Article

Musicae Scientiae
2019, Vol. 23(4) 426–441

DOI: 10.1177/1029864917752545
journals.sagepub.com/home/msx
SAGE

Unusualness as a predictor of music preference

Anthony Chmiel and Emery Schubert
The University of New South Wales, Australia

Abstract
This paper investigates the role of unusualness ratings in predicting music preference. In addition, the variables complexity and familiarity were rated for five music stimuli covering a range of styles. Ninety-four participants were exposed to each stimulus ten times over a three-week period. The three variables were tested as predictors of preference using linear and quadratic curve-fitting procedures. A linear increasing relationship was observed for familiarity, and inverted-U relationships were observed for unusualness and complexity. These results are consistent with Berlyne's inverted-U model, or a segment of the inverted-U in the case of familiarity. Unusualness was a good indicator of music preference, and explained more variance than complexity or familiarity. Furthermore, the two stimuli that scored highest in unusualness produced consistently low ratings of preference independent of exposure, which appears to be a hallmark of "extreme" music stimuli.

Keywords
Music preference, empirical aesthetics, unusualness, inverted-U, Daniel Berlyne

Introduction

As music listening occurs in a multitude of settings, and for a range of purposes, the overall question of *why* we like music and *how* we can predict what music we like is a multifaceted one. In recent years much of the focus has been centered on ecologically-based approaches, such as the situational context of listening (Krause & North, 2017; North & Hargreaves, 2000b; Schellenberg, Peretz, & Vieillard, 2008), the emotions produced by the music (Garrido & Schubert, 2011; Huron, 2011; North & Hargreaves, 1997; Schubert, 2007; Vuoskoski & Eerola, 2012), and listener personality types (Hunter & Schellenberg, 2011) just to name a few. While these areas have provided fruitful insights, a recent review (Chmiel & Schubert, 2017a) on the reported relationships between preference and *collative variables*—the more "traditional" parameters used in experimental aesthetics—has highlighted that these variables may still have an important role to play in predicting music preference.

Collative variables were introduced by Berlyne (1960, 1971, 1974) as perceivable variables that can be analyzed and compared (or simply, collated). The most commonly tested collative

Corresponding author:
Emery Schubert, Empirical Musicology Laboratory, School of the Arts and Media, The University of New South Wales, Sydney, New South Wales 2052, Australia.
Email: e.schubert@unsw.edu.au

variables in studies of music preference are *complexity* and *familiarity/novelty* (Chmiel & Schubert, 2017a; Finnäs, 1989), although Berlyne also refers to change, suprisingness, ambiguity, and puzzlingness as other examples (Berlyne, 1971, p. 69). Berlyne proposed that collative variables are the primary determinants of preference, introducing an *inverted-U model* of preference to explain the relationship between the variables. According to the model, preference will trace out a parabolic relationship with a collative variable, meaning that a stimulus containing a moderate amount of a collative variable will be the most preferred. Increases or decreases in the level of a collative variable from this optimal point will be less preferred (see, Berlyne, 1971, p. 89).

Importantly, in cases where only a small range of a collative variable is used, only a segment of the overall inverted-U curve may be observed (Berlyne, 1974, p. 176; Heyduk, 1975, p. 84; Walker, 1973, p. 69). As an example, a piece of music that is exposed to a participant only a handful of times may produce an increase in preference without any decline. If the same participant was exposed to the same piece of music many times, a complete inverted-U relationship may be observed. Similarly, if a participant is repeatedly exposed to a piece of music they are already familiar with, only a decrease in preference might be observed as a function of exposure. We refer to this concept as the *segmented* inverted-U model. Chmiel and Schubert's (2017a) review of the literature found that nearly 90% of previous studies reported preference trajectories compatible with an inverted-U relationship, or a segment thereof, as a function of a collative variable. The data suggest that when other variables (such as ecological, referring to variables associated with meaningfulness and associative value) are kept constant, collative variables are able to explain a significant amount of preference response to music. In other words, the segmented inverted-U model can be thought of as an overarching trend that can be used to predict the general trajectories of music preference. However, in some circumstances preference trajectories may run counter to the model. In the following review, we attempt to identify one such exception.

Music stimuli used in previous investigations

One of the limitations of past empirical research on music preference is concerned with the nature of the stimuli used. These studies have primarily used mainstream, accessible examples such as popular music, classical music, and traditional styles of jazz (Chmiel & Schubert, 2017a). Alternatively, they used highly artificial tone sequences explicitly designed to test theories of music preference (that is, stimuli lacking in ecological credibility). Apart from these, only a handful of studies have investigated music that could be considered "ecologically realistic" but unusual, and even fewer have used music that could be considered extreme cases (Chmiel & Schubert, 2017b). A conceptual problem exists in that the typical, accessible styles of music generally used in aesthetic studies (such as those mentioned above) should only produce a small range of a collative variable such as complexity (North & Hargreaves, 2000a). Analyses exclusively containing these types of stimuli therefore might not allow a complete inverted-U relationship to appear in many cases. Furthermore, the majority of papers (almost 95%, according to the review by Chmiel and Schubert, 2017a) have based their investigations primarily or solely on the two collative variables *complexity* and *familiarity/novelty*. Investigation of preference trajectory for collative variables should therefore include a sufficiently wide range of values for each variable.

Contextually extreme music (henceforth referred to as "extreme music") will be defined as music that pushes the boundaries of familiar musical idioms or consists of radically unfamiliar idioms for the listener. For example, the style of free jazz may incorporate non-conventional aspects of composition and performance for those only exposed to other styles. Additionally, music belonging to a culture that the listener is not familiar with may produce a similar result; For a Western listener, elements such as wide vibrato, large glissandi, radical tonal changes, and the presence of a foreign language might be perceived as contextually extreme. For further discussion on this, see Chmiel and Schubert (2017b). In Berlyne's conceptualisation we would assume "extreme music" to refer to any example of music that produces an extremely high perception of the collative variable complexity. However, the term may also be applicable to other concepts, such as the level of *unusualness* that a listener perceives when listening to the music. We therefore propose that empirical investigations should include music stimuli that putatively vary in terms of *unusualness*, which in itself is a previously untested variable in music preference to our knowledge.

Extreme music is a particularly interesting area of aesthetics because it appears to challenge the robustness of the inverted-U preference path that is found in the literature as a general function of exposure (Chmiel & Schubert, 2017b). Rather, the few studies using extreme stimuli seem to suggest a floor-effect trend of preference, in which preference ratings remain at or close to the minimal rating regardless of subsequent exposures, counter to the predictions of the inverted-U model. Should such an exception to the model be identified, it could signify the start of new perspective on the inverted-U, in which the model is seen as a useful overall trend of music preference except in cases where specific criteria are met (such as in cases where contextually the stimulus is extreme, for one). Our examination of the literature identified only two studies that used realistic (non-abstract) stimuli that could be considered as such.

In the first of the two studies using extreme stimuli, Hargreaves (1984, Experiment 2) examined preference responses for two "typical" excerpts (one excerpt of popular music, and one excerpt of classical music), alongside an excerpt of avant-garde jazz. The two typical stimuli produced inverted-U trajectories of preference across 12 exposures, whereas preference for the remaining stimulus remained close to the minimum preference rating for the entire 12 exposures. Hargreaves' study listed the avant-garde stimulus as a piece by the duo A Touch of the Sun. We were not able to find a copy of the music for analysis; however, we were able to talk to Simon Mayo (clarinettist of A Touch of the Sun) via personal correspondence (see also Chmiel and Schubert, 2017b), and ask him to rate the two variables complexity and unusualness for the music on 11-point scales (0 to 10) for a general audience. Mayo described the piece as "a chaotic combination of influences from jazz and avant-garde classical areas…for the man in the street, the album would have been considered at the time maximum extremeness (10,10) for both features (complexity and unusualness)" (S. J. Mayo, personal correspondence, April 7, 2016). We can therefore assume this stimulus to be an example of extreme music.

In the second extreme study, Downey and Knapp investigated preference ratings for nine Western music stimuli that they describe as "represent[ing] the principle forms of music" (Downey and Knapp, 1927, p. 224). Participants were exposed to the stimuli at five weekly sessions, and recorded preference on 9-point scales. A number of these nine stimuli were well-known Western classical works, such as by Tchaikovsky and Mendelssohn, and upon our examination of these stimuli we noted them all to be of a typical and accessible nature; we classified none of them as contextually extreme. Participants were also exposed to a tenth stimulus—an unidentified Cantonese piece—although this tenth stimulus was only included in four of the five weekly sessions. The arrangement of this piece was described as Chinese voice and Chinese orchestra, containing "…dissonance and irregular time, in strong contrast to the melodious and orderly selections [of the other stimuli]." Downey and Knapp further noted that "The Cantonese song was felt to be definitely unpleasant. A few persons complained that an anticipation of hearing it operated to reduce their enjoyment of the programme as a whole"

(Downey and Knapp, 1927, p. 228), and one of the participants described their response to this piece as "Interesting, but made head ache" (Downey and Knapp, 1927, p. 237). While this piece remains unidentified, and thus a definitive analysis is not possible, the responses strongly suggest that (in contextual terms for these listeners, at the time of their listening experiences) this stimulus acted as an example of extreme music.

Downey and Knapp (1927, p. 229) reported descriptive statistics for their ten stimuli across the five subsequent exposures. The mean preference trajectories for the nine Western stimuli all occurred at moderate preference levels, whereas the trajectory for the Cantonese stimulus remained at or close to the minimal preference rating. According to the descriptive statistics, mean preference appears to slightly increase across the subsequent exposures. Downey and Knapp did not perform any inferential analysis, and it therefore seems likely from the reported data that this trajectory would not have shown a statistical increase between the first and last exposures if a repeated measures analysis were performed. We therefore assume that this stimulus was responded to in a similar manner to the extreme stimulus from Hargreaves' (1984) study.

The results of these two examples of music seem to suggest that in extreme cases, preference responses adhere to a floor-effect in which preference does not increase as a function of exposure (as appears to be the case for Hargreaves' (1984) study, however we must be cautious in assuming this for Downey and Knapp's (1927) study, which only included four exposures). What is significant about this possible floor-effect, is that it cannot be explained by well-established models of preference in music psychology, such as the inverted-U model, or Zajonc's (1968, 2001) mere exposure effect. The mere exposure effect suggests that repeated exposures alone are sufficient to increase preference with a monotonic increasing relationship, although this effect has primarily been linked with highly distributed exposures or subliminal exposures (Moreland & Zajonc, 1977; Szpunar, Schellenberg, & Pliner, 2004), and therefore holds far less relevance to studies concerned with focused listening. Indeed, one possible explanation for such results could be that when exposures are distributed far enough, the seemingly monotonic relationship with preference is in fact a "straddling" of the optimal point of exposure. In other words, this could in fact be a facet of Berlyne's segmented model in which the variable familiarity does not reach oversaturation. A similar interpretation could be made for subliminal exposures (Szpunar, Schellenberg, & Pliner, 2004) if we assume that familiarity might not increase at the same rate for subliminal exposures to realistic music as it does for focused exposures. For detailed reviews of the mere exposure effect see Bornstein (1989), Martindale (1984, 1988), and Hargreaves and North (2010).

Many explanations of aesthetics, such as the two noted above, propose that at some point additional exposures should create an increase in preference from the minimum rating level (for reviews on alternative explanations to the inverted-U, see Chmiel & Schubert, 2017a; Finnäs, 1989; Hargreaves, 1986; Hargreaves & North, 2010). The extreme examples of music reported above may form an exception to the overarching tendencies of music preference. Alternatively, it is possible that extreme music requires a substantially larger number of subsequent exposures (in comparison to typical music) before an eventual increase in preference occurs. Regardless, this area of music aesthetics has received minimal attention. We therefore decided to perform an empirical investigation using stimuli that we assumed would receive moderate to high ratings of variables such as complexity. The study tested the relationship between preference and the variables *complexity*, *familiarity*, and *unusualness*. We specifically aimed to ensure a wide range of collative variables to allow complete inverted-U relationships to emerge. To our knowledge, unusualness has not been previously used as a predictor for music preference. Self-reported ratings of unusualness, alongside complexity, allowed us to subjectively define extreme/unusual stimuli. Further, such ratings also cater for individual differences and experiences, in which participants who are unfamiliar with a style or piece of music may rate it as extreme, while participants highly familiar might not rate it as unusual at all. The use of an unusualness rating could be related to complexity, but not exclusively so. Given the little research in the area, we decided it would be wise to include a rating for both. Our investigation of the literature on extreme music produced two hypotheses:

(H1) Preference varies as an inverted-U function of collative variables, or a segment thereof; and

(H2) Extreme music produces preference trajectories remaining at or close to the minimal rating (a floor-effect) as a function of exposure.

Method

Materials

Five stimuli were used, covering a range of styles to provide broad variance in the independent variables. The excerpt details and durations are listed in Table 1. Our stimuli included two examples of Western popular music, although our examples were specifically chosen to evoke extreme ends of (un)familiarity for this style. The first piece of popular music (*Happy*) can be regarded as well-known, in that it was the most successful commercial selling song of 2014 (http://www.billboard.com). The second piece (*Red ribbon*) fits the paradigm of Western popular music, however it was never released on a commercial level. We also used a piece of classical music (*Tallawarra*); we specifically selected an example that we believed would not previously be familiar to general listeners. Finally, we included two pieces intended to act as extreme examples. First, we selected a piece of atonal music (*Etwas bewegte*), under the assumption that many listeners, particularly non-musicians, would find the tonality and concepts such as dissonance to be extreme. We selected our second extreme piece (*Megalon*) from the website *The Weirdest Band in the World* (Manson, 2013). The music was described as a "...math-rock/metal/dub/videogame-soundtrack combo featuring double drums, lots of choppy guitar, some trumpet and an instrument that sounds like a Theremin but I think is just some kind of vintage synth."

Table 1. Details of stimuli used.

Abbreviated title	Stimulus details	Excerpt duration
Etwas bewegte	A. Webern, "Etwas bewegte Achtel" and "Bewegt," from Six pieces for large orchestra, Op. 6. On *Schoenberg, Webern, Berg: Orchestral works* [CD]. London: Warner Classics, 1909.	2:28
Red ribbon	Bright Young Things, "Red ribbon." On *The great Lonesome* [CD]. Rockhampton: Soulmate Records, 2010.	2:21
Megalon	Godswounds, "Megalon." On *Death to the babyboomers* [CD]. Sydney: Sonichimaera, 2014.	1:52
Tallawarra	J. Peterson, "Tallawarra." On *Works by various Australian composers* [CD]. Melbourne: Australian Music Unit, ABC Classic FM, 2000.	2:10
Happy	P. Williams, "Happy." On *Girl* [CD]. New York: I am other, 2013.	2:09

Note: due to their short duration of approximately one minute each, both movements by A. Webern were combined into one continuous excerpt, which will henceforth collectively be referred to as *Etwas bewegte*.

All excerpts were shortened versions of the original stimuli. That is, all excerpts began at the starting point of each stimulus, but were edited to fade out at an appropriate point after several minutes had elapsed. The exception to this was *Etwas bewegte* (see Table 1 for details).

Participants

Ninety-four participants were recruited from an undergraduate elective course containing a mixture of music students and non-music students. We asked participants to rate their musical background on a six-point scale (1: non-musician, $n = 5$; 2: music-loving non-musician, $n = 40$; 3: amateur musician, $n = 22$; 4: serious amateur-musician, $n = 14$; 5: semi-professional musician, $n = 12$; and 6: professional musician, $n = 1$) based on an item from Ollen's (2006) Musical Sophistication Index. Participants ranged in age from 18 to 59 years (mean = 21.6, standard deviation = 5.2), with 39 males (41.5%) and 55 females (58.5%).

Procedures

The experiment took place over a three-week period. Participants were exposed to the entire set of excerpts three times a week: once during an "in-class" environment, and an additional two times while at home. The excerpts were identical for both the in-class and online sessions. The two "home exposures" were accessed online, allowing the experimenters to monitor the completion of these exposures. After the third week, a final, fourth week meeting consisted of only the in-class exposure, being the tenth and final exposure. Responses were recorded only for the in-class exposures, which consisted of exposures 1, 4, 7, and 10. For the in-class exposures, participants were seated in a room with approximately 30 other participants and listened to the excerpts over common loudspeakers. Each participant was seated at a computer or personal device such as a laptop or tablet. Participants were asked to report their responses for the variables preference, complexity, familiarity, and unusualness ("I like this piece"; "The music sounds complex"; "This piece is familiar"; and "The piece is unusual") by using eleven-point rating scales ("Strongly Agree [10]"; "Strongly Disagree [0]"; and "Neither Agree nor Disagree [5]") for each item. Demographic information, such as age, sex, and music background, was recorded with an online survey in the week before the first exposure. All exposures also included listening response tasks, in which participants were required to include open-ended responses describing the sounds of the stimuli, and/or any musical qualities that became apparent/noticeable on that specific exposure—this served as an extra measure to check the completion of online exposures. Further analysis of open-ended responses is not reported. All 94 participants completed these tasks for all exposures.

Ethics

Prospective participants all agreed to participate and completed a written consent form. The study received ethics approval as part of the Music in My Life Project (University of New South Wales Human Ethics Approval HC13015).

Results and discussion

Curve-fitting of collapsed variables

Linear and quadratic regression analyses onto preference were performed separately for the independent variables complexity, familiarity, and unusualness. In each of the three analyses,

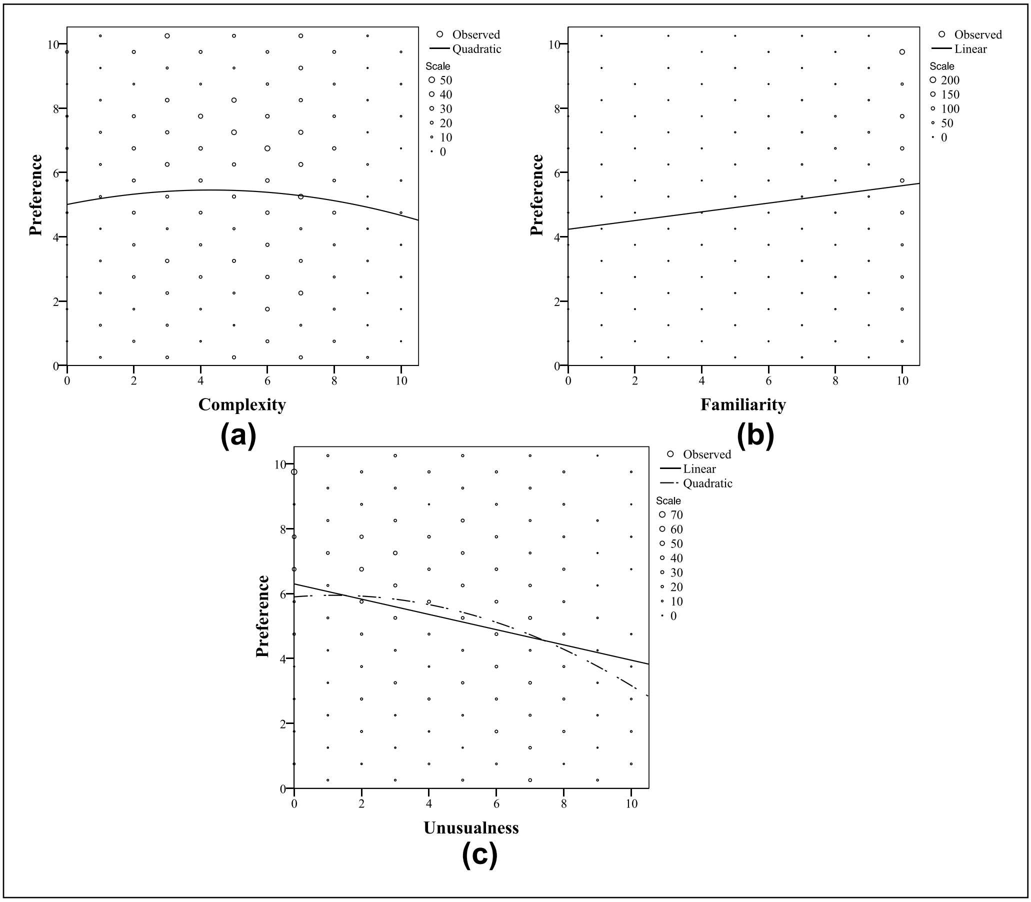

Figure 1. Curve estimations of the relationships between (a) preference and complexity; (b) preference and familiarity; and (c) preference and unusualness. Only curves with significant coefficients at $p < 0.05$ are shown.

preference was the dependent variable and participant responses were collapsed across exposures and stimuli. The results were collapsed across the five pieces in order to include a wide range of responses for the three collative variables. Preference and complexity produced no significant linear relationship with each other ($F(1, 1878) = 0.53$, $p = 0.464$), $R^2 < 0.001$. A significant quadratic relationship was observed between preference and complexity ($F(2, 1877) = 3.16$, $p = 0.042$), $R^2 = 0.003$. The quadratic term coefficient can be interpreted as an inverted-U relationship ($\beta = -0.19$, $t = -2.41$, $p = 0.016$), and can be observed in Figure 1(a). Preference and familiarity produced a significant linear relationship ($F(1, 1878) = 34.05$, $p < 0.001$), $R^2 = 0.018$. The linear term coefficient can be interpreted as the first, increasing segment of the inverted-U ($\beta = 0.13$, $t = 5.84$, $p < 0.001$), and can be observed in Figure 1(b). There was no significant quadratic relationship observed between preference and familiarity; the analysis of variance (ANOVA) was significant ($F(2, 1877) = 17.03$, $p < 0.001$), $R^2 = 0.18$, but the quadratic term coefficient was not ($\beta = -0.01$, $t = -0.16$, $p = 0.874$). Preference and unusualness produced a significant linear relationship ($F(1, 1878) = 96.02$, $p < 0.001$), $R^2 = 0.049$, and also a significant quadratic relationship ($F(2, 1877) = 57.52$, $p < 0.001$), $R^2 = 0.058$. The

Table 2. Mean (SD) for each variable. The "Overall" row indicates the average of the four measured exposures.

	Exposure	Preference	Familiarity	Complexity	Unusualness
Tallawarra	1	7.29 (1.95)	3.83 (3.00)	5.78 (2.37)	4.10 (2.15)
	4	7.09 (2.21)	8.67 (1.61)	6.11 (1.82)	4.53 (2.30)
	7	6.69 (2.60)	9.01 (1.58)	5.74 (2.15)	4.35 (2.29)
	10	7.12 (2.25)	9.13 (1.38)	5.39 (2.02)	3.59 (2.38)
	Overall	7.05 (2.25)	7.66 (1.89)	5.75 (2.09)	4.14 (2.28)
Red ribbon	1	5.64 (2.70)	3.45 (3.11)	3.18 (1.98)	2.73 (2.19)
	4	5.50 (2.92)	8.41 (1.98)	3.47 (1.93)	2.37 (2.12)
	7	5.39 (2.81)	9.10 (1.61)	3.21 (1.78)	2.49 (1.94)
	10	5.12 (3.05)	9.23 (1.50)	3.13 (1.93)	2.47 (2.05)
	Overall	5.41 (2.87)	7.55 (2.05)	3.25 (1.90)	2.51 (2.07)
Megalon	1	3.43 (2.61)	1.84 (2.25)	5.30 (2.34)	7.06 (2.23)
	4	2.97 (2.75)	8.09 (2.23)	5.18 (2.25)	5.99 (2.96)
	7	2.83 (2.75)	8.72 (2.10)	5.80 (2.27)	6.53 (2.19)
	10	2.77 (3.08)	8.93 (2.04)	5.41 (2.14)	6.18 (2.92)
	Overall	3.00 (2.80)	6.89 (2.15)	5.42 (2.25)	6.44 (2.57)
Etwas bewegte	1	4.07 (2.72)	3.99 (3.28)	7.05 (2.22)	6.40 (2.55)
	4	3.97 (2.68)	7.76 (2.38)	7.18 (1.65)	6.00 (2.51)
	7	3.67 (2.84)	8.43 (1.99)	6.78 (1.80)	6.14 (2.44)
	10	3.61 (3.02)	8.93 (1.51)	6.59 (2.04)	5.98 (2.75)
	Overall	3.83 (2.81)	7.28 (2.29)	6.90 (1.93)	6.13 (2.56)
Happy	1	7.45 (2.28)	9.57 (1.44)	3.31 (1.97)	2.40 (2.54)
	4	7.31 (2.40)	9.48 (1.46)	3.07 (2.22)	2.32 (2.32)
	7	7.05 (2.55)	9.71 (0.71)	3.23 (2.12)	2.27 (2.07)
	10	6.84 (2.67)	9.45 (1.38)	2.80 (2.07)	1.87 (1.84)
	Overall	7.16 (2.47)	9.55 (1.25)	3.10 (2.09)	2.21 (2.19)

linear term coefficient could be interpreted as the second, decreasing segment of the inverted-U (β = -0.22, t = -9.80, p < 0.001), and the quadratic term coefficient could be interpreted as an inverted-U relationship (β = -0.31, t = -4.26, p < 0.001). Both of these relationships are depicted in Figure 1(c). The quadratic (inverted-U) coefficient for the relationship between preference and unusualness has a relatively high effect size.

Analysis of variables between pieces and over subsequent exposures

Descriptive statistics for preference, familiarity, complexity and unusualness over the subsequent exposures are reported in Table 2, and the estimated marginal means across subsequent exposures for each variable are plotted in Figure 2. A mixed-design ANOVA was used to test the overall sample for differences in preference, with sex as a between-subjects factor. There was no significant difference in preference by sex (F(1, 92) < 0.01, p = 0.942, η^2 < 0.01); we therefore pooled the sample by sex. Two-way repeated-measures ANOVAs were performed for each of the four variables preference, familiarity, complexity, and unusualness, with *Piece* and *Exposure* as the within-subjects factor. Post hoc tests were subjected to Bonferroni correction to adjust for multiple comparisons. For preference, there was a significant omnibus (F(4, 90) = 67.10, p < 0.001, η^2 = 0.75), and a significant main effect for *Piece* (F(3.09, 287.24) = 67.45,

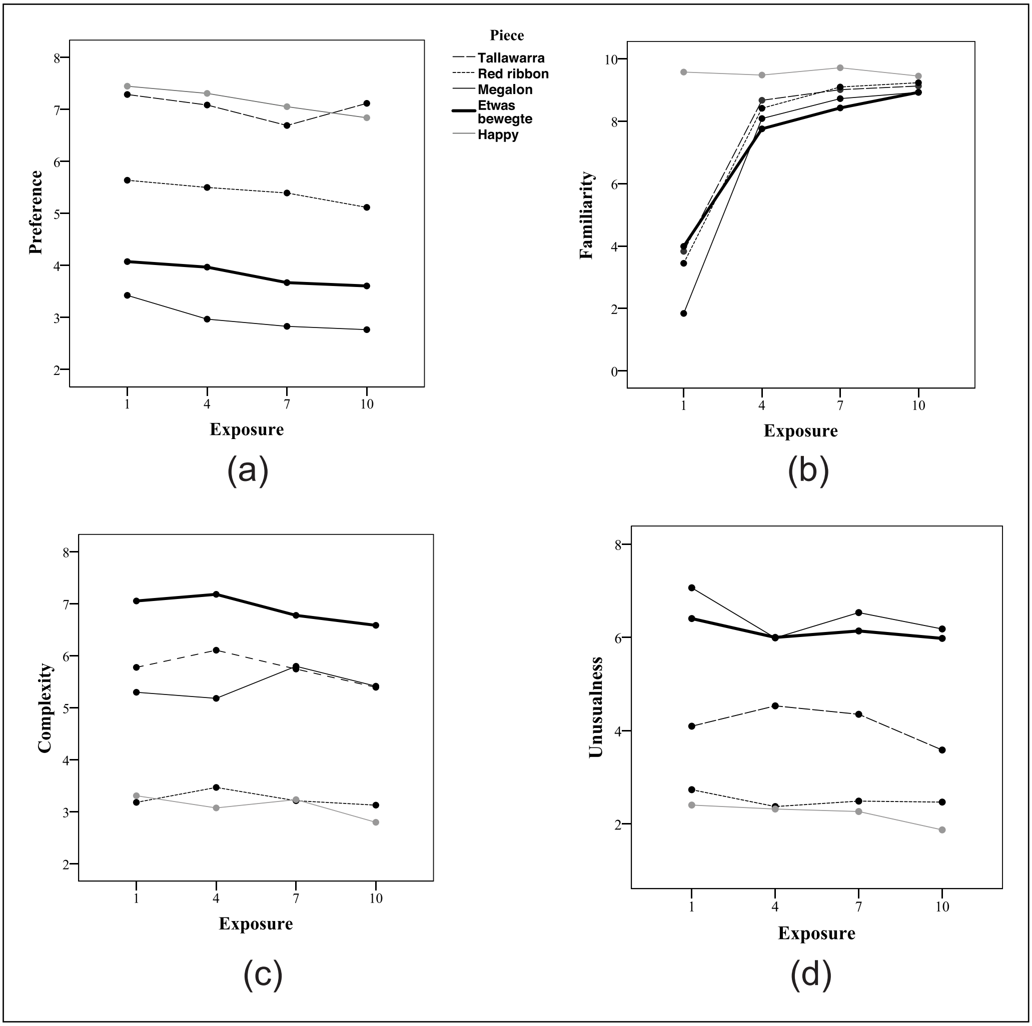

Figure 2. Plotted means for variables by piece and exposure: (a) preference; (b) familiarity; (c) complexity; and (d) unusualness.

$p < 0.001$, $\eta^2 = 0.42$). Post hoc paired-samples *t*-tests showed statistical differences between preference ratings for all pieces ($p < 0.001$), except between *Tallawarra* and *Happy* ($p > 0.999$, $d = 0.44$), and between *Megalon* and *Etwas bewegte* ($p = 0.06$, $d = 3.14$). For familiarity, there was a significant omnibus ($F(4, 90) = 102.46$, $p < 0.001$, $\eta^2 = 0.82$), and a significant main effect for *Piece* ($F(4, 372) = 126.22$, $p < 0.001$, $\eta^2 = 0.58$). Post hoc paired-samples *t*-tests showed statistical differences between familiarity ratings for all pieces ($p < 0.05$), except between *Tallawarra* and *Red ribbon* ($p > 0.999$, $d = 0.04$), and between *Red ribbon* and *Etwas bewegte* ($p = 0.23$, $d = 0.11$). For complexity, there was a significant omnibus ($F(4,90) = 92.3$, $p < 0.001$, $\eta^2 = 0.80$), and a significant main effect for *Piece* ($F(3.5, 320.63) = 138.15$, $p < 0.001$, $\eta^2 = 0.59$). Post hoc paired-samples *t*-tests showed statistical differences between complexity ratings for all pieces ($p < 0.001$), except between *Tallawarra* and *Megalon* ($p > 0.99$, $d = 1.18$), and between *Red ribbon* and *Happy* ($p > 0.999$, $d = 0.75$). For unusualness, there was a significant omnibus ($F(4, 90) =$

Table 3. Repeated-measures analysis of variance results for each stimulus and variable, with "exposure" as the within-subjects factor (*df* = 3279).

Stimulus	Preference	Familiarity	Complexity	Unusualness
Tallawarra	$F = 2.92$	$F = 180.32$	$F = 3.28$	$F = 4.39$
	$p = 0.043$	$p < 0.001$	$p = 0.025$	$p = 0.005$
	$\eta^2 = 0.03$	$\eta^2 = 0.66$	$\eta^2 = 0.03$	$\eta^2 = 0.04$
Red	$F = 1.79$	$F = 188.75$	$F = 1.04$	$F = 0.87$
ribbon	$p = 0.158$	$p < 0.001$	$p = 0.376$	$p = 0.448$
	$\eta^2 = 0.02$	$\eta^2 = 0.67$	$\eta^2 = 0.01$	$\eta^2 = 0.01$
Megalon	$F = 2.73$	$F = 282.10$	$F = 2.04$	$F = 5.05$
	$p = 0.051$	$p < 0.001$	$p = 0.120$	$p = 0.003$
	$\eta^2 = 0.03$	$\eta^2 = 0.75$	$\eta^2 = 0.02$	$\eta^2 = 0.05$
Etwas	$F = 1.98$	$F = 96.06$	$F = 3.01$	$F = 1.04$
bewegte	$p = 0.131$	$p < 0.001$	$p = 0.031$	$p = 0.370$
	$\eta^2 = 0.02$	$\eta^2 = 0.51$	$\eta^2 = 0.03$	$\eta^2 = 0.01$
Happy	$F = 4.10$	$F = 0.92$	$F = 2.20$	$F = 2.24$
	$p = 0.011$	$p = 0.422$	$p = 0.088$	$p = 0.084$
	$\eta^2 = 0.04$	$\eta^2 = 0.01$	$\eta^2 = 0.02$	$\eta^2 = 0.02$

82.18, $p < 0.001$, $\eta^2 = 0.78$), and a significant main effect for *Piece* ($F(2.9, 269.9) = 153.65$, $p < 0.001$, $\eta^2 = 0.62$). Post hoc paired-samples *t*-tests showed statistical differences between unusualness ratings for all pieces ($p < 0.001$), except between *Red ribbon* and *Happy* ($p = 0.67$, $d = 1.51$), and between *Megalon* and *Etwas bewegte* ($p > 0.999$, $d = 0.86$).

Main effect results for repeated-measures ANOVAs on *Exposure* are reported in Table 3. For preference, *Tallawarra* ($p = 0.043$, $\eta^2 = 0.03$) and *Happy* ($p = 0.011$, $\eta^2 = 0.04$) produced significant results, although *Megalon* could also be considered marginally significant ($p = 0.051$, $\eta^2 = 0.03$). Post hoc paired-samples *t*-tests for *Tallawarra* showed a significant decrease in preference ($p = 0.019$, $d = 0.26$) from exposures 1 to 7, and then a significant increase for the last two measured exposures ($p = 0.017$, $d = 0.18$). Post hoc tests confirmed the decrease in preference between the first and last exposures as significant for *Megalon* ($p = 0.029$, $d = 0.23$) and *Happy* ($p = 0.008$, $d = 0.25$). Preference results for *Red ribbon* ($p = 0.061$, $d = 0.18$) and *Etwas bewegte* ($p = 0.103$, $d = 0.16$) were not significant, however the reported means show a decrease. The preference results of *Red ribbon, Megalon, Etwas bewegte*, and *Happy* appear to outline the second, decreasing curve of the inverted-U. Additional exposures for *Red ribbon* and *Etwas bewegte* may have produced significant results.

The decreasing preference trajectories observed for *Red ribbon, Megalon, Etwas bewegte*, and *Happy* could be interpreted as the result of over-exposure; familiarity ratings for *Happy* were consistently high from the first exposure, and familiarity ratings for the remaining stimuli increased substantially between the first two measured exposures (exposures 1 and 4). A number of alternative interpretations are possible. For example, as subsequent exposures have been linked with producing decreases in perceived levels of complexity (Heyduk, 1975; Walker, 1973), such a change may also have moved the perceived level of complexity further away from the optimal complexity level (assuming that these stimuli were rated below this optimal level). A similar inference could be made for a decrease in unusualness with subsequent exposures; this is supported by the decreasing mean values of unusualness reported in Table 2. Therefore, a combination of both over-exposure and a decrease in perceived unusualness and/or complexity could have contributed to these decreasing trajectories of preference. Alternatively, these

Table 4. Correlations between unusualness and collative variables, collapsed across the four measured exposures ($n = 376$).

Stimulus	Unusualness and complexity	Unusualness and familiarity
Tallawarra	$r = 0.23, p < 0.001$	$r = –0.04, p = 0.435$
Red ribbon	$r = 0.35, p < 0.001$	$r = –0.11, p = 0.034$
Megalon	$r = 0.29, p < 0.001$	$r = –0.15, p = 0.003$
Etwas bewegte	$r = 0.35, p < 0.001$	$r = –0.08, p = 0.099$
Happy	$r = 0.32, p < 0.001$	$r = –0.07, p = 0.175$

changes in preference could be the combination of collative factors such as over-exposure, alongside other non-collative factors.

The results for *Tallawarra* appear to form a "standard-U shape." This result is therefore counter to our hypothesis, although a number of interpretations are possible. The increase in preference between exposures 7 and 10 may be attributed to some external non-collative factor. Alternatively, it is possible that interactions between multiple collative variables, which may contain individual inverted-U relationships with preference, have hidden the inverted-U relationship in the overall preference trajectory, such as the effects of exposure on complexity and unusualness described above. Alternatively, additional exposures from the tenth exposure may have produced subsequent decreases in preference for this stimulus, in which preference would have produced an overall decreasing trajectory, with a small "upward hump" around the tenth exposure. Indeed, this stimulus may also generate responses counter to Berlyne's model. In such a case, we suggest that the inverted-U can be best be thought of as a useful, general trend that tends to fit music preference, but may contain exceptions.

Correlation among independent variables

A Pearson product-moment correlation was performed between unusualness and complexity, and also between unusualness and familiarity. Correlations were performed for each stimulus, collapsed over the four measured exposures. This allowed us to determine the similarity between unusualness and each of the existing collative variables, if any. The results are reported in Table 4. Based on Cohen's (1992, p. 157) study, examination of effect size showed a small to moderate positive relationship between unusualness and complexity for each stimulus. No statistically significant relationship was observed between unusualness and familiarity for the stimuli *Tallawarra, Etwas bewegte*, and *Happy; Red ribbon*, and *Megalon* produced small negative relationships.

Identification of extreme stimuli

We identified which stimuli, if any, would be the most likely candidates as extreme music. The post hoc paired samples *t*-tests between stimuli (above) show that *Megalon* and *Etwas bewegte* were rated highest in terms of unusualness. In terms of complexity, *Etwas Bewegte* was rated significantly higher than all other stimuli. *Tallawarra* and *Megalon* were rated next highest in complexity, and post hoc tests showed that there was no significant difference in complexity between these two stimuli; however, they were rated significantly higher than *Red ribbon* and *Happy*. In other words, while there was no statistical difference in complexity between *Megalon* and *Tallawarra, Megalon* received a significantly higher rating of unusualness than *Tallawarra*, and is therefore a better candidate as an extreme stimulus. The same conclusion is reached by comparing the results of the "Overall" rows in Table 2. With congruous results from both types

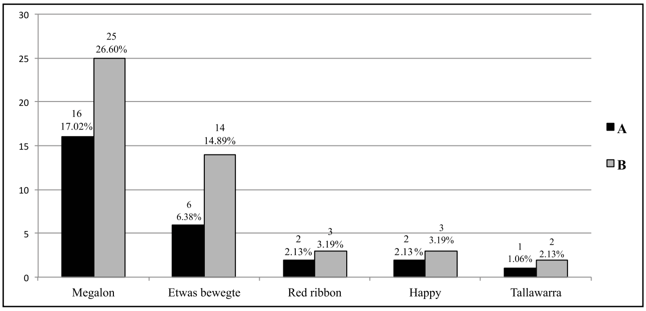

Figure 3. Bar chart depicting the number of individual floor-effect trajectories counted for each stimulus. Counts are shown for both Method *A* (preference remaining at 0 and/or 1 across subsequent exposures) and Method *B* (preference ratings remaining at 0, and/or 1, and/or 2). The percentages shown are in regards to the entire sample size.

of analysis, we propose *Megalon* and *Etwas bewegte* to be the most susceptible stimuli in this experiment to floor-effect responses of preference.

None of the stimuli produced floor-effect trajectories for the plotted mean preference scores depicted in Figure 2. However, this can be explained by comparing the ratings of complexity and unusualness received by our stimuli with the ratings provided by Simon Mayo for the duo A Touch of the Sun in our literature review (a score of 10 out of 10 for both variables). Both *Megalon* and *Etwas bewegte* were the strongest candidates of extreme music in our pool; however, the "Overall" rows in Table 2 show that these variables received only moderate ratings (overall mean ranging between 5.18 and 7.18) in comparison to the avant-garde stimulus used by Hargreaves (1984). Our stimuli should therefore not be considered extreme, as the previous examples in the literature can be. Regardless of this, we considered these stimuli sufficiently unusual to investigate the hypothesis in comparison to the remaining three stimuli.

Investigation of floor-effect trajectories

As we note above, *Megalon* and *Etwas bewegte* are the two best candidates of extreme stimuli; however, they only received moderate ratings of these variables. Preference trajectories were examined on an individual basis (that is, participant-by-participant) for each exposure and stimulus. We used two methods to investigate instances of floor-effects for each stimulus; we decided to include both methods due to the moderate ratings the "extreme" stimuli received. Method *A*, being the more conservative approach, considered preference trajectories that contained only the values "0" and/or "1" across the four measured exposures as floor-effect responses. Method *B* considered preference trajectories that contained the values "0," and/or "1," and/or "2" across the four measured exposures as floor-effect responses being based on the low mean preference ratings reported by Hargreaves (1984). For both methods, all other responses were classified as not exhibiting a "floor-effect." The number of floor-effect trajectories for each method were counted for each stimulus, and are reported in Figure 3.

In line with our above hypothesis, the pieces *Megalon* and *Etwas bewegte* produced a substantially higher count of floor-effect trajectories than the other three stimuli, which only produced either one or two counts. This result was observed regardless of which method was employed. This analysis therefore supports the hypothesis that the most extreme stimuli would be more susceptible to floor-effect trajectories, regardless of the fact that they did not receive extreme ratings for the variables complexity and unusualness. Additionally, for the 27 floor-effect responses recorded with Method *A*, 18 responses (66.67%) belonged to participants who rated themselves in categories 1, 2, or 3 in the musical background description, being the three categories associated with less musical training (see the Method section). The remaining nine responses (33.33%) for Method *A* belonged to participants who rated themselves in categories 4, 5, or 6. Similarly, for the 47 floor-effect compatible responses recorded with Method *B*, 31 (65.96%) belonged to participants who rated themselves in categories 1, 2, or 3, and the remaining 16 responses (34.04%) belonged to participants who rated themselves in categories 4, 5, or 6.

Conclusions

This study explored the use of unusualness as a predictor of preference, alongside complexity and familiarity. Preference was hypothesized (H1) to trace out an inverted-U as a function of each of these three collative variables, or a segment thereof. With curve-fitting analysis, an inverted-U relationship was observed between preference and unusualness, and between preference and complexity, although unusualness explained more of the variance. A linear increasing relationship was observed between preference and familiarity. While a linear result can occur with ten distributed exposures and still be considered part of the inverted-U (Finnäs, 1989), additional exposures could produce a complete inverted-U, as is often seen in studies including larger numbers of non-subliminal exposures to realistic music, such as up to 32 exposures in Schellenberg et al. (2008), and up to 64 exposures in Szpunar et al. (2004).

As noted in the literature review, there exists a conceptual problem when investigating the relationship between preference and collative variables while using stimuli that do not broadly sample each collative variables. While preference tends to trace out an inverted-U path as a function of a collative variable, when only a small range of this collative variable is present it may prevent the full inverted-U path from being produced. By collapsing the data across the five stimuli we were able to examine the relationship between preference and a wider range of these independent variables, such as complexity and unusualness, than is possible from the use of exclusively typical stimuli.

Our second hypothesis (H2) proposed that any extreme stimuli would produce floor-effect preference responses. The results partially supported this hypothesis. While floor-effects were not observed in any of the estimated marginal means for preference across subsequent exposures, this can be explained by the non-extreme ratings of complexity and unusualness for the stimuli that we a priori intended as examples of extreme music. Regardless, upon individual investigation of preference trajectories we observed substantially higher counts of floor-effect trajectories for the two stimuli that were rated highest in terms of complexity and unusualness. This result seems to suggest that the number of floor-effect responses received by a stimulus might be related in an increasing manner to the extreme nature of the stimulus. That is, as an example of music becomes more and more unusual, the inverted-U model may become less and less accurate at predicting preference as a function of exposure, up to a point where it no longer becomes a valid explanation. We therefore propose that future studies containing extreme stimuli (specifically, stimuli that receive substantially higher ratings of complexity and/or unusualness than ours did) might produce results with a higher count of floor-effect responses of preference, reminiscent of the results reported in the literature.

This is the first study that has provided evidence of a novel measure of preference—unusualness—that appears to behave in a manner consistent with collative variables, but may act independently of the often-studied collative variable "exposure." Our data also suggest that the variable unusualness is distinct from complexity or familiarity—there may be circumstances when unusualness is rated relatively high without an accompanying high score of complexity (and vice versa), such as a piece that is in a style highly foreign to the listener, but objectively uncomplicated. However, the present study had limitations. First, while 10 exposures are comparable to many similar studies in the literature, it is a relatively small number of exposures given the hypothesised proposal of an indefinite floor-effect as a function of exposure. We therefore recommend further study in this area with additional exposures. Regardless, even if further research finds an eventual climb in preference with additional, distributed exposures (for example), it would still indicate a very slowly rising segment of the inverted-U that is atypical with respect to the current understanding of music preference research (Chmiel & Schubert, 2017a). Further, the inclusion of additional variables might aid in the detailed analysis of the relationship between preference and collative variables. For example, a rating of the familiarity and/or interest that a respondent has for the *style* of music that each stimulus belongs to might aid in analysis, such as with over-exposure. With this inclusion, if a particular piece is not very familiar but the style is very familiar, an early onset of over-exposure could be expected. Further study may also benefit from including variables incorporating typicality, as it is conceptually plausible for unusualness to be negatively correlated with ecological variables such as typicality and/or meaningfulness, where quadratic functions of preference are not always found (Martindale, 1984, 1988).

Finally, future research designs could restrict stimuli to exclusively extreme examples of music, to reduce the possibility of a contrast effect (Schwarz & Bless, 1992), where the extreme stimuli are rated artificially lower because they are inadvertently compared against the non-extreme stimuli, rather than because of the independently perceived attribute of the stimulus itself. On the other hand, some kind of comparison or control was useful to ensure that the inverted-U relationship can emerge with regular stimuli in parallel with the extreme ones. Research may also benefit from recruiting exclusively non-musician participants, as it is generally understood that musicians (those with higher musical background ratings) hold a higher tolerance for variables such as musical complexity and unconventionality (Orr & Ohlsson, 2005). This is particularly relevant considering that both Methods *A* and *B* reported above recorded substantially higher counts of floor-effect trajectories for the first three categories of musical background descriptions, which refer to the lowest amount of musical training.

In summary, the findings of this study highlight the use of collative variables and the essentially forgotten inverted-U model of preference (Hargreaves & North, 2010). The results suggest that when a broad sample of music stimuli are used, collative variables can still predict the broad trajectory that preference will follow. However, the results also suggest that extreme music may prove an exception to the inverted-U, and indeed other existing explanations of music preference. The implications of such a finding may call for the development and reworking of some of these models. We therefore recommend further investigation with a larger number of exposures and unusual stimuli, including some that reach extreme levels of collative variables such as complexity and unusualness.

Funding

This work was supported by the Australian Research Council (Grant Number FT120100053).

References

Berlyne, D. E. (1960). *Conflict, arousal and curiosity*. New York, NY: McGraw-Hill.
Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: Appleton-Century-Crofts.

Berlyne, D. E. (Ed.) (1974). *Studies in the new experimental aesthetics*. Washington, DC: Hemisphere.
Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*(2), 265–289.
Chmiel, A., & Schubert, E. (2017a). Back to the inverted-U for music preference: A review of the literature. *Psychology of Music*, *45*(6), 886–909.
Chmiel, A., & Schubert, E. (2017b). Emptying rooms: When the inverted-U model of preference breaks down – An investigation using music with collative extremes. *Empirical Studies of the Arts*. Advance online publication. doi:10.1177/0276237417732683.
Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.
Downey, J. E., & Knapp, G. E. (1927). The effect on a musical programme of familiarity and of sequence of selections. In M. Schoen (Ed.), *The effects of music: A series of essays* (pp. 223–243). New York, NY: Books For Libraries Press.
Finnäs, L. (1989). How can musical preferences be modifed? A research review. *Bulletin of the Council for Research in Music Education*, *102*(1), 1–58.
Garrido, S., & Schubert, E. (2011). Individual differences in the enjoyment of negative emotion in music: A literature review and experiment. *Music Perception*, *28*(3), 279–296.
Hargreaves, D. J. (1984). The effects of repetition on liking for music. *Journal of Research in Music Education*, *32*(1), 35–47.
Hargreaves, D. J. (1986). *The developmental psychology of music*. Cambridge, UK: Cambridge University Press.
Hargreaves, D. J., & North, A. C. (2010). Experimental aesthetics and liking for music. In P. N. Juslin & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 515–546). New York, NY: Oxford University Press.
Heyduk, R. G. (1975). Rated preference for musical compositions as it relates to complexity and exposure frequency. *Perception and Psychophysics*, *17*(1), 84–91.
Hunter, P. G., & Schellenberg, E. G. (2011). Interactive effects of personality and frequency of exposure on liking for music. *Personality and Individual Differences*, *50*(2), 175–179.
Huron, D. (2011). Why is sad music pleasurable? A possible role for prolactin. *Musicæ Scientiæ*, *15*(2), 146–158.
Krause, A. E., & North, A. C. (2017). Pleasure, arousal, dominance, and judgements about music in everyday life. *Psychology of Music*, *45*(3), 355–374.
Manson, J. (2013). Weird band of the week: Godswounds. Retrieved from https://weirdestbandintheworld.com/2013/05/29/weird-band-of-the-week-godswounds/
Martindale, C. (1984). The pleasures of thought: A theory of cognitive hedonics. *Journal of Mind and Behavior*, *5*(1), 49–80.
Martindale, C. (1988). Aesthetics, psychobiology, and cognition. In F. H. Farley & R. W. Neperud (Eds.), *The foundations of aesthetics, art, and art education* (pp. 7–42). New York, NY: Praeger Publishers.
Moreland, R. L., & Zajonc, R. B. (1977). Is stimulus recognition a necessary condition for the occurence of exposure effects? *Journal of Personality and Social Psychology*, *35*(4), 191–199.
North, A. C., & Hargreaves, D. J. (1997). Liking, arousal potential, and the emotions expressed by music. *Scandinavian Journal of Psychology*, *38*(1), 45–53.
North, A. C., & Hargreaves, D. J. (2000a). Collative variables versus prototypicality. *Empirical Studies of the Arts*, *18*(1), 13–17.
North, A. C., & Hargreaves, D. J. (2000b). Musical preferences during and after relaxation and exercise. *The American Journal of Psychology*, *113*(1), 43–67.
Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings.* The Ohio State University, Unpublished dissertation.
Orr, M. G., & Ohlsson, S. (2005). Relationship between complexity and liking as a function of expertise. *Music Perception*, *22*(4), 583–611.
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. (2007). The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music*, *35*(3), 499–515.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: Assimilation and contrast effects in social judgement. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgements* (pp. 217–245). Hillsdale, NJ: Lawrence Erlbaum Associates.

Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 370–381.

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.

Walker, E. L. (1973). Psychological complexity and preference: A hedgehog theory of behavior. In D. E. Berlyne & K. B. Madsen (Eds.), *Pleasure, reward, preference* (pp. 65–97). New York, NY: Academic Press.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplement*, *9*(2), 1–27.

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, *10*(6), 224–228.