring affective states. This is one of the only papers that differentiates mood and emotion, which is an important distinction that should be noted in future research questions. Garrido’s review focused on the specific mood or emotion investigated by other researchers in 95 articles, as well as how they defined the conceptual model used in designing the study. The features of individual musical stimuli are not discussed.

Each of these reviews covers music and emotion with a different perspective and all are useful for researchers investigating music and emotion. Despite the importance of these studies, however, no systematic analysis of the actual musical stimuli used in emotion research has been published. Although several of these review papers discuss musical stimuli, they concentrate more on *how* stimuli (broadly) might affect emotion or mood, rather than characterizing features of the stimuli themselves. The current study aims to supplement these major reviews with a survey of the types of stimuli used in previous research of musical expression and evocation and a discussion of emotional language that clarifies how emotional terminology has been utilized in the literature. As the ability to identify and utilize appropriate musical samples is a task that all researchers using musical stimuli face, it is crucial to identify general trends of stimulus choices. The validity, reliability, and power of the study to find an effect are all directly affected by stimulus choice.

Accordingly, the central aim of this paper is to conduct a literature review based on a publicly-available database of over 22,000 musical stimuli, the Previously-Used Musical Stimuli (PUMS) database (Warrenburg, 2019). I aim to present readers with a summary of how musical emotional stimuli have been used in the past, with the intention of facilitating the choice of stimuli in future studies. In order to do this, I will describe features like how past researchers have selected their stimuli and the types of emotion that have been studied, as well as how stimuli characteristics vary across induced emotion and

perceived emotion studies. Furthermore, I examine how stimuli have been used from 1928 to 2018, showing that there are certain trends in music research that come and go. By examining the relationship among methodological factors and types of musical stimuli used, I highlight the importance that is given by study authors to specific combinations of methodological and musical forms. This type of analysis is needed because it will help future researchers carefully select musical stimuli in future studies and will provide context when examining the methodology of other published research.

The review does not cover the difference between mood and emotion, how music compares with speech, or performance factors that imbue music with emotional meaning. It should be noted that the majority of the stimuli are from Western music and therefore, cross-cultural comparisons must be left for future research.

## Methods

### AN OVERVIEW OF THE PUMS DATABASE

The reviews listed above tend to concentrate on *how* stimuli (broadly) might affect emotion or mood, rather than summarizing features of the stimuli themselves. There does not appear to be any published database that identifies which musical stimuli have been used in previous emotion-related studies. Rather, researchers have often summarized stimuli characteristics without providing references to the stimuli themselves. This section describes the publicly available database of Previously-Used Musical Stimuli (PUMS), which summarizes the types of musical passages that have been used in studies of emotional expression and evocation from 1928 through 2018 (Warrenburg, 2019). The PUMS database also provides a resource for researchers designing new emotion-related studies.

The PUMS database contains reference information for 22,417 musical stimuli that were used in 306 studies on music and emotion. In order to create this database, the author conducted a systematic analysis of the literature, where she examined 654 papers on music and emotion according to various inclusion criteria. For each of the stimuli in the PUMS database, the author reported characteristics such as its designated emotion or mood, duration of the sample, whether it was an excerpt or a full work, whether it was used in a study of induced or perceived emotion, type (or style) of music, the selection method used by the original researchers, the publication date of the original study, and the familiarity of the participants with the stimulus. When information was available in the original papers, the composer, name, performer, track number, and

measure numbers (or duration markings) of the work were recorded.

### CONDUCTING A LITERATURE REVIEW

A systematic review of the coded features of the 22,417 stimuli in the PUMS database was conducted. The aim of the literature review is to provide insight into questions researchers might have when choosing suitable stimuli for their studies. The characteristics detailed in the PUMS database are focused on how and why music researchers selected certain stimuli for their studies. Audio features of the stimuli, such as the fundamental frequency, tempo, and pitch contour are not analyzed. Future research should examine the stimuli in the PUMS database for these features, as this information would prove useful to future researchers studying music and emotion.

In the creation of the PUMS database, I noticed that not all researchers provide information about which musical stimuli they use in their studies. Depending on the aims of the study and the audience of the journal, it may not be relevant to include features like the specific measure numbers of their stimuli. Other researchers did not denote which methodology they used to select their stimuli or how familiar the participants tended to be with the musical samples. In constructing the PUMS database, then, it was not always possible to find every feature for a particular stimulus. Accordingly, when describing the results of the literature review, the number of stimuli for which the information was available is always noted. For example, in the section on perceived versus induced emotion, we will see that 74% (13,895) of the studies focused on perceived emotion, 24% (4,607) focused on induced emotion, and only 2% (316) focused on both perceived and induced emotion. In total, then, there was information about the study locus (perceived/induced) for 18,818 musical stimuli, but this information was not available for 3,599 musical stimuli.

In the systematic review of the PUMS database, I will first present descriptive statistics regarding the coded features, such as the most popular emotion studied in musical stimuli and the proportion of studies that regard induced versus perceived emotion. Next, I will describe the interactions among these features. For example, I will explore whether studies about *sad* music tend to use participant-selected musical stimuli more often than studies about *happy* music. Finally, I will explore some longitudinal trends of studies using emotional music stimuli. The longitudinal analysis will help us see whether certain emotions are more represented in music studies in some decades and whether methods like pilot testing are used equally across the years.

The goal of the literature review is to provide insight into common questions that may arise when researchers are choosing the most suitable stimuli for their studies, such as the following:

- (1) Should one make use of extant musical recordings or create more controlled stimuli that are composed specifically for the study?
- (2) What is the optimum duration of a stimulus; should stimuli be long or short? Additionally, when using existing excerpts, should they be excerpts from longer works or the full works?
- (3) Does the style or genre of the music matter?
- (4) Should one avoid music that is familiar to the participants?
- (5) Finally, of critical importance is the question of how to operationalize the mood or emotion conveyed or evoked by a particular passage. How have other researchers selected emotional stimuli in past studies? For example, how have other researchers operationalized an emotion like “sadness”?

## Results: Analyzing Features of Emotion-Related Musical Stimuli

### INDIVIDUAL FEATURES

**Familiarity.** In the PUMS database, 55% (2,557) of the stimuli were explicitly used because they were unfamiliar to participants, 35% (1,619) of the stimuli were used because they were familiar to participants, and 10% (482) of the stimuli were familiar to some participants and unfamiliar to other participants.

**Excerpt/full work designation.** The review indicates that 55% (10,832) of the musical stimuli were musical excerpts or passages and 45% (8,873) were full musical works.

**Duration.** The musical stimuli ranged from two seconds in duration to 45 minutes in duration. Fifty percent (4,238) of the stimuli were less than 30 seconds in duration, 28% (2,381) of the stimuli were between 30 seconds and 59 seconds in duration, 9% (769) were between one and two minutes in duration, 2% (165) were between two and three minutes in duration, and 8% (704) were between three and four minutes in duration. Stimuli four minutes and over (in one-minute increments) each represented 1% or less of the sample.

**Induced or perceived.** Seventy-four percent (13,895) of the stimuli in the PUMS database were part of studies that focused on perceived emotion, 24% (4,607) of the stimuli were part of studies that focused on induced emotion, and only 2% (316) of the stimuli were part of

studies that focused on both perceived and induced emotion.

The fact that there are many more stimuli used in perceived emotion studies than in induced emotion studies could be a testament to the inclusion of a few Music Emotion Recognition (MER) studies, which typically contain large numbers of stimuli. Excluding some of these larger studies (i.e., Eerola, 2011; Schuller, Dorfner, & Rigoll, 2010; Weninger, Eyben, Schuller, Mortillaro, & Scherer, 2013), the number of emotional stimuli used in perceived and induced emotion studies are closer to equal, with 4,564 stimuli used in induced emotion studies and 8,163 stimuli used in perceived emotion studies.

**Emotional terms.** The PUMS database includes stimuli that were categorized as evoking or representing 114 emotions. Of the 114 emotions listed, 23% (2,013) of the stimuli were deemed *sad*, 19% (1,614) were *happy*, 10% (869) were *angry*, and 9% (775) were *relaxed*. Furthermore, 8% (731) pertain to *chills/pleasure* and 4% (332) pertain to *fear*. *Chills (no associated pleasure)* made up 4% (318) of the sample, *peace* (172) and *groove* (148) each made up 2% of the sample, and *negative valence* (119) made up 1% of the stimuli.

These findings suggest that over half of music-related emotion stimuli have been designated as representing or inducing three emotions: *sadness*, *happiness*, and *anger*. In addition, only nine emotional terms (*sad*, *happy*, *anger*, *relaxed*, *chills/pleasure*, *fear*, *chills*, *peace*, and *groove*) have been related to musical stimuli more than 1% of the time. The term *tender* appeared about 1% of the time, which is surprising given that *tenderness* or *tender/peaceful* are sometimes considered to be a primary emotion of music (e.g., Horn & Huron, 2015; Juslin, 2013b). A complete list of the emotion terms is given in Table 1.

**Selection method for emotional stimuli.** In the PUMS database, the way the researchers selected their emotional music stimuli is broken into six categories: *previous studies*, *experimenter/expert chosen*, *pilot tested*, *composed for study*, *participant chosen*, and *professionals asked to play/express emotion*. The analysis of the PUMS database suggests that 37% (8,046) of stimuli were chosen because they were used in past studies, whereas 46% (10,051) of the stimuli were selected because of experimenter or expert opinion. Seven percent (1,547) of the stimuli were participant selected and only 3% (744) were pilot tested. An additional 1% (206) of the stimuli were specifically composed for the study, and 1% (175) were performed by professionals (e.g., drummers, singers, pianists, guitarists) who were asked to play or express certain emotions. For the list of

TABLE 1. Complete List of Emotion Terms in a Review of 306 Studies on Music and Emotion

Term	Number	Term	Number	Term	Number
Sad	2013	Scary	12	Comforting, Relaxing	2
Happy	1614	Spiritual	12	Exaggerated	2
Anger	869	Pleasant, Joyful	10	Active; Agitated	1
Relaxed	775	Serene	10	Active; Alert	1
Chills, Pleasure	731	Calm	9	Agitated; Anger	1
Fear	332	Content	7	Agitated; Restless	1
Chills	318	Disappointment, Remorse, Frustration	7	Anger; Hate	1
Peace	172	Distress, Sad	7	Anger; Joy; Sad	1
Groove	148	Fear, Startle, Nervous	7	Arousal	1
Negative Valence	119	Gratification, Satisfaction	7	Arousal; Negative Valence	1
Positive Valence	116	Gratitude	7	Bored; Listless	1
Negative Valence, Low Arousal	100	Hate, Disgust	7	Bored; Unstimulating	1
Neutral	94	Hope, Joy, Happy, Gloating, Surprise, Excited	7	Chills; Happy	1
Positive Valence, High Arousal	93	Lonely	7	Chills; Sad	1
Negative Valence, High Arousal	90	Love	7	Depressed; Sad	1
Positive Valence, Low Arousal	89	Pity	7	Energizing	1
Tender	86	Pleasant; Trivial	7	Exalted	1
Fear; Threatening	56	Pride, Admiration	7	Excitative	1
Happy; Sad	53	Relief	7	Excited; Festive	1
Joy	49	Resentment, Anger	7	Expressive	1
Pleasant	46	Shame, Reproach	7	Happy; Angry/Agitated	1
Arousal; Potency; Valence	45	Surprise	6	Happy; Fear; Sad	1
Low Arousal	30	Exuberant	5	Irritation	1
Negative Tension	30	Negative/Positive Valence, Low/High Arousal	5	Joy; Pleasant	1
Positive Energy	30	Active	4	Negative/Positive Valence, Low Arousal	1
Positive Tension	30	Joy; Lively	4	Normal	1
Depressed	27	Negative Valence, Low/High Arousal	4	Nostalgia	1
Sad, Beautiful	27	Pleasant; Trivial; Unpleasant	4	Pain	1
Arousal; Valence	21	Positive Valence, Low/High Arousal	4	Peace; Relaxing	1
Lively	21	Sedated	4	Pleasant; Unpleasant	1
Solemn	20	Tension	4	Restrained	1
Lump in Throat; Tears	19	Agitated	3	Serene; Tranquil	1
Anxiety	16	Anger; Fear	3	Solemn; Unpleasant	1
Elated	15	Disgust	3	Sorrow	1
Unpleasant	14	Excited	3	Standard	1
Affection	12	Neutral; Sad	3	Stimulating	1
Humor	12	Amusement	2	Tranquil	1
Longing	12	Arousing	2	Unsettling	1

stimuli that were chosen because they were used in past studies, the reader is referred to Table 2.

*Style of music and composers.* In the PUMS database, 57% (10,705) of stimuli were classified as popular music, 10% (1,910) were classified as Western Art Music, and 7% (1,397) were classified as film or film-like music (see Table 3). The large focus on popular music might be attributed to several large studies, each with over 2,000 stimuli. Three of these studies (Eerola, 2011; Schuller et al., 2010; Weninger et al., 2013) were

removed, in order to observe the influence of popular music without these large studies, which, in total, accounted for 8,448 stimuli (7,944 of which were labeled as popular music). Without these three studies, the influence of popular music remains, with 4,081 stimuli, compared to Western Art Music (WAM), which contained 1,564 stimuli and film music, which contained 807 stimuli.

In addition to counting the number of stimuli that belong to a certain style of music, another way of

TABLE 2. List of "Past Studies" Where Experimenters Gathered Their Stimuli

Source	Number of Stimuli	Source	Number of Stimuli
Albersnagel, 1998	6	<i>Last.FM</i>	3024
Altenmüller et al., 2002	160	Levitin & Menon, 2003	16
Bachorik et al., 2009	288	Liégeois-Chauvel et al., 2014	40
Bigand et al., 2001	24	<i>Magnatune.com</i>	800
Bigand et al., 2005	89	Martin, 1990	10
Billboard Hot 100 Year End Charts	8	McFarland, 1985	2
Blood et al., 1999	1	Nagel et al., 2007	7
Bower & Mayer, 1989	2	Nater et al., 2006	3
Darling, 1982	9	Nyklíček et al., 1997	25
Clark, 1983	9	Omar et al., 2010	40
Clark & Teasdale, 1985	7	Panksepp, 1995	1
Costa-Giomi et al., 2008	2	Peretz et al., 1998	248
Dalla Bella et al., 2001	18	Peretz et al., 2001	42
Dibben, 2004	8	Ramos et al., 1996	1
Dolgin & Adelson, 1990	1	Roy et al., 2008	30
Eerola & Vuoskoski, 2011	338	Schaffrath, 1995	12
Filipic et al., 2010	40	Schellenberg, 1996	2
Flores-Gutiérrez, 2001	2	Schuller et al., 2010	5296
Gentile, 1998	10	Spackman et al., 2005	4
Gosselin et al., 2005	168	Speck et al., 2011	240
Gosselin et al., 2006	18	Smithsonian Collection of Classic Country Music, 1981	21
Gomez & Danuser, 2004	16	Song et al., 2012	80
Grewe et al., 2007	10	Song et al., 2013	80
Hevner, 1935	10	Stratton & Zalanowski, 1991	10
Hevner, 1936	10	Thomson Music Index Demo	108
Hunter et al., 2008	60	Vieillard et al., 2008	198
Iwanaga et al., 1996	2	Vines et al., 2006	1
Iwanaga & Tsukamoto, 1997	2	Vuoskoski & Eerola, 2012	3
Juslin, 1997c	10	Wanderley, 2002	3
Kawakami et al., 2013	2	Wu & Jeng, 2006	76
Koelsch et al., 2006	28	Wu & Jeng, 2008	37
Kreutz et al., 2003	10	Yang & Chen, 2011	60
Krumhansl, 1997	3	Yang et al., 2006	390
		Yang, Su, Lin, & Chen, 2007	60

Note: Not all of studies labeled which study they used to choose their stimuli. The sources in italics were labeled as "experimenter/expert chosen." For the complete reference, please see the References.

examining the styles of musical samples is to investigate the number of studies that made use of these style classifications. In the PUMS database, 34% (103) of research studies used Western Art Music as their stimuli, compared to 9% (27) of research studies that used film music and 9% (28) of studies that used popular music.

In the PUMS database, the ten most popular composers were Mozart (168), Beethoven (135), J. S. Bach (129), Bernard Bouchard (112), Chopin (91), Schumann (69), Mendelssohn (62), Schubert (57), Brahms (49), and Marsha Bauman (48). Bernard Bouchard and Marsha Bauman composed works specifically for studies on music and emotion. By comparison, the ten most popular composers in music theory journals in rank order (according to Duinker & Léveillé Gauvin, 2017) are Schoenberg, Beethoven, Brahms, J. S. Bach, Mozart,

Webern, Stravinsky, Wagner, Schubert, and Chopin. Although there is overlap between these two lists, the music theory journals tend to concentrate more on later composers, whereas musical emotion studies tend to rely on WAM composers only through the Romantic period.

#### INTERACTIONS

*Duration and perceived/induced.* A chi-square test suggests that the distributions of durations (collapsed into < 30 s, 30 s to 1 min, > 1 min) were used differently across induced, perceived, and both induced/perceived conditions,  $X^2 = 3150.20$ ,  $df = 4$ ,  $p < .01$ . The most common duration of musical stimuli across both induced and perceived studies was less than 30 seconds. However, whereas perceived studies had used 65% of their stimuli being less than 30 seconds (3,178) and 98%

TABLE 3. List of Musical Styles or Genres Used in 306 Studies on Music and Emotion

Style	Stimuli	Style	Stimuli	Style	Stimuli
Popular Music	10705	Western Art Music	1910	Film Music	1397
Chinese Music	1241	MIDI	766	Keyboard Instruments	461
Electronic	363	Rock	289	Jazz	193
Folk Music	151	Natural Sounds	136	Voice	131
Guitar	88	Rock & Pop	69	Indian Music	59
Soul	58	Drum Beat	48	Ambient	42
Mafa	40	Funk/Soul	36	Japanese Music	34
Russian Music	32	Hip-Hop	28	Violin	27
Country	25	Flute	24	World	24
Bassoon	20	Improvisation	20	Saxophone	20
Latin	18	Children's Music	16	TV Programs	16
Blues	15	Atonal	13	Religious	13
Heavy Metal	12	Hymn	12	New Age	11
Rap & Hip-Hop	11	Alternative	10	Sentograph	10
Pygmy Music	8	R&B	7	Traditional Music	6
Gospel	5	Techno	5	Death Metal	4
Foreign	4	Indie Pop/Rock	4	Disco	3
Trumpet	3	African Music	2	Background Music	2
Dance Music	2	Minimal Music	2	Post-Rock	2
Randomly Generated	2	Krygystani Music	2	Navaho Indian Music	2
Avant-garde	1	Bossa Nova	1	Computer Game Music	1
Emo Rock	1	Gothic Rock	1	House	1
Industrial Rock	1	Metal	1	Modern Tonal	1
Psychedelic Trance	1	Romantic	1	Single Tones	1
Trance	1				

of their stimuli being less than one minute (4,757), induced studies only used 27% of their stimuli as less than 30 seconds (776) and 45% of their stimuli as less than one minute (1,314). Almost all of the perceived studies used stimuli under three minutes in length; however, induced studies included about 29% of stimuli longer than 3 minutes (833). The exact durations for induced/perceived musical stimuli is summarized shown in Table 4.

*Duration and style.* I also examined the durations of stimuli with regard to genre (Western Art Music, Film or Film-like music, Popular music, and Jazz). A chi-square test comparing these genres to the collapsed time periods (< 30 s, 30 s to 1 min, > 1 min) suggests that the distributions of times were used differently across styles of music,  $X^2 = 411.72$ ,  $df = 6$ ,  $p < .01$ . One finding is that while film music, popular music, and jazz music almost exclusively relied on stimuli under 30 seconds,

TABLE 4. Comparison of Induced-perceived Stimuli with Their Duration in 306 Studies on Music and Emotion

<i>Both</i>		<i>Induced</i>		<i>Perceived</i>	
Time Length	Number	Time Length	Number	Time Length	Number
< 30 s	38	< 30 s	776	< 30 s	3178
30 s - 1 min	193	30 s - 1 min	538	30 s - 1 min	1579
1 - 2 min	30	1 - 2 min	639	1 - 2 min	65
2 - 3 min	1	2 - 3 min	102	2 - 3 min	38
3 - 4 min	7	3 - 4 min	694	3 - 4 min	2
4 - 5 min	4	4 - 5 min	5	4 - 5 min	2
5 - 6 min	2	5 - 6 min	14	5 - 6 min	1
6 - 7 min	1	6 - 7 min	41	6 - 7 min	1
7 - 8 min	27	7 - 8 min	32	7 - 8 min	0
8 - 9 min	0	8 - 9 min	11	8 - 9 min	0
9 - 10 min	0	9 - 10 min	8	9 - 10 min	1
10 min +	0	10 min +	28	10 min +	1

72% for film (577), 75% for jazz (125), and 80% for popular music (941), Western Art Music had more of a spread of durations used for musical passages. Only 48% (556) of the Western Art Music stimuli were less than 30 seconds, with 26% (300) lasting between 30 seconds and 1 minute, and 26% lasting longer than one minute.

*Perceived/induced and selection method.* A chi-square test suggests that the way researchers selected emotional music stimuli varied across perceived, induced, and both perceived/induced studies,  $X^2 = 4881.89$ ,  $df = 12$ ,  $p < .01$ . Whereas studies of perceived emotion primarily relied on stimuli from previous studies (51%, 7,009) and experimenter/expert chosen stimuli (38%, 5,187), studies of induced emotion relied more on experimenter/expert chosen stimuli (35%, 1,533), and participant chosen stimuli (30%, 1,343), rather than on previous studies, (19%, 826). One interesting finding is that over twice as many studies utilized stimuli that had been pilot tested before the main experiment for induced emotion than for perceived emotion (6%, 267 for induced emotion and 3%, 460 for perceived emotion). An additional fact of note is that all of the stimuli that asked professionals to play a work with a certain emotion were studied in a perceived emotion context.

*Perceived/induced and familiarity.* A chi-square test suggests that the familiarity of the participants with the musical stimuli varied across perceived, induced, and both perceived/induced studies,  $X^2 = 792.40$ ,  $df = 4$ ,  $p < .01$ . An analysis of the relationship of the use of familiar music with the methodology of perceived emotion versus induced emotion shows that induced studies use more familiar music (50%, 1,367) than unfamiliar music or music with mixed familiarity, whereas perceived studies tend to rely on unfamiliar music (83%, 1,270). Studies that measured both perceived and induced emotions used familiar music only 22% of the time (28) and used unfamiliar music 78% of the time (100).

*Emotional terminology and familiarity.* Familiarity of the participants with the musical stimuli differed across the kinds of emotion studied (only the nine terms used more than one percent of the time were examined). A chi-square test suggests that the familiarity of the participants with the musical stimuli varied across emotional terms.  $X^2 = 2751.48$ ,  $df = 16$ ,  $p < .01$ . Angry music was primarily unfamiliar to participants (71%, 73). Fearful music, as well, was mainly unfamiliar to participants (85%, 170), with only 9% of fearful stimuli classified as familiar to participants (17). The happy music was familiar to participants only 2% (8) of the time and was unfamiliar to participants 75% (303) of the time. In contrast to happy music, the sad category was primarily

familiar to participants (45% 329), although this could be due in large part to the Taruffi and Koelsch (2014) paper, where participants were asked to select music that made them feel sad. The sad music was unfamiliar to participants 42% of the time (307), while participants had mixed familiarity with the sad music 13% of the time (92). Music associated with chills (92%, 280) and chills/pleasure (100%, 731) were overwhelmingly familiar to the participants, whereas participants had mixed familiarity with all of the groove stimuli (148). Peaceful music was largely unfamiliar to the participants (93%, 149).

*Emotional categories and selection method.* The number of emotions studied appears to vary by the selection method used by the researchers. A chi-square test comparing the nine most common emotion terms with the selection method categories suggests that the selection methods of the emotion varied across different kinds of emotion,  $X^2 = 6202.01$ ,  $df = 48$ ,  $p < .01$ . Only eight emotion terms were used to describe stimuli composed specifically for a study (anger, fear, happy, happy/sad, negative valence, neutral, peace, sad) and only 12 terms were used for stimuli when professionals were asked to play/express emotions (anger, anger/hate, fear, happy, joy, neutral, pain, sad, solemn, sorrow, surprise, tender). Participant-selected stimuli also relied on a limited number of affective terms (anger, chills (no pleasure), chills (pleasureable), disgust, fear, happy, joy, lively, lump in throat/tears, relaxed, sad, surprise).

When stimuli were chosen by the experimenter, however, a broader spectrum of emotional terms were used, with 54 distinct terms appearing in the PUMS corpus for experimenter/expert chosen stimuli and 40 terms appearing for previous studies. The terms for the experimenter/expert chosen stimuli were affection, anger, anxiety, arousal/valence, arousing, chills, chills/happy, chills/sad, comforting/relaxing, content, depressed, energizing, excitative, excited, expressive, exuberant, fear, groove, happy, happy/angry/agitated, happy/sad, humor, irritation, joy, joy/pleasant, longing, low arousal, negative tension, negative valence, negative valence/high arousal, negative valence/low arousal, negative valence/low/high arousal, negative/positive valence/low arousal, negative/positive valence/low/high arousal, neutral, nostalgia, peace, pleasant, positive energy, positive tension, positive valence, positive valence/high arousal, positive valence/low arousal, positive valence, low/high arousal, relaxed, sad, scary, sedated, solemn, spiritual, surprise, tender, tension, and unpleasant.

The terms for previous studies included agitated, anger, anxiety, arousal, arousal/negative valence, arousal/valence, arousing, calm, comforting/relaxing, depressed,

*elated, fear, fear/threatening, happy, happy/sad, joy, low arousal, negative tension, negative valence, negative valence/high arousal, negative valence/low arousal, neutral, peace, pleasant, pleasant/joyful, positive energy, positive tension, positive valence, positive valence/high arousal, positive valence/low arousal, relaxed, sad, sad/beautiful, scary, serene, stimulating, tender, tension, tranquil, and unpleasant.*

*Perceived/induced and emotional categories.* A chi-square test comparing the nine most common emotion terms (minus *groove*, which had no classifications of induced/perceived) with the classifications of stimuli into *induced, perceived, and both induced/perceived* suggests that the emotional terms differed across the type of study design,  $X^2 = 1946.70$ ,  $df = 14$ ,  $p < .01$ . For perceived emotion studies and studies that included both perceived and induced emotion, *sadness* was the most commonly studied emotion, with *happiness* coming in second. Induced emotion studies were topped by *chills/pleasure*, with *sadness* and *happiness* coming next on the list. The induced studies had a total of 58 separate emotional terms, perceived studies had 61 separate emotional terms, and studies of both

induced and perceived emotion contained 16 emotional terms.

Induced studies had 22 emotions with 10 or more counts and 24 categories with only one emotion term. Perceived studies used 27 emotion terms with 10 or more counts and 15 categories with only one emotion term. Studies that examined both perceived and induced emotions had 16 emotion term categories, five of which had 10 or more stimuli present. The reader is referred to Table 5 to see the order of the top 25 terms for the different methodological loci (perceived/induced emotion).

#### LONGITUDINAL TRENDS

*Emotions examined.* A chi-square test comparing the nine most common emotion terms with dates from the 1920s-1990s, 2000s, and 2010s suggests that the emotional terms differed across the decade studied,  $X^2 = 2539.08$ ,  $df = 16$ ,  $p < .01$ . *Sad* is consistently one of the most commonly-studied emotions. *Sadness* represented 15% of the stimuli in the 1980s (after *neutral* and *depressed*), 31% in the 1990s, 17% in the 2000s, and 27% in the 2010s. *Happiness* is commonly the next most

TABLE 5. Top 25 Emotions by Condition (Induced, Perceived, Both) in 306 Studies on Music and Emotion, Excluding Generic Terms Like "Arousal" or "Valence"

<i>Induced</i>		<i>Perceived</i>		<i>Both</i>	
Emotion	Count	Emotion	Count	Emotion	Count
Chills, Pleasure	731	Sad	524	Sad	36
Sad	666	Happy	467	Happy	33
Happy	347	Fear	270	Anger	23
Chills	311	Anger	174	Relaxed	20
Negative Valence	89	Peace	153	Fear	12
Positive Valence	86	Tender	62	Happy; Sad	8
Negative Valence, Low Arousal	80	Fear; Threatening	56	Peace	7
Positive Valence, High Arousal	71	Arousal; Potency; Valence	45	Negative Valence, High Arousal	5
Negative Valence, High Arousal	69	Neutral	34	Negative Valence, Low Arousal	5
Positive Valence, Low Arousal	69	Low Arousal	30	Positive Valence, High Arousal	5
Neutral	60	Negative Tension	30	Positive Valence, Low Arousal	4
Fear	37	Negative Valence	30	Excited	3
Anger	30	Positive Energy	30	Joy	3
Happy; Sad	27	Positive Tension	30	Sedated	3
Tender	24	Positive Valence	30	Negative Valence, Low/High Arousal	1
Depressed	22	Joy	27	Positive Valence, Low/High Arousal	1
Pleasant	21	Sad, Beautiful	27		
Lump in Throat; Tears	19	Lively	21		
Arousal; Valence	18	Solemn	20		
Elated	15	Positive Valence, High Arousal	17		
Unpleasant	14	Negative Valence, High Arousal	16		
Scary	12	Positive Valence, Low Arousal	16		
Calm	9	Negative Valence, Low Arousal	15		
Joy	9	Affection	12		
Peace	9	Humor	12		

commonly studied emotion, with 15% of the research in the 1980s, 23% in the 1990s, 16% in the 2000s, and 20% of the research in the 2010s. The 2000s saw a massive spike in the study of *chills*, with 30% of all emotion studies regarding the presence of *frisson* experiences (whether pleasurable or not). In the 2010s, 2% of all the emotional stimuli studied were *chills*; in the 1990s, 4% of all the stimuli examined were *chills*. The study of *chills* was virtually absent in the other decades. In general, then, most research has focused on *happiness* and *sadness*, which has remained constant throughout the decades.

The number of emotional terms used to describe musical stimuli was 5 in the 1920s, 7 in the 1970s, 9 in the 1980s, 38 in the 1990s, 70 in the 2000s, and 51 in the 2010s (2010–2018). This trend shows a massive proliferation of emotional terms as time passes.

*Induced/perceived locus.* A chi-square test comparing the study design of *induced*, *perceived*, and *both induced/perceived* and the decades 1920s–1990s, 2000s, and 2010s suggests that the type of study design varied across the decade studied,  $X^2 = 2369.83$ ,  $df = 4$ ,  $p < .01$ . This data should be interpreted with caution, as researchers before the 1990s may not have made the distinction between perceived and induced studies clear to their participants. All of the studies examined in the 1920s and 1930s only examined perceived emotion, and all the studies examined in the 1970s only examined induced emotion. In the 1980s, 93% (53) of the studies examined induced emotions. In the 1990s and 2000s, induced and perceived emotion were studied more equally, with 51% (334) studying perceived emotion and 41% (265) studying induced emotion in the 1990s and 50% (2,324) studying perceived emotion and 48% (2,258) studying induced emotion in the 2000s. In the 2010s, there was a huge spike in studying perceived emotion, with 84% studying perceived emotion (11,144).

*Duration.* A chi-square test comparing the duration of the stimuli (< 30 s, 30 s to 1 min, and > 30 s) and the decades 1920s–1990s, 2000s, and 2010s suggests that the duration of the stimuli varied across the decades,  $X^2 = 1775.61$ ,  $df = 4$ ,  $p < .01$ . In the studies from 1928–1999, there was a relatively even split of studies of different durations. Thirty-five percent of the stimuli (145) were less than 30 seconds, 21% of the stimuli (86) were between 30 seconds and one minute in duration, and 45% of the stimuli (188) were greater than one minute in duration. In the 2000s, most of the stimuli were shorter than 30 seconds (64%, 2397), and 29% (1100) were greater than one minute. In the 2010s, this trend continued, with only 13% (533) of the stimuli having durations of longer than one minute.

*Excerpt/full work.* A chi-square test comparing whether a stimulus was an excerpt or full work and the decades 1920s–1990s, 2000s, and 2010s suggests that the type of stimuli varied across the decades,  $X^2 = 2272.90$ ,  $df = 2$ ,  $p < .01$ . From 1928–2009, most musical stimuli were excerpts of longer works (75% or 412 from 1928–1999 and 87% or 3458 in the 2000s). In the 2010s, this trend switched so that there was a roughly even spread of excerpts and full works used (46% or 6962 were excerpts).

*Familiarity.* The familiarity of participants with the stimuli was also examined across time. A chi-square test comparing the familiarity of the participants with the stimuli and the decades 1920s–1990s, 2000s, and 2010s suggests that the type of stimuli varied across the decades,  $X^2 = 958.97$ ,  $df = 4$ ,  $p < .01$ . Before 2000, there was a roughly even spread of participant familiarity with the stimuli (25% or 67 were familiar, 35% or 95 were unfamiliar, and 40% or 107 had mixed familiarity). In the 2000s, the majority of the stimuli were familiar to participants (57% or 1062), while in the 2010s, most of the stimuli were unfamiliar to participants (69% or 1751).

*Style of music.* The style of music (film, jazz, popular, or WAM) was also compared across decades. A chi-square test comparing the style across the decades 1920s–1990s, 2000s, and 2010s suggests that the type of stimuli varied across the decades,  $X^2 = 4383.98$ ,  $df = 6$ ,  $p < .01$ . Whereas 93% (274) of all music studied from 1928–1999 was Western Art Music, the 2000s saw a relatively more even spread of genres, with 25% (384) of film music, 6% (117) of jazz, 27% (419) of popular music, and 40% (616) of WAM. In the 2010s, most of the music studied was popular (87%, 8949).

*Selection method.* Once again, a chi-square test was conducted, comparing the selection methods across the decades 1920s–1990s, 2000s, and 2010s. The results are consistent with the idea that the way researchers selected emotional music stimuli varied across the decades,  $X^2 = 9588.42$ ,  $df = 12$ ,  $p < .01$ . Before 2000, 38% (293) of stimuli were chosen by the experimenter or another expert, 17% (128) of the stimuli were pilot tested, 15% (113) of the stimuli were performed by professionals explicitly asked to express those emotions, 10% (76) of the stimuli were chosen because they were used in a previous study, 9% (67) of the stimuli were selected by the participants, 1% (10) of the stimuli were composed for the study, and the remaining 11% (82) were selected in two or more ways.

In the 2010s, the stimuli were selected in the following ways: 23% (1,042) were also used in previous studies, 22% (990) were participant-selected works, 20% (905) were chosen by experts or the experimenters, 13% (584)

were pilot tested, 2% (91) were composed for the study, 1% (62) were performed by professionals asked to play or express the emotion, and 18% (833) were selected in more than one way.

From 2010-2018, 53% (8,833) of stimuli were chosen by the experimenter or an expert, 42% (6,928) of stimuli were chosen because they were used in previous studies, 3% (490) of stimuli were selected by participants, 1% (105) of stimuli were composed for the study, less than 1% (32) of stimuli were pilot tested, and there were no stimuli used that asked professionals to specifically express an emotion. The remaining 1% (165) of the stimuli were selected in more than one way.

## Discussion

The analysis of the PUMS database shed light on the kinds of musical stimuli that have been used in studies about emotion. Studies use musical stimuli of variable durations, from shorter than 10 seconds to longer than 10 minutes. Musical stimuli were selected in different ways by various researchers, many of whom chose stimuli that had been used in previous studies. The musical stimuli are biased towards Western popular and art music genres, but stimuli were also used from jazz and film genres. In addition to naturalistic recordings of music, some stimuli were synthesized by computers. In this section, I offer some possible interpretations of the findings presented above.

### HOW HAVE RESEARCHERS SELECTED EMOTIONAL MUSIC STIMULI IN THE PAST?

*Familiarity.* Musical familiarity is important because a participant is likely to have certain associations with the stimulus (e.g., personal or movie references) and any emotional response to a stimulus may be colored by these associations (Juslin & Västfjäll, 2008; Schellenberg, Peretz, & Viellard, 2008; Schubert, 2007). When studying experienced emotion, however, often times using familiar music is preferred, as the researcher can better examine how a person's physiology is changing due to intense emotional responses rather than studying (presumably) weaker emotional responses due to experimenter-selected music (Ali & Peynircioğlu, 2010; Panksepp, 1995). The fact that researchers used unfamiliar musical stimuli 55% of the time and used familiar stimuli 35% of the time suggests that researchers' goals differ widely among studies.

The familiarity of participants with the musical stimuli differed across studies of perceived and induced emotion. The analysis showed that 50% of stimuli used in induced emotion studies were familiar to participants,

whereas 83% of stimuli used in perceived emotion studies were unfamiliar to participants. These different trends between induced and perceived emotion may speak to the importance authors place on familiar (participant-selected) stimuli in order to induce an emotion in listeners (Ali & Peynircioğlu, 2010). Certainly, studies examining *frisson* experiences or music-induced *sadness* may wish to rely on familiar stimuli that evoke a specific emotion in the listeners (Panksepp, 1995; Tarrufi & Koelsch, 2014).

*Duration.* We saw that researchers used a wide range of stimuli durations in studies on music and emotion, which is consistent with the idea that researchers have different theoretical or empirical aims. Short musical stimuli allow a researcher to study music that is more likely to be affectively homogeneous—meaning that the music may portray (or evoke) a single emotion. These short stimuli may therefore exhibit internal validity. Recall that 78% of the stimuli in the PUMS database are shorter than one minute in duration.

The use of longer stimuli makes it impossible to know which musical event or sequence of musical events may evoke a particular affective response. However, longer stimuli become important when studying experienced emotion or mood, as it takes more time to induce and maintain an affective response in a listener (e.g., Eerola & Vuoskoski, 2013). Furthermore, listening to longer stimuli may allow the listener to respond in a continuous fashion, where their responses can be measured over time. For example, Schubert (2001) discusses examining a person's continuous response to a piece of music, allowing him to investigate how a participant emotionally responds to different musical events without varying the genre or composer. The analysis indicated that researchers used stimuli longer than 1 minute 22% of the time, with some stimuli lasting more than 40 minutes. Similarly, researchers used musical excerpts 55% of the time, but used full musical works 45% of the time, suggesting that stimulus choice has been driven by a wide variety of research goals.

*Induced or perceived locus.* There is some evidence that perceived emotions are more intense than the corresponding induced emotion (Schubert, 2013). It is also generally accepted that people perceive and recognize emotion through different mechanisms than the way they experience emotion (Huron, 2015; Juslin, 2013a, 2013b; Juslin & Laukka, 2003; Zentner, Grandjean, & Scherer, 2008). The analysis of the PUMS database indicates that 74% of the stimuli were used in studies on perceived emotion. The fact that perceived emotion seems to take precedence over induced emotion could mean that the mechanisms behind perceived emotion

are better studied or understood. Future research might focus on induced emotion, although these studies are notably complicated by the fact that experienced emotion depends on demographic information, personal associations, familiarity with the stimulus, etc. (Juslin & Västfjäll, 2008; Schellenberg et al., 2008; Schubert, 2007). The longitudinal analyses showed that the focus on induced emotion in the 1980s and 1990s has shifted to a focus on perceived emotion in the 2000s and 2010s.

It also appears that there are more uses of valenced terms (*positive valence*, *negative valence*) in the induced studies than in the perceived studies. This trend could be due to the fact that more dimensional models of emotion are used in the induced emotion studies than in the perceived emotion studies (see Eerola & Vuoskoski, 2013, for more on dimensional vs. discrete models of emotion). In short, the kinds of emotions studied in perceived and induced loci matters because it has implications for the mechanisms of emotion that are being investigated by researchers.

The kinds of emotion studied are also important because it could indicate which emotions have been thought to be successfully perceived or induced in music listeners. If an emotion is studied often, it could be that there has been success in evoking/representing that emotion to listeners. For example, there are more stimuli classified as *fear* and *anger* in perceived studies than in induced studies. This finding is consistent with the idea that listeners are less likely to experience these emotions in response to music (Huron, 2006; Zentner et al., 2008). Huron notes that there are certain auditory and musical stressors that can cause *fear* in listeners, like loud sounds, extensity (i.e., many sounds occurring at once, including a rich orchestral texture), low pitches, scream-like sounds, acoustic proximity (e.g., sudden crescendos), and surprise (e.g., sudden musical changes, like changes in harmonies, texture, dynamics, and tempo) (Huron, 2006). In his Suppressed-Fear Theory, Huron hypothesizes that when a person listens to music, initial—and subconscious—feelings of *fear* can be quickly transformed via cognitive appraisals into positive feelings, like *awe*. Namely, an initial *fear* reaction is suppressed and instead the listener will report positive affective responses to *scary* or *fearful* music.

Eerola and Vuoskoski (2013) found that stimuli around 15 seconds are ideal for studies of perceived emotion, whereas stimuli over around 30 seconds are better suited for studies of induced emotion. It takes a certain amount of time to initiate an emotional response and therefore, longer stimuli are more appropriate for the induced emotion locus. Recall that in the PUMS database, musical stimuli in studies on perceived emotion

were shorter than one minute 98% of the time, whereas musical stimuli in studies on induced emotion were shorter than one minute 45% of the time. This finding shows an awareness of the music community that it takes longer to evoke an emotion than to perceive an emotion in music. Nevertheless, some experiments may be using stimuli that may be too short to reliably induce an emotion in listeners and therefore may suffer from weakened statistical power.

*Emotional terminology.* The particular emotional terminology that researchers use to describe the stimuli are important for theoretical and practical reasons. It is possible that multiple stimuli that have been characterized as representing or eliciting the same kind of emotion (e.g., *happy*) could share similar music theoretic or structural characteristics (Juslin & Laukka, 2003; Warrenburg, 2019). Recall that in the PUMS database, there were 114 emotions listed, suggesting that researchers are using some degree of nuance in differentiating emotional classes. One finding was that there are similar numbers of emotional terms in both the perceived and induced studies. This trend is interesting because it has been noted that listener agreement is high only for a few emotions in perceived emotion (e.g., *sadness*, *fear*, *tenderness*, *anger*, *happiness*; Juslin, 2013b), whereas listeners are thought to experience a larger number of emotions in response to music (e.g., the GEMS of Zentner et al., 2008). Even if it is true that listeners can only perceive around five categories of emotion in music, the results of the PUMS database analysis indicate that investigators are exploring other emotions that might be able to be perceived in music.

Close analysis, however, revealed that over half of music-related emotion stimuli concentrated on only three emotions: *sadness*, *happiness*, and *anger*. For example, researchers have made use of at least 2,000 musical stimuli that they have labeled as *sad*. This collection of *sad* music includes seemingly disparate musical selections, such as Barber's *Adagio for Strings*, Tori Amos's *Icicle*, and Miles Davis's *Summer Night*. An important question for music research is whether all of these stimuli express *sadness* in the same way. In a series of studies, Warrenburg (2019) showed that some of these *sad* musical passages contained different musical structures and resulted in different perceived and experienced emotions among listeners. Sandra Garrido, David Huron, and Jonna Vuoskoski have also written about various types of sadness that can be experienced in response to music (Garrido, 2017; Huron & Vuoskoski, 2019).

When stimuli with different characteristics are summarized using a single term, problems may arise when

comparing statistics across stimuli of this class. In general, using more distinct and nuanced (“*emotionally granular*”) terms may benefit the field as a whole (see Warrenburg, 2019 for more details on this idea). Pioneers of the field of music psychology, such as Kate Hevner and Malcolm Rigg, found that musical expression of emotion can be broad and imprecise, such that different listeners do not always agree on more exact emotion characterizations (Hevner, 1936; Rigg, 1964). Another possibility, however, is that the current definitions of music-related emotions do not provide enough information to describe exactly what is experienced subjectively or semantically. Participants may agree on more nuanced terms, provided that they are told specifically what a researcher means by a possibly nebulous term such as *melancholy*. In other words, the field of music and emotions could suffer from semantic underdetermination.

The longitudinal analysis suggested that there were nine emotional terms studied in the 1980s, 38 in the 1990s, 70 in the 2000s, and 51 in the 2010s (2010–2018). The expansion of emotion categories over time could be due to a number of factors, including primarily studying basic emotions throughout the 1980s and the use of dimensional models in the 1990s onwards (Eerola & Vuoskoski, 2013). It could also mean that people broke down a previous category of emotion (like *happiness*) into more than one term (like *joy* and *tender* and *peaceful*).

*Selection method for emotional stimuli.* Another important choice a researcher makes when designing a study on music and emotion is how to choose the particular emotional stimuli. If a researcher wants to study *sad* music, for example, they must choose the appropriate stimuli. There are many ways of choosing these stimuli. Some of the common ways of selecting emotional stimuli include expert opinion, pilot testing, or asking professional musicians to express emotions when playing a stimulus. Using different operationalizations of emotional terminology—like how researchers define and select *sad* music—will similarly impact the results of the study. There are pros and cons for each selection method. Using previously validated materials, for example, makes it possible to provide inferences across a broad literature. For example, researchers who are interested in developmental or cultural differences may be especially drawn to using previously used stimuli to enable comparison with published measurements of Western adult undergraduates. On the other hand, by using stimuli curated for another study and by relying on the way other researchers have defined and selected their emotional music stimuli, a researcher may be using an unintentionally biased sample.

As one example of how researchers could be unintentionally using a biased sample of music, research suggests that music previously labeled as *sad* can be better defined by using at least two emotional terms: *melancholic* and *grief-like* music (Huron, 2015; Warrenburg, 2019). Although affective responses to musical stimuli are influenced by variables such as personality characteristics, listening context, and familiarity with the musical passage, Warrenburg (2019) has found that there are differences in the musical structure of *melancholic* and *grieving* music and that listeners perceive and experience different emotions in response to these two types of musical passages. *Melancholy* music tends to be quieter, lower-in-pitch, and contains narrow pitch intervals, whereas *grief-like* music tends to contain sustained tones, gliding pitches, and harsh timbres (Warrenburg, 2019). However, in many studies, these different types of musical emotions are both labeled as *sad*. If one relies on others’ definition of *sad* music, they may accidentally use a *grief-like* stimulus when they are truly interested in *melancholy-like* stimuli. Nevertheless, whether an expert chooses a *sad* stimulus because it was slow and in the minor mode, or because a performer attempted to convey sadness in his or her rendition of a song, it is important to understand the way researchers defined and selected their emotional stimuli.

The analysis of the PUMS database suggests that 46% of musical stimuli were chosen based on expert or experimenter opinion, 37% were selected because they were used in past studies, 7% were participant selected, 3% were pilot tested, 1% were specifically composed for the study, and 1% were performed by professionals who were asked to play or express certain emotions. Recall also that stimuli in studies of perceived emotion were selected because they were used in previous studies 51% of the time and were selected because of experimenter/expert opinion 38% of the time, whereas stimuli in studies of induced emotion were chosen because of experimenter/expert opinion 35% of the time, were selected by participant chosen stimuli 30% of the time, and were chosen because of their use in previous studies 19% of the time. It could be that studies of perceived emotion focus on previously validated stimuli, rather than pilot testing, because some authors are writing algorithms to recognize emotion in music (MER studies). Perhaps, however, previously validated stimuli are also used because the experimenters wish to regulate certain parameters, such as major mode and fast tempi for *happiness*.

*Style of music.* The PUMS database analysis suggested that 57% of musical stimuli were classified as popular music, 10% of musical stimuli were classified as Western Art Music, and 7% of musical stimuli were

film or film-like music. The focus on Western Art Music is notable. One of the potential confounds of conflating studies that have used popular music and Western Art Music is that popular music may be more likely to contain emotional lyrics than the typically instrumental WAM. Of course, whether or not the use of music with lyrics causes a confound with music without lyrics depends on the specific research question. Popular music may also employ a different harmonic vocabulary compared with WAM (de Clercq & Temperley, 2011), so that it becomes more difficult to compare specific musical features that may be giving rise to participants' emotional states.

The increasing use of film music is a trend that is important because this music may be composed specifically to help convey or induce an emotion in listeners (during film viewing), which may correspond with the film scene (Hoeckner, Wyatt, Decety, & Nusbaum, 2011). The small number of cross-cultural studies (or studies that use music from non-Western cultures) is also notable. Future research should address this shortcoming of the PUMS database. The longitudinal analyses showed that the reliance on Western Art Music from the 1920s through the 1990s has given way to a new focus on Popular Music. Future analyses should examine whether or not these trends are due to a shift in theoretical focus of the researchers.

#### HOW SHOULD FUTURE RESEARCHERS SELECT EMOTIONAL MUSIC STIMULI?

*Methods of stimulus selection and emotional term operationalization.* In the creation of the PUMS database, I found that researchers selected musical stimuli in different ways, including *experimenter/expert chosen*, *previous studies*, *professionals asked to play/express emotions*, *pilot tested*, and *composed for study*. As mentioned earlier, there are benefits to each of these methods and the best method of stimulus selection will depend on the aim of the study. One word of caution about using relying on curated excerpts used in previous studies. As documented in past research (Eerola & Vuoskoski, 2011, 2013), many stimuli were chosen as a representative sample of emotion in early studies and then were used in a large number of subsequent studies by the same and other researchers. The PUMS database analysis indicates that 37% of stimuli used in research studies have been chosen because they were used in previously published papers. This trend is increasing, from 10% in the 1990s and before, to 23% in the 2000s, and to 42% in the 2010s. A potential problem with using music from previous studies could occur when many studies rely on the same stimuli. In these

cases, although the stimuli may have been well-chosen for an initial study, these samples of music may be (unintentionally) biased and lead to misrepresentative results in subsequent studies.

As one example of this complication, studies using the Music Mood Induction Procedure (MMIP) tend to rely on Delibes' *Coppelia* to induce happy or elated moods and on Prokofieff's *Russia under the Mongolian Yoke* (notably played at half speed) to induce sad or depressed moods (see Västfjäll, 2002, and Warrenburg, 2019, for more details on the music in MMIP). The protocol for MMIP includes asking the participants to think of past unhappy experiences, engage in fantasies about death, dance to the music, and think about possible future pleasant events. These instructions serve the purpose of MMIP studies well. From a musical perspective, however, it is impossible to know if these musical stimuli alone are able to alter a person's mood state.

One way to avoid using a (possibly) biased sample is to use the PUMS database. Instead of duplicating the stimuli from a single study, future researchers can use stimuli across multiple studies to create a new sample of emotional music stimuli. If the researcher wishes to use 30 musical stimuli, for example, they can use the PUMS database to identify two stimuli from ten different studies. The result of this process will be a new sample of music that has been validated by past researchers. By creating new samples for each study, findings in the field of music and emotion will be based on a more representative sample and will be less likely to promote findings based on an (unintentionally) biased sample curated for a single study.

*Implications for reviews and meta analyses.* We have seen that there have been varied approaches to sampling musical stimuli in the emotion literature. Musical stimuli in the research literature range from 2 seconds to about 45 minutes in duration, studies use musicians and nonmusicians as participants, experimenters recruit responses among individuals with different degrees of familiarity with the stimuli, and methodologies include different types of experimental conditions (e.g., ecologically-valid vs. laboratory-controlled). In any large review or meta-analysis of music and emotion studies (including this one), it is unavoidable to conflate results across these disparate research approaches. In other words, although individual studies may show certain trends, when combined with other studies, these trends may be weakened or disappear. It is possible, then, that some of the inconclusive results from previous reviews may be due to factors like the inconsistent use of emotion terms throughout the music community.

## Conclusions

This paper presented an analysis of the PUMS database to explore how research on music and emotion has utilized musical stimuli in the past. The careful selection of stimuli will benefit correlational, exploratory, and experimental designs and result in more power in statistical analyses. In the introduction, I asked several questions regarding how researchers should choose musical excerpts for their studies.

First, I asked if one should make use of extant musical recordings or create more controlled stimuli that are composed specifically for a study. In this paper, I presented data that suggests that about one percent of music and emotion studies have opted to use stimuli composed specifically for the study. I suggested that if familiarity with a musical stimulus will confound the results of the study, researchers should seek a composer to create stimuli for them. Composers, in turn, can examine the various operationalizations of emotion used in the PUMS (e.g., using features like major mode, fast tempo, and loud dynamics) to aid in their compositional process.

Second, I asked whether it would be better to use long or short stimuli, and whether stimuli should be excerpts or full works. I found that half of the stimuli in music and emotion studies are less than 30 seconds in duration, while only about 12% were two minutes or longer. I further found that 65% of stimuli in perceived emotion studies utilized stimuli less than 30 seconds in length, with almost 98% of these perceived emotion stimuli were less than one minute in length. If one wishes to study induced emotion, one should make use of longer stimuli. Fifty-five percent of induced emotion studies utilize musical stimuli longer than 1 minute, consistent with this idea. Although 55% of the stimuli were reported to be excerpts, only 8% of the stimuli longer than one minute in duration were full works, suggesting that most of the full works used were relatively short.

Third, I analyzed the style of the music used in studies of music and emotion. Although there was a wide spread of genres in the PUMS database, ranging from Heavy Metal to Mafa to Psychedelic Trance, over half of music used in these studies were considered to be Popular Music. A further 10% of these stimuli were Western Art Music. Despite this trend, the top ten composers sampled in these studies included Mozart, Beethoven, J. S. Bach, Chopin, Robert Schumann, Mendelssohn, Schubert, and Brahms. The concentration on Western Art Music composers the Top 10 list might suggest that the WAM music sampled in the literature relies on a more narrow range of music than does the popular

music sampled in the literature. In fact, while there are 134 WAM composers in the PUMS and 204 Popular Music composers in the PUMS, 12 WAM composers account for 50% of the WAM stimuli, while 57 Popular Music composers account for 50% of the Popular Music Stimuli. The style of the music may matter, depending on the goals of the study. For example, while 96% of Popular Music was used in studies of perceived emotion, only 38% of Western Art Music was used in studies of induced emotion. Additionally, over 99% of Film Music was unfamiliar to participants, while only 74% of Western Art Music was unfamiliar to participants.

Fourth, the familiarity of participants with the musical stimuli was examined. Although 55% of the stimuli in the PUMS were explicitly used because they were unfamiliar to participants, 35% were used because they were familiar to participants. However, while 83% percent of stimuli used in perceived emotion studies were unfamiliar to participants, only 43% of stimuli used in induced emotion studies were unfamiliar to participants. It makes sense that when studying experienced emotion, researchers may want to utilize stimuli with which participants have previous emotional connections. Consistent with this claim, 38% of familiar music was chosen by the participants themselves.

Finally, and most importantly, I investigated how researchers have selected their stimuli and how they may have operationalized the mood or emotion in a particular musical passage. The analysis of the PUMS database is consistent with the idea that 46% of musical stimuli are chosen by experimenters or other experts, 37% are selected because they were used in previous peer-reviewed publications, and only 3% of musical stimuli are pilot tested. When studying music-induced sadness, 47% of the stimuli were participant chosen, 13% were pilot tested, and none of the stimuli were performed by professionals with the intent of inducing sadness in listeners. On the other hand, when studying how music can represent sadness to others, only 4% of the stimuli were selected by participants, 4% were pilot tested, and 11% were played by professionals asked to directly express sadness. When researchers aimed to induce fear in listeners, however, they did not use any participant-selected works. Instead, the researchers primarily used pilot testing (39%) and expert/experimenter opinion (32%) to select musical passages that might induce fear in listeners. These comparisons highlight how study design depends on the emotion(s) in question and the specific aims of the researcher.

In summary, the current paper aimed to examine how researchers have selected and used emotional music stimuli since the 1920s. I showed that the literature has

relied on nine emotional terms (see Eerola & Vuoskoski, 2013, and Warrenburg, 2019, for detailed reviews of how other researchers have measured emotional responses). The implications of previous research conflating multiple emotional states are profound, as the ability to discriminate different emotions affects all music and emotion literature, including meta-analyses. By observing trends of music stimuli over 90 years, we have seen how design methodologies have developed over time and vary by the specific emotional terminology used by researchers. The hope of this

review is that researchers will continue to carefully choose their stimuli and report on the selection process, aiding in the field of music and emotion for decades to come.

### Author Note

Correspondence concerning this article should be addressed to Lindsay Warrenburg, Weigel Hall, 1866 College Road N #110, Columbus, OH 43210. E-mail: warrenburg.3@osu.edu

### References

- ALBERSNAGEL, F. A. (1988). Velten and musical mood induction procedures: A comparison with accessibility in thought associations. *Behaviour Research and Therapy*, 26(1), 79–96.
- ALI, S. O., & PEYNIRCIOĞLU, Z. F. (2010). Intensity of emotions conveyed and elicited by familiar and unfamiliar music. *Music Perception*, 27, 177–182.
- ALTENMÜLLER, E., SCHUERMANN, K., LIM, V. K., & PARLITZ, D. (2002). Hits to the left, flops to the right: Different emotions during listening to music are reflected in cortical lateralization patterns. *Neuropsychologia*, 40(13), 2242–2256.
- 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*, 26, 355–364.
- BIGAND, E., TILLMANN, B., POULIN, B., D'ADAMO, D. A., & MADURELL, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition*, 81(1), B11–B20.
- 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(8), 1113–1139.
- BLOOD, A. J., ZATORRE, R. J., BERMUDEZ, P., & EVANS, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature neuroscience*, 2(4), 382–387.
- BOWER, G. H., & MAYER, J. D. (1989). In search of mood-dependent retrieval. *Journal of Social Behavior and Personality*, 4(2), 121–156.
- CBS INTERACTIVE. (2012). Play music, find songs, and discover artists. Retrieved from <https://www.last.fm/>.
- CLARK, D. M. (1983). On the induction of depressed mood in the laboratory: Evaluation and comparison of the Velten and musical procedures. *Advances in Behaviour Research and Therapy*, 5(1), 27–49.
- CLARK, D. M., & TEASDALE, J. D. (1985). Constraints on the effect of mood on memory. *Journal of Personality and Social Psychology*, 48, 1595–1608.
- COSTA-GIOMI, E., COHEN, L. B., SOLAN, D., & BROCK, A. (2008). Melodic categorization during the first year of life. In Miyazaki, K., Hiraga, Y., Adachi, M., Nakajima, Y., & Tsuzaki, M. (Eds.), *Proceedings of the 10<sup>th</sup> International Conference in Music Perception and Cognition*. Sapporo, Japan.
- DALLA BELLA, S., PERETZ, I., ROUSSEAU, L., & GOSSELIN, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, 80(3), B1–B10.
- DARLING, D. L. (1982). *The construction of a semantic differential scale designed to measure attitude toward selected styles of music*. Dissertation, University of Illinois at Urbana-Champaign.
- DE CLERCQ, T., & TEMPERLEY, D. (2011). A corpus analysis of rock harmony. *Popular Music*, 30(1), 47–70.
- DIBBEN, N. (2004). The role of peripheral feedback in emotional experience with music. *Music Perception*, 22, 79–115.
- DOLGIN, K. G., & ADELSON, E. H. (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. *Psychology of Music*, 18(1), 87–98.
- DUINKER, B., & LÉVEILLÉ GAUVIN, H. (2017). Changing content in flagship music theory journals, 1979–2014. *Music Theory Online*, 23(4), 3.
- EEROLA, T. (2011). Are the emotions expressed in music genre-specific? An audio-based evaluation of datasets spanning classical, film, pop and mixed genres. *Journal of New Music Research*, 40(4), 349–366.
- EEROLA, T., & VUOSKOSKI, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 39(1), 18–49.
- EEROLA, T., & VUOSKOSKI, J. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception*, 30, 307–340.
- FILIPIC, S., TILLMANN, B., & BIGAND, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review*, 17(3), 335–341.

- GABRIELSSON, A., & LINDSTRÖM, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 367–400). New York: Oxford University Press.
- GARRIDO, S. (2014). A systematic review of the studies measuring mood and emotion in response to music. *Psychomusicology: Music, Mind, and Brain*, 24(4), 316–327.
- GARRIDO, S. (2017). *Why are We Attracted to Sad Music?* Cham, Switzerland: Palgrave Macmillan.
- GOMEZ, P., & DANUSER, B. (2004). Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology*, 53(2), 91–103.
- GOSSELIN, N., PERETZ, I., NOULHIANE, M., HASBOUN, D., BECKETT, C., BAULAC, M., & SAMSON, S. (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, 128(3), 628–640.
- GOSSELIN, N., SAMSON, S., ADOLPHS, R., NOULHIANE, M., ROY, M., HASBOUN, D., ET AL. (2006). Emotional responses to unpleasant music correlates with damage to the parahippocampal cortex. *Brain*, 129(10), 2585–2592.
- FLORES-GUTIÉRREZ, E.O., 2001. *La respuesta emocional a la música: atribución de términos de la emoción a segmentos musicales*. Thesis for Master in Neurobiology. México: Universidad Nacional Autónoma de México.
- GENTILE, D. (1998). *Infants' discrimination of musical affective expressiveness*. Dissertation, University of Minnesota.
- GREWE, O., NAGEL, F., KOPIEZ, R., & ALTENMÜLLER, E. (2007). Listening to music as a re-creative process: Physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Perception*, 24(3), 297–314.
- HEVNER, K. (1935). The affective character of the major and minor modes in music. *The American Journal of Psychology*, 47(1), 103–118.
- HEVNER, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology*, 48(2), 246–268.
- HEVNER, K. (1937). The affective value of pitch and tempo in music. *The American Journal of Psychology*, 49(4), 621–630.
- HOECKNER, B., WYATT, E. W., DECETY, J., & NUSBAUM, H. (2011). Film music influences how viewers relate to movie characters. *Psychology of Aesthetics, Creativity, and the Arts*, 5(2), 146–153.
- HORN, K., & HURON, D. (2015). On the changing use of the major and minor modes 1750–1900. *Music Theory Online*, 21(1).
- HUNTER, P. G., SCHELLENBERG, E. G., & SCHIMMACK, U. (2008). Mixed affective responses to music with conflicting cues. *Cognition and Emotion*, 22, 327–352.
- HURON, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Boston: MIT press.
- HURON, D. (2015). Affect induction through musical sounds: An ethological perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(20140098).
- HURON, D., & VUOSKOSKI, J. K. (2019). *On the enjoyment of sad music: Pleasurable Compassion Theory and the role of trait empathy*. Paper presented at the biannual national conference of the Society for Music Perception and Cognition, New York, NY.
- IWANAGA, M., IKEDA, M., & IWAKI, T. (1996). The effects of repetitive exposure to music on subjective and physiological responses. *Journal of Music Therapy*, 33(3), 219–230.
- IWANAGA, M., & TSUKAMOTO, M. (1997). Effects of excitative and sedative music on subjective and physiological relaxation. *Perceptual and Motor Skills*, 85(1), 287–296.
- JUSLIN, P. N. (1997). Perceived emotional expression in synthesized performances of a short melody: Capturing the listener's judgement policy. *Musicae Scientiae*, 1(2), 225–256.
- JUSLIN, P. N. (2013a). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10, 235–266.
- JUSLIN, P. N. (2013b). What does music express? Basic emotions and beyond. *Frontiers in Psychology*, 4, 596.
- JUSLIN, P. N., & LAUKKA, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129(5), 770–814.
- JUSLIN, P. N., & VÄSTFJÄLL, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31(5), 559–575.
- KAWAKAMI, A., FURUKAWA, K., KATAHIRA, K., & OKANOYA, K. (2013). Sad music induces pleasant emotion. *Frontiers in Psychology*, 4, 311–326.
- KOELSCH, S., FRITZ, T., CRAMON, D. Y., MÜLLER, K., & FRIEDERICI, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping*, 27(3), 239–250.
- KREUTZ, G., RUSS, M. O., BONGARD, S., & LANFERMANN, H. (2003). Cerebral correlates of music listening. An fMRI-study on the effects of “happy” and “sad” classical music. *Nervenheilkunde*, 22(3), 150–156.
- KRUMHANSL, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, 51(4), 336–352.
- LEVITIN, D. J., & MENON, V. (2003). Musical structure is processed in “language” areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, 20(4), 2142–2152.
- LIÉGOIS-CHAUVEL, C., BÉNAR, C., KRIEG, J., DELBÉ, C., CHAUVEL, P., GIUSIANO, B., & BIGAND, E. (2014). How functional coupling between the auditory cortex and the amygdala induces musical emotion: a single case study. *Cortex*, 60, 82–93.
- MAGNATUNE. (2014). A music service that isn't evil. Retrieved from <http://magnatune.com/>

- MARTIN, M. (1990). On the induction of mood. *Clinical Psychology Review*, 10(6), 669–697.
- McFARLAND, R. A. (1985). Relationship of skin temperature changes to the emotions accompanying music. *Biofeedback and Self-Regulation*, 10(3), 255–267.
- NAGEL, F., KOPIEZ, R., GREWE, O., & ALTENMÜLLER, E. (2007). EMuJoy: Software for continuous measurement of perceived emotions in music. *Behavior Research Methods*, 39(2), 283–290.
- NATER, U. M., ABBRUZZESE, E., KREBS, M., & EHLERT, U. (2006). Sex differences in emotional and psychophysiological responses to musical stimuli. *International Journal of Psychophysiology*, 62(2), 300–308.
- NYKLÍČEK, I., THAYER, J. F., & VAN DOORNEN, L. J. (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, 11, 304–321.
- OMAR, R., HAILSTONE, J. C., WARREN, J. E., CRUTCH, S. J., & WARREN, J. D. (2010). The cognitive organisation of music knowledge: a clinical analysis. *Brain*, 133(4), 1200–1213.
- PANKSEPP, J. (1995). The emotional sources of “chills” induced by music. *Music Perception*, 13, 171–207.
- PERETZ, I., BLOOD, A. J., PENHUNE, V., & ZATORRE, R. (2001). Cortical deafness to dissonance. *Brain*, 124(5), 928–940.
- PERETZ, I., GAGNON, L., & BOUCHARD, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68(2), 111–141.
- Prometheus Global Media, LLC. (2005). Billboard Hot 100 Songs - Year-End. Retrieved from <https://www.billboard.com/charts/year-end/2005/hot-100-songs>.
- RAMOS, J., GUEVARA, M. A., MARTÍNEZ, A., ARCE, C., DEL RIO, Y., AMESCUA, C., & CORSI-CABRERA, M. (1996). Evaluación de los estados afectivos provocados por la música. *Revista Mexicana Psicología*, 13(2), 131–145.
- RICKARD, N. S. (2004). Intense emotional responses to music: A test of the physiological arousal hypothesis. *Psychology of Music*, 32(4), 371–388.
- RIGG, M. G. (1964). The mood effects of music: A comparison of data from four investigators. *The Journal of Psychology*, 58(2), 427–438.
- ROY, M., PERETZ, I., & RAINVILLE, P. (2008). Emotional valence contributes to music-induced analgesia. *Pain*, 134(1–2), 140–147.
- SCHAFFRATH, H. (1995). The Essen folksong collection. In D. Huron (Ed.), *Database containing 6,255 folksong transcriptions in the Kern format and a 34-page research guide* [computer database]. Menlo Park, CA: CCARH.
- SCHELLENBERG, E. G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition*, 58(1), 75–125.
- SCHELLENBERG, E. G., PERETZ, I., & VIEILLARD, S. (2008). Liking for happy-and sad-sounding music: Effects of exposure. *Cognition and Emotion*, 22(2), 218–237.
- SCHUBERT, E. (2001). Continuous measurement of self-report emotional response to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 393–414). New York: Oxford University Press.
- SCHUBERT, E. (2007). The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music*, 35(3), 499–515.
- SCHUBERT, E. (2013). Emotion felt by the listener and expressed by the music: Literature review and theoretical perspectives. *Frontiers in Psychology*, 4, 837.
- SCHULLER, B., DORFNER, J., & RIGOLL, G. (2010). Determination of nonprototypical valence and arousal in popular music: Features and performances. *EURASIP Journal on Audio, Speech, and Music Processing*, 2010, 1–20.
- SHERMAN, M. (1928). Emotional character of the singing voice. *Journal of Experimental Psychology*, 11(6), 495–497.
- Smithsonian Collection. (1981). *The Smithsonian Collection of Classic Country Music* [CD]. New York: Smithsonian Collection of Recordings.
- SONG, Y., DIXON, S., & PEARCE, M. (2012). Evaluation of musical features for emotion classification. In Gouyon, F., Herrera, P., Martins, L. G., & Müller, M. (Eds.), *Proceedings of the 13<sup>th</sup> International Society for Music Information Retrieval Conference* (pp. 523–528). Porto: FEUP Edições.
- SONG, Y., DIXON, S., PEARCE, M., & FAZEKAS, G. (2013). Using tags to select stimuli in the study of music and emotion. In Luck, G., & Brabant, O. (Eds.), *Proceedings of the 3rd International Conference on Music & Emotion*. Jyväskylä: University of Jyväskylä, Department of Music.
- SPACKMAN, M. P., FUJIKI, M., BRINTON, B., NELSON, D., & ALLEN, J. (2005). The ability of children with language impairment to recognize emotion conveyed by facial expression and music. *Communication Disorders Quarterly*, 26(3), 131–143.
- SPECK, J. A., SCHMIDT, E. M., MORTON, B. G., & KIM, Y. E. (2011). A comparative study of collaborative vs. traditional musical mood annotation. In Klapuri, A., & Leider, C. (Eds.), *Proceedings of the 12<sup>th</sup> International Society for Music Information Retrieval Conference* (pp. 549–554). Miami.
- STRATTON, V. N., & ZALANOWSKI, A. H. (1991). The effects of music and cognition on mood. *Psychology of Music*, 19(2), 121–127.
- TARUFFI, L., & KOELSCH, S. (2014). The paradox of music-evoked sadness: An online survey. *PLoS ONE*, 9(10), e110490.
- Thomson Multimedia. (2002). Metadata for audio: automatic music classification project description. Metadata for Audio: Automatic Music Classification Project Description. Rennes, France: Thomson Multimedia.
- VÄSTFJÄLL, D. (2002). Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae*, 5(1\_suppl): 173–211.

- VIEILLARD, S., PERETZ, I., GOSSELIN, N., KHALFA, S., GAGNON, L., & BOUCHARD, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion*, 22(4), 720–752.
- VINES, B. W., KRUMHANSL, C. L., WANDERLEY, M. M., & LEVITIN, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, 101(1), 80–113.
- VUOSKOSKI, J. K., & EEROLA, T. (2012). Can sad music really make you sad? Indirect measures of affective states induced by music and autobiographical memories. *Psychology of Aesthetics, Creativity, and the Arts*, 6(3), 204–213.
- WANDERLEY, M. M. (2001). Quantitative analysis of non-obvious performer gestures. In I. Wachsmuth & T. Sowa (Eds.), *Proceedings of the International Gesture Workshop* (pp. 241–253). Berlin: Springer.
- WARRENBURG, L. A. (2019). *Subtle semblances of sorrow: Exploring music, emotional theory, and methodology* (PhD dissertation). The Ohio State University.
- WENINGER, F., EYBEN, F., SCHULLER, B. W., MORTILLARO, M., & SCHERER, K. R. (2013). On the acoustics of emotion in audio: What speech, music, and sound have in common. *Frontiers in Psychology*, 4, 292–304.
- WU, T. L., & JENG, S. K. (2006). Automatic emotion classification of musical segments. In Eerola, T., Luck, G., & Toiviainen, P. (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition* (pp. 385–393). Bologna.
- WU, T. L., & JENG, S. K. (2008). Probabilistic estimation of a novel music emotion model. In S. Satoh, F. Nack, & M. Etoh (Eds.), *Proceedings of the Advances in Multimedia Modeling, 14th International Multimedia Modeling Conference* (pp. 487–497). Kyoto, Japan: Springer.
- YANG, Y. H., & CHEN, H. H. (2011). Prediction of the distribution of perceived music emotions using discrete samples. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(7), 2184–2196.
- YANG, Y. H., LIN, Y. C., SU, Y. F., & CHEN, H. H. (2007). Music emotion classification: A regression approach. In *Proceedings of the IEEE International Conference on Multimedia & Expo (ICME)* (pp. 208–211). Beijing, China: IEEE.
- YANG, Y. H., LIU, C. C., & CHEN, H. H. (2006). Music emotion classification: A fuzzy approach. In *Proceedings of the 14th ACM International Conference on Multimedia* (pp. 81–84). Santa Barbara, CA: ACM.
- YANG, Y. H., SU, Y. F., LIN, Y. C., & CHEN, H. H. (2007). Music emotion recognition: The role of individuality. In *Proceedings of the 15th ACM International Workshop on Human-Centered Multimedia* (pp. 13–22). Augsburg: ACM.
- ZENTNER, M., GRANDJEAN, D., & SCHERER, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, 8(4), 494–521.