Back to the inverted-U for music preference: A review of the literature

Anthony Chmiel and Emery Schubert

Empirical Musicology Laboratory, School of the Arts and Media,

The University of New South Wales, Sydney, New South Wales 2052, Australia

Correspondence: e.schubert@unsw.edu.au

NOTE:

Back to the inverted-U for music preference: A review of the literature. (2017). *Psychology of Music 45*(6), pp. 886-909.

DOI: 10.1177/0305735617697507

Abstract

This study investigated the inverted-U model of preference for music as a function of collative variables (especially familiarity and complexity) over the last 115 years. The results of 57 studies on music preference were categorized according to their patterns of preference. Fifty of the 57 studies (88%) were categorized as compatible with an overarching (segmented) inverted-U model, while the results of five studies (9%) were interpreted as mixed, showing both compatible and incompatible results. Two studies (3%) were categorized as completely incompatible with the model. In contrast to authors who describe the model as defunct, this review has observed that studies producing results compatible with the inverted-U are still prevalent. We propose that while there may be inconsistencies with Berlyne's psychobiological theory from a scientific, arousal-based standpoint, the inverted-U model is able to explain a considerable amount of data. Rather, it seems that research interests have moved elsewhere, but caution is urged in asserting denial or dismissal of the relationship in music preference research.

Keywords: collative variables, Daniel Berlyne, experimental aesthetics, inverted-U, literature review, music response, preference, psychobiological theory

Back to the inverted-U for music preference: A review of the literature

Introduction

Why do we like certain pieces or types of music, and does our liking change over time? In the 1960s Daniel Berlyne's psychobiological theory provided a general explanation of aesthetic preference that could directly address these questions. He proposed that preference varies systematically as a function of *collative variables*, such as familiarity/novelty and complexity. Having links to research conducted by Wundt nearly a century earlier, the theory came to dominate all aspects of psychological research on aesthetic preference, including music preference, for over two decades (Hargreaves & North, 2010). This dominance subsided in the 1980s as new theories provided more sophisticated accounts of music preference, which expanded or set aside the collative variables, and some literature reviews give the impression that Berlyne's theory became of little or no relevance after that time (e.g., Hargreaves & North, 2010, p. 522; Silvia, 2005, p. 345).

Despite the increased sophistication of more recent work, this article seeks to investigate whether the findings predicted by Berlyne were rejected because they are not replicable, or for some other reason. Specifically, we examine literature concerned with finding a relationship between music preference and collative variables. In brief, Berlyne's theory argues that preference for aesthetic stimuli, and specifically music in the present case, maps out an inverted-U pattern as a function of a collative variable. First, this inverted-U model is explained, followed by a presentation of views and findings that appear to contradict the model, or render it defunct. We then present a detailed literature review to further investigate the question of an inverted-U relationship.

The inverted-U model

Berlyne's psychobiological theory focuses on the role of arousal as the primary determinant of aesthetic preference (Berlyne, 1960, 1967, 1969, 1970, 1971, 1974). Central to this thesis are three types of variables hypothesized to determine the level of arousal produced by a stimulus, and in doing so determine the level of preference felt toward that stimulus. These are (a) *psychophysical variables*, (b) *ecological variables*, and (c) *collative variables*. Psychophysical variables refer to perceivable differences in the features of a stimulus, such as brightness or loudness (Berlyne, 1971, pp. 68-69), whereas ecological variables refer to previous associations and meanings (Berlyne, 1971, p. 69) such as a nostalgic memory triggered by a particular stimulus. Collative variables refer to evaluative properties of a stimulus, which Berlyne proposes will be collated and compared by a respondent. As a non-exhaustive list, collative variables include complexity, novelty/familiarity, change, conflict, surprisingness, uncertainty, interestingness and ambiguity (Berlyne, 1960, p. 44; 1971, p. 69). While Berlyne proposed that all three types of variables contribute towards aesthetic preference, his legacy is the notion that collative variables are the "most significant" determinants of preference (Berlyne, 1971, p. 69).

The biological driver of the inverted-U comes from the influence of two opposing mechanisms: a reward system that responds to initial increases of stimulus arousal, and an aversion system that gradually becomes activated beyond a critical level of arousal, and that opposes the effect of the reward system (Berlyne, 1969, p. 1068). With the collative variable of exposure, Berlyne argued that the reward system is activated by increasing arousal with exposure to a stimulus, but after a time the aversion system opposes this activation, leading to an increasing dominance of the aversion system as arousal continues to increase as a result of subsequent exposure. For this reason, Berlyne's experiments, and many of the experiments that followed him, have focused partly or entirely on the role collative variables play on preference. Two such variables have received particular attention: *complexity* and *familiarity*.

Berlyne argued that an organism is aroused by various aspects of the external environment, and that this arousal drives the approach and aversion systems of the organism. The amount of arousal a stimulus can invoke is referred to as *arousal potential* (Berlyne, 1960, 1971). Berlyne proposed that an intermediate level of arousal is preferred, and that preference is an index of this level of arousal. In other words, a stimulus exhibiting a low level of a collative variable, such as lacking familiarity, will be less preferred than a stimulus exhibiting a moderate level of this collative variable (a moderate level of familiarity). However, once familiarity (in this case) surpasses an optimal point, the level of preference will begin to decrease. This proposed inverted-U function was likened to the pattern outlined in Figure 1, also known as the Wundt curve (Berlyne, 1960, pp. 200-201; 1971, pp. 86-89). Wilhelm Wundt first proposed an inverted-U relationship between stimulus intensity and preference (Wundt, 1874, p. 468), although as Martindale and Moore point out (1989, p. 433), this was based solely on casual observations and it was not until decades later that the inverted-U was identified under controlled, experimental conditions.

Figure 1 here

Limitations of the inverted-U model

One of the controversies of the inverted-U model is that it is not clear what constitutes its replication. Some studies report a monotonic increasing relationship between preference and exposure, for example, and use this as evidence for the failing of the inverted-U trajectory. A well-known example of this is Zajonc's (1968) *mere exposure* effect, which proposes that additional exposure to a stimulus strictly increases preference towards it. Another important example is *prototypicality* as proposed by Martindale and colleagues (Martindale & Moore, 1988, 1989; Martindale, Moore, & Borkum, 1990; Martindale, Moore, & West, 1988), where preference is related to the ecological *typicality* of a stimulus in a monotonic increasing fashion, which can be primed (reinforced) with subsequent exposures. Martindale also proposed a monotonic increasing relationship between exposure and preference when the stimulus in question is perceived in temporally distributed presentations, rather than through massed exposure. Specifically, with massed exposure (i.e., listening to the same piece of music repeatedly and in close succession) the inverted-U pattern of preference emerges, but when a piece of music is exposed to the listener with large periods of time in between listenings, the inverted-U pattern may disappear (Martindale, 1984, 1988).

However, how can we be certain that preference can increase indefinitely as a function of exposure, even in the case of distributed exposure? Consider a thought experiment on an "infinitely" repeated, but distributed, exposure to a stimulus. The preference ratings should continue to increase. If it is rated on a scale that reflects the amount of stimulus preference then, according to the monotonic increasing model, eventually the rating scale will reach the peak value of the scale and cannot be increased upon subsequent exposure due simply to the limit imposed by the maximum value of the rating scale. Furthermore, we would need to demonstrate that the preference rating *never* declines despite repeated exposures. In the absence of such evidence, the inverted-U curve cannot be completely ruled out. The gradual increase in preference with additional exposure could simply be the realization of the opening segment of the curve as it makes its trajectory toward the optimal level of preference. Certain conditions (such as prototypicality and distributed exposure) may well slow its progress down considerably, but this does not demonstrate that the inverted-U model is *necessarily* inadequate. It could just mean that the conditions under which it appears were not satisfied. In other words, the inverted-U model can still be applied in the absence of evidence that a decline in preference after the preference peak *cannot* occur. Indeed, the inverted-U model can be seen as consisting of three segments (Berlyne, 1971, p. 194; 1974, p. 176; Heyduk, 1975, p. 84; Walker, 1973, p. 69). These three segments are (a) an increase in preference as discussed above, (b) a decrease in preference (the right-hand segment of the curve), which may be observed in specific cases such as where the familiarity of a stimulus is already high, and (c) a complete inverted-U trajectory. The first two segments of the curve are depicted in Figure 1 with dashed arrows.

Martindale also identified an "isohedonic trap" (Martindale, 1984; Martindale & Moore, 1989) where, according to the psychobiological theory, preference should be identical for any stimuli that produce the same level of arousal. For example, if the arousal generated by a piece of music, white noise, and a bright light produce the same level of arousal, they should each be liked the same amount. One way to limit this problem is to strictly consider stimuli that can be classified as being in the same modal category (in this case, music). As a further example of the controversy of the inverted-U model, consider the relationship *among* the collative variables complexity and preference. According to the model, there is an optimal point of complexity where preference will peak (illustrated in Figure 1). Subsequent exposures to a stimulus interact with subjective complexity to create a decrease in a respondent's subjective (perceived) level of complexity for that stimulus (Berlyne, 1974, p. 176; Heyduk, 1975, p. 84). That is, over time a repeated stimulus can become subjectively less and less complex. For a stimulus containing a subjectively low level of complexity below the optimal preference point for a particular respondent (such as Piece A in Figure 1), additional exposure should lead to a decrease in preference as the complexity level becomes even lower due to the additional exposures. Conversely, for a stimulus containing an initially high subjective level of complexity (such as Piece B in Figure 1), additional exposure should lead to an increase in preference as the level of subjective complexity is lowered towards the optimal point (Heyduk, 1975). From these examples, it is clear that a number of factors may influence the direction of aesthetic preference, and that a complete inverted-U curve may only be observed in specific circumstances such as when a large range of a collative variable is sampled. Importantly, one collative variable can interact with another, and so the inverted-U pattern may be hidden among the individual collative variables.

Another of Martindale's criticisms on the inverted-U is that collative variables are not the most important predictors of preference. He reported that, instead, typicality accounts for more variance than collative variables (Martindale & Moore, 1989; Martindale et al., 1990). However, as Konečni pointed out (1996, p. 131), regardless of a possible "overemphasis" on the relative importance of collative variables, the validity of the inverted-U can hardly be falsified on these grounds. Essentially, while prototypicality presents another important development in understanding the psychological foundations of music preference, it does not necessarily replace the inverted-U explanation. If all variables, including level of prototypicality, are held constant or controlled, the inverted-U relationship may still provide a satisfactory explanation of preference.

Alternatives to the inverted-U model

Researchers in the in the decades since the 1970s and 1980s have noted various limitations of traditional experimental aesthetics, such as the use of laboratory conditions and abstract stimuli that hold little relevance to realistic music experiences (Gardner, 1974, p. 208; Martindale et al., 1990, p. 54; North & Hargreaves, 1996b, p. 535). Generally speaking, the field has expanded from the traditional experiments characterized by Berlyne and Zajonc towards ecologically richer ones, several of which we discuss here. Hargreaves and colleagues (Hargreaves, 2012; Hargreaves, MacDonald, & Miell, 2005; Hargreaves & North, 2010) have proposed one such alternate approach with a multicomponent theory, the Reciprocal-feedback model of responses to music (henceforth RFM), that includes Berlyne's collative variables as a determinant of preference alongside numerous other co-determinants. As a non-exhaustive list, other determinants in the RFM include prototypicality, spreading activation (see, also Schubert, Hargreaves, & North, 2014), and the specific context and situation (such as the social environment, and the current emotional and arousal state of the listener). Such an approach highlights the inverted-U model as a single, yet useful, part of a multifaceted picture.

One of the components of the RFM belonging to "contexts and situations" is the influence of choice in determining arousal. An important source of evidence for these uses of music is Konečni's (1982, pp. 500-501) theory describing music as a tool for *optimizing* mood and emotion. He argued that a person in a state of high arousal will prefer music inducing a low level of arousal, to "level out" their arousal to an optimal level. Conversely, a person in a state of low arousal is assumed to prefer music inducing a higher level of arousal. This idea has been expanded by Hargreaves and colleagues in the RFM, who noted situations (e.g., North & Hargreaves, 2000) in which a person may instead choose to match their *desired* arousal level with the musical stimulus. As an example, a person involved in high-arousal exercise may choose to listen to fast, loud music, inducing additional levels of high arousal, to help them reach their goal state. In other words, Hargreaves noted that music may be used as a tool to reach various levels of arousal, which may either contrast with or match the current arousal state of the listener. But these interactions between the arousal of the music and the desirable state of the individual do not directly discount the simple positive relationship between single collative variables and preference. Rather, they point to a factor that may manipulate the kind of music to which the individual chooses to listen. If one listens to a piece of music for the purpose of increasing their arousal level, does it mean that the same piece can be used indefinitely? Or could it be that, as with distributed exposure, it slows down the rate at which the optimal point of preference is retained? That is, might the listener using the piece to maintain a high arousal level not decide after an extended period of time to replace that piece with another, less frequently heard, yet still high-arousal piece? If so, the inverted-U relationship is not necessarily incorrect. Rather, it is complicated by and interacts with the influence of contextual desirability.

A second approach we will discuss is *cognitive fluency*, which associates increased processing ease with increased hedonic response. Fluency is characterized by the speed and accuracy with which a stimulus is processed, and this may be influenced by variables such as priming, presentation duration, and repetition (Belke, Leder, Strobach, & Carbon, 2010; Leder, Belke, Oeberst, & Augustin, 2004; Reber, Schwarz, & Winkielman, 2004). Cognitive fluency has recently been examined in relation to repetitive lyrics in popular music (Nunes, Ordanini, & Valsesia, 2015). A distinction must be made between repetition of multiple exposures to an entire stimulus, being the approach frequently used to investigate preference for stimuli, and the internal (intraopus) repetition examined by Nunes et al. In their first two studies, Nunes et al. reported a positive relationship between lexical repetition and fluency, as predicted from their review of the literature. Lexical repetition was manipulated through multiple versions of the same piece, with different versions containing increasing amounts of lyrical repetitiveness, while fluency was determined with ratings of novelty. For study three, which was split into three parts, Nunes et al. investigated the impact of lexical repetition on chart performances of pieces from *Billboard*'s Hot 100 singles charts between 1958 and 2012. For part A, a total of 2048 pieces were compared from two categories: those that reached the #1 spot, and those that never climbed above #90. Increased lexical repetition was reported to significantly increase the chances of a piece belonging to the #1 category. Part B investigated whether lexical repetition influenced the speed of chart ascension for 939 #1 pieces, measured in weeks. Greater repetition was reported to increase chart adoption speed. For the final part of study three the "initial adoption speed", rationalized as whether or not a #1 piece debuted in the Top 40 (being an indicator of exceptional commercial success), was assessed for 149 pieces. Repetition (used to manipulate fluency) was reported to facilitate faster success up to an optimal point, after which increased fluency nullified the positive effect. As the authors note (Nunes et al., 2015, p. 196) this result is reminiscent of Berlyne's inverted-U. As such, we can see evidence of an inverted-U relationship with collative variables persevering within the context in more recent, ecologically based approaches.

To summarize, much recent literature has overtly or implicitly rejected Berlyne's inverted-U relationship for preference, while arguing for the need to consider a more multifaceted approach to aesthetic (and therefore music) preference (e.g., Leder & Nadal, 2014). As Levitin notes (2006, p. 240), the inverted-U model was not intended as a singular, comprehensive explanation in which collative variables are the only reason for which one might like or dislike a piece of music. Rather, the model may be best thought of as a general trend allowing a broad application across music contexts. The review above also points out that the contribution of collative properties may still be a reliable, possibly significant contributor to the narrative. As such, it may be fruitful to examine the literature in the context of the inverted-U model preference as a function of collative variables. We do this by conducting a review of the literature.

Literature Review

We wanted to test whether studies in music psychology have observed results compatible with one or more of the three segments of the inverted-U, even if not explicitly reported. Our approach was to review the literature on aesthetic preference of music across a long time period, and analyze those findings with respect to the three segments of the inverted-U curve. Two critical categories were then identified: whether the data from a study on preference and a collative variable could be explained by at least one of the three segments, or not (i.e., cannot be explained at all by the inverted-U model). With such an approach, we expected to be able to answer the question of whether the inverted-U model itself still provides a good explanation for preference in terms of collative variables, with all other variables controlled for or held constant.

Method

Procedure

The inclusion criteria for the literature to be reviewed were as follows: (a) the article needed to collect data on preference ratings or equivalent (liking, enjoyment, etc.) in response to at least one auditory stimulus – limiting stimuli to one modal category addresses Martindale's concern with regard to the isohedonic trap; (b) the independent variable needed to include at least one collative variable, or allow extraction of such a variable from the published data;[1] (c) studies published between the years 1901 and 2015 (inclusive) were analyzed. Literature was identified using various combinations of general and keyword searches, such as "music", "preference", "aesthetics", and "inverted-U". Analysis of articles cited in these papers and citation indexes were also accessed to ensure as broad a spread as possible of papers that satisfied the inclusion criteria.

Coding

Studies presenting data on preference as a function of one or more collative variables were coded according to one of five categories for each independent variable, as shown in Table 1. The first three categories are regarded as commensurate with some part or all of the inverted-U, as discussed above, and the remaining two categories (4 and 5) are regarded as contrary to the inverted-U model:

[1] A number of studies were excluded due to insufficient data reported (e.g., Downey & Knapp, 1927; Flowers, 1980; Pereira, Teixeira, Figueireido, Xavier, & Brattico, 2011).

1. Preference increases (denoted in the table with ↗).
2. Preference decreases (denoted in the table with ↘).
3. Preference increases, then decreases in an inverted-U shape (denoted in the table with ∩).
4. Preference decreases, then increases in a standard-U shape (denoted in the table with ∪).
5. Preference remains flat. This may be either (a) a statistically static result, or (b) cases with non-significant results, where no specific pattern(s) can be determined (both denoted in the table with **–**).

Table 1 here.

Results

Re-analyzed data

We performed additional analysis on three studies included in the review. The first of these (Gilliland & Moore, 1924) published the raw data of enjoyment ratings on a 10-point scale by 35 participants after the 1st and 25th hearing of four pieces of music (two classical pieces, labeled here as A and B, and two popular pieces, labeled here as C and D). Mean preference ratings were reported, however no inferential statistical analysis was performed. Reported mean preference scores (1st exposure, 25th exposure) increased for three pieces (M_A = 5.88, 6.09; M_B = 5.60, 6.94; and M_D = 4.37, 4.50) and decreased for the remaining piece (M_C = 5.00, 4.91)[2]. We re-analyzed these data (see Table 2), omitting data for one participant with incomplete responses. Our re-analysis revealed that

[2] Gilliland and Moore (1924) did not report *SD* values.

only Piece B significantly increased in preference score according to a paired samples *t*-test (coded in Table 1 as category 1), while the other three pieces did not change mean enjoyment score significantly at $p = .05$ (coded in Table 1 as category 5). Some studies have strictly cited the results of Gilliland and Moore's article as producing a positive relationship between preference and familiarity (e.g., Bornstein, 1989, p. 265; Finnäs, 1989, p. 12; Hunter & Schellenberg, 2011, p. 175). With this re-analysis in mind, however, the results of three of the four pieces do not support a strictly increasing relationship. Subsequent literature has reported that peak enjoyment may occur after much fewer than 25 exposures (Bartlett, 1973; Getz, 1966; Heyduk, 1975; Krugman, 1943; K. C. Smith & Cuddy, 1986), which suggests that reports of enjoyment after an intermediate number of exposures, for example at the third, sixth or ninth exposure, may have revealed an inverted-U pattern, but after a larger number of exposures, preference for all pieces were returning towards the initial enjoyment level. Hence an inverted-U relationship cannot be eliminated by this study either.

Paired samples *t*-tests were also performed on the data reported by Mull (1957), who collected preference responses on a seven-point scale from −3 to +3 in response to the works of Hindemith and Schoenberg made by 16 participants. Despite the non-significant *p*-value for the Hindemith stimulus, we have coded both stimuli as '1' due to the effect size, with non-significance possibly attributable to the small number of participants (see Table 2 for details[3]). Our final additional statistical analysis was performed on a study by Hamlen and

[3] The pieces are listed in Table 2 as they are reported by Mull (1957), although in the original publication both stimuli appear to contain misspellings of the work numbers. The Schoenberg stimulus most likely refers to the String Quartet No. 3, Op. 30. The Hindemith stimulus may refer to either the String Quartet No. 4, Op. 22, or the String Quartet No. 5, Op. 32.

Shuell (2006). This study contained groups of participants that were exposed to classical music, either with or without accompanying visual material. The majority of inferential tests reported in this study focused on differences between the audio or audio-visual groups, and as such the results were difficult to code with respect to the simple effects of music preference and the collative variable. Consequently, in our reanalysis, the audio and audio-visual condition responses were collapsed. This re-analysis produced a positive correlation between preference and ratings of familiarity ($r(24) = .866$, $p < .001$), and was coded as ‘1’.

Table 2 here.

Variables captured by the review of the literature

Fifty-seven studies were identified that satisfied the selection criteria. Of these, 54 (95%) tested the variables complexity or familiarity, or both. Of the three remaining studies, one examined general arousal (North & Hargreaves, 1997b), the second examined age brackets of children with the assumption that increasing age would hold a positive correlation with stylistic familiarity for common examples of music (Hargreaves & Castell, 1987), and the third study examined the participants’ level of musical awareness (training and knowledge) of Asian-Indian music, in relation to preference for stimuli of the same nature (Erdmonston, 1969). Erdmonston’s study does not strictly report a collative variable; however this variable is related to stylistic familiarity and expectation. Thirty-three (58%) studies tested familiarity alone. Two primary variables relating to familiarity were identified: (a) studies using explicit ratings of familiarity, identified in 16 studies (28%); (b) studies assuming an increase of familiarity with subsequent exposures (at times referred to as repetition), identified in 20 studies (35%). Additionally, two studies reported ratings of familiarity/exposure in terms of radio "plugging" (Erdelyi, 1940; Wiebe, 1940), one study equated familiarity/exposure with the amount of time a stimulus spent in the music charts (Eerola & North, 2000), and two studies examined "stylistic familiarity", each referring to the variable as genre specific (Hargreaves & Castell, 1987; Shehan, 1985). Fourteen studies (25%) tested the influence of complexity alone. Finally, three studies investigated complexity and uncertainty (Bragg & Crozier, 1974; Crozier, 1974; Martindale & Moore, 1989).

Stimuli consisted of both abstract and realistic (i.e., musical, or ecologically plausible) types of auditory stimuli. The realistic types included music from a range of styles, time periods and cultures, chord progressions, and "cover versions", in which existing pieces were changed to different musical styles. The abstract stimuli types included tone sequences and rhythmic sequences, some of which were randomly generated.

Main findings

Fifty studies (88%) were coded exclusively in categories 1, 2 and 3 (either exclusively in one category, or a mixture thereof). These 50 studies are therefore interpreted as part of the over-arching, segmented inverted-U model. Of these 50 studies, 16 (28% of the total studies) were coded exclusively in category 1, four studies (7% of the total studies) were coded exclusively in category 2, and 15 studies (26% of the total studies) were coded exclusively in category 3. This left

15 remaining studies that contained a mixture of categories 1, 2, and/or 3. However, it must be noted that only the 15 studies coded exclusively in category 3 can be considered as *genuine* support for Berlyne's theory. While the remaining 35 studies contain results compatible with the segments of the inverted-U, they do not definitively support the model and we therefore refer to these as *secondary* supporting studies.

Of the seven remaining studies (12%), five studies (9% of the total studies) were coded as a mixture of categories 1–3 and 4–5 (i.e., a mixture of compatible and incompatible results), and two studies (3% of the total studies) were coded as exclusively incompatible with the three segments of the inverted-U model (i.e., coded exclusively as category 5). The chronological distribution of all included studies is shown as a histogram with a five-year bin size in Figure 2. The plot also presents a visual chronological overview of the number of studies according to four distinctions:

1. Genuine support for the inverted-U, found in 15 studies
2. Secondary support for the inverted-U, found in 35 studies
3. Mixed results, found in five studies
4. Incompatible results, found in two studies

Figure 2 here

Of the 16 studies coded exclusively into category 1 (preference rising as the positively framed collative variable level rises), 15 investigated manipulation of familiarity alone. Eight of these used explicit ratings of familiarity, whereas for six of these studies subsequent exposures were used as the independent variable, as previously outlined. The remaining two studies used radio plugging, and musical awareness, which as previously noted is related to stylistic familiarity. These results could suggest that the collative variable familiarity has a stronger tie to the first, increasing slope of the inverted-U curve than complexity, as consistent with Zajonc, Crandall, Kail, and Swap (1974, p. 688). However, it is also possible that these studies did not expose the stimuli to the participants enough times to surpass the optimal level of familiarity. When examining the number of exposures of these 16 studies, one study used an ambiguous radio plugging period of 13 weeks (Erdelyi, 1940). Of the remaining 15 studies, seven used only a single exposure, and another four studies used five exposures or fewer. As noted by Berlyne (1974), Heyduk (1975), and Walker (1973) in the introduction to this chapter, studies using only a small sample of a collative variable can be expected to produce monotonic increasing or decreasing results.

Of the four studies (7%) coded exclusively as category 2 (preference decreasing with an increase in a collative variable), three of these manipulated complexity as the independent variable (Eerola & North, 2000; Russell, 1982; J. D. Smith & Melara, 1990), and the fourth reported decreasing preference as a function of increasing familiarity (Cui, Collett, Troje, & Cuddy, 2015). Category 2 appeared alongside other categories in 11 studies (making a total of 15 studies; 26% of the total studies).

Examination of the types of statistical analyses used

Of the overall 57 studies, 32 (56%) only used linear analysis methods, such as a single correlation analysis or a single *t*-test. In comparison, only 14 studies (25%) included inferential non-linear analysis (13 of these included both linear and non-linear analyses, while the remaining study used a solely curvilinear analysis). Of these 14 studies, eight produced significant curvilinear results, three produced non-significant curvilinear results, and three produced a mixture of significant and non-significant curvilinear results. The implication of such a relatively small percentage of studies using non-linear analysis methods is that a number of significant quadratic results may be hidden in the data of the other, linear-only analysis studies. With this in mind, the number of studies classified as concave down quadratic (inverted-U, category 3) may have been considerably underestimated. Furthermore, of the 16 studies coded exclusively as category 1, only one study (North & Hargreaves, 1997a) used a non-linear analysis. This category, which was the most frequently represented in our analysis, may have been considerably smaller had additional non-linear methods been employed.

Conclusion

The literature review identified 57 experiments investigating the relationship between music preference and one or more collative variables—typically complexity and familiarity/exposure—that could be interpreted through predictions made by Berlyne's inverted-U model. Categorization of results showed 50 of these studies as compatible with an overarching inverted-U theory consisting of three possible segments, and 15 of these 50 studies producing strictly inverted-U results. Furthermore, in contrast to the narrative portrayed by several reviews, the number of studies in which genuine or secondary supporting results were reported or identified in the last 25 years (1990–2015, 22 studies) is a relatively similar value to the number of studies reported in the 30 years before that, the heyday of Berlyne's theory and peak influence (1960–1989, 28 studies). If the results are so consistent with Berlyne's overarching model, why have some of the most influential reviews of music preference been dismissive of Berlyne's ideas?

First, it must be reiterated that Berlyne and others clearly identified the inverted-U as consisting of up to three segments of a curve, as noted in our introduction. However, a number of articles reporting monotonic results (e.g., Bradley, 1971; Heingartner & Hall, 1974; Lieberman & Walters, 1968) have solely been interpreted as rejecting the inverted-U (either by the authors, or in subsequent reviews) rather than supporting both monotonic increase *and* a segment of an overarching inverted-U relationship. Second, the majority of articles in this review were limited to strictly linear analyses meaning that quadratic relationships may have remained hidden in the data. This may be expected of studies conducted prior to Berlyne's work, for example those which measured preference at only the first and last exposures (e.g., Gilliland & Moore, 1924; Mull, 1957). However, preference needs to be measured so as to produce a reasonable degree of variance, over at least three points of the collative variable under investigation—for example several times (at least three) over the course of the exposure period—for a curvilinear relationship to be identifiable, should one exist. Regardless, the majority of post-Berlyne studies have not included curvilinear analyses. Third, the linear relationships may have been a result of insufficient variation in the collative variable. That is, the insufficient variation of the collative variable (not capturing very low, intermediate and very high levels) did not allow the inverted-U to fully emerge.

The conclusion drawn from our analysis of the literature is that the inverted-U explanation of preference as a function of collative variables is a robust manner of explaining data, in particular when all other variables are held constant or controlled. It may be that because of the strong association of the inverted-U model with Berlyne's psychobiological theory, the inverted-U model has been judged guilty by association. In other words, the theory's fundamental reliance on the concept of arousal may be at the heart of the demise of interest in the collative variable. The concept of arousal has generated extraordinary confusion in the literature. Most researchers in the reviewed studies equate, either explicitly or implicitly, arousal to preference. Landers (1980), for example, noted that:

> The intensity level of behavior is termed arousal. The construct of arousal, which is often used interchangeably with other intensity-related terms such as drive, tension, and activation, refers to the degree of energy release of the organism, which varies on a continuum from deep sleep to high excitement (p. 77).

A contradiction is immediately evident between the arousal as an excitement indicator versus excitement as a preference indicator, because it assumes that enjoyment of low-arousal (low-excitement) activity is not possible. Furthermore, both of these definitions indicate degrees of neglect of the psychobiological origins of arousal, where arousal refers to neurobiological activity, but may also refer to wakefulness (for a more detailed discussion, see Mashour & Alkire, 2013). The link Berlyne makes between arousal, in the neurophysiological (biological) sense, and preference is a theoretical one upon which psychobiological theory hinges.

The concept of arousal stemmed from an attempt to simplify investigations of the poorly defined phenomena of emotion, drives and motives (Neiss, 1988). Arousal packaged in the psychobiological theory provided a pathway to the discovery of the illusive biological mechanism responsible for generating pleasure and preference. However, the logic of the pathway was unsustainable. Modern neuroscience has to a large extent filled the gaps left by the demise of arousal (and other) theories of preference through the formulation of the "reward system" and pleasure/reward-inducing neurotransmitters (e.g., Blood & Zatorre, 2001; Chanda & Levitin, 2013; Huron, 2001; Schultz, 2015; Yager, Garcia, Wunsch, & Ferguson, 2015). What remain, in regards to music preference at least, are the data, which happen to be well explained by the inverted-U model, but which need something more fashionable than the psychobiological theory to understand them.

In conclusion, our analysis of the literature calls for a reassessment and refinement of the view that Berlyne's theory is inadequate or no longer relevant. Research outputs continue to validate the overarching inverted-U model, and the confusion between the model and the theory should be balanced with the actual data at hand. We therefore recommend that Berlyne's inverted-U model of preference for collative variables be accepted as a well-established explanation, rather than a dated view to be brushed aside, at least until rigorous research can control collative variables in such a way as to allow proper falsification. But until that occurs, the state of the art must be that preference varies with a collative variable to form some or all segments of an inverted-U curve, provided all other variables are held constant.

References

Bartlett, D. L. (1973). Effect of repeated listenings on structural discrimination and affective response. *Journal of Research in Music Education, 21*, 302-317.

Belke, B., Leder, H., Strobach, T., & Carbon, C. C. (2010). Cognitive fluency: High-level processing dynamics in art appreciation. *Psychology of Aesthetics, Creativity, and the Arts, 4*(4), 214-222.

Berlyne, D. E. (1960). *Conflict, arousal and curiosity*. New York: McGraw-Hill.

Berlyne, D. E. (1967). Arousal and reinforcement. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 1-110). Lincoln: University of Nebraska Press.

Berlyne, D. E. (1969). Arousal, reward and learning. *Annals of the New York Academy of Sciences, 159*, 1059-1070.

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics, 8*(5A), 279-286.

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

Berlyne, D. E. (Ed.) (1974). *Studies in the new experimental aesthetics*. Washington, DC: Hemisphere.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *PNAS, 98*(20).

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin, 106*(2), 265-289.

Bradley, I. L. (1971). Repetition as a factor in the development of musical preferences. *Journal of Research in Music Education, 19*(3), 295-298.

Bragg, B. W. E., & Crozier, J. B. (1974). The development with age of verbal and exploratory responses to sound sequences varying in uncertainty level. In D. E. Berlyne (Ed.), *Studies in the new experimental aesthetics* (pp. 91-108). New York: Wiley.

Brittin, R. V. (1996). Listener's preference for music of other cultures: Comparing response modes. *Journal of Research in Music Education, 44*(4), 328-340.

Burke, M. J., & Gridley, M. C. (1990). Musical preferences as a function of stimulus complexity and listener's sophistication. *Perceptual and Motor Skills, 71*, 687-690.

Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in Cognitive Sciences, 17*, 179-193.

Crozier, J. B. (1974). Verbal and exploratory responses to sound sequences varying in uncertainty level. In D. E. Berlyne (Ed.), *Studies in the new experimental aesthetics* (pp. 27-90). New York: Wiley.

Cui, A. X., Collett, M. J., Troje, N. F., & Cuddy, L. L. (2015). Familiarity and preference for pitch probability profiles. *Cognitive Processing, 16*(2), 211-218.

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: Books For Libraries Press.

Eerola, T., & North, A. C. (2000). *Expectancy-based model of melodic complexity.* Paper presented at the Proceedings of the Sixth International Conference on Music Perception and Cognition, Keele, Staffordshire.
Erdelyi, M. (1940). The relation between 'radio plugs' and sheet sales of popular music. *Journal of Applied Psychology, 24*, 696-702.
Erdmonston, W. E. J. (1969). Familiarity and musical training in the esthetic evaluation of music. *The Journal of Social Psychology, 79*, 109-111.
Finnäs, L. (1989). How can musical preferences be modifed? A research review. *Bulletin of the Council for Research in Music Education, 102*, 1-58.
Flowers, P. J. (1980). Relationship between two measures of music preference. *Contributions to Music Education, 8*, 47-54.
Gardner, H. (1974). Aesthetics and psychobiology by D.E. Berlyne. Review by: Howard Gardner. *Curriculum Theory Network, 4*(2/3, Curriculum in Art), 205-211.
Getz, R. P. (1966). The effects of repetition on listening response. *Journal of Research in Music Education, 14*, 178-192.
Gilliland, A. R., & Moore, H. T. (1924). The immediate and long-time effects of classical and popular phonograph selections. *Journal of Applied Psychology, 8*(3), 309-323.
Gordon, J., & Gridley, M. C. (2013). Musical preferences as a function of stimulus complexity of piano jazz. *Creativity Research Journal, 25*(1), 143-146.
Hamlen, K. R., & Shuell, T. R. (2006). The effects of familiarity and audiovisual stimuli on preference for classical music. *Bulletin of the Council for Research in Music Education, 168*, 21-34.
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. (2012). Musical imagination: Perception and production, beauty and creativity. *Psychology of Music, 40*(5), 539-557.
Hargreaves, D. J., & Castell, K. C. (1987). Development of liking for familiar and unfamiliar melodies. *Bulletin of the Council for Research in Music Education, 91*, 65-69.
Hargreaves, D. J., MacDonald, R. A. R., & Miell, D. (2005). How do people communicate using music? In D. E. Miell, R. A. R. Macdonald, & D. J. Hargreaves (Eds.), *Musical communication* (pp. 1-25). Oxford: Oxford 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: Oxford University Press.
Heingartner, A., & Hall, J. V. (1974). Affective consequences in adults and child repeated exposure to auditory stimuli. *Journal of Personal and Social Psychology, 29*, 719-723.
Heyduk, R. G. (1975). Rated preference for musical compositions as it relates to complexity and exposure frequency. *Perception & 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*, 175-179.

Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences, 930*, 43-61.
Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's syndrome patients acquire affective reactions? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 22-36.
Johnston, R. R. (2016). The effect of repetition on preference ratings for select unfamiliar musical examples: Does preference transfer? *Psychology of Music, 44*(3), 514-526.
Konečni, V. J. (1982). Social interaction and musical preference. In D. Deutsch (Ed.), *The Psychology of music*. New York: Academic Press.
Konečni, V. J. (1996). Daniel E. Berlyne (1924-1976): Two decades later. *Empirical Studies of the Arts, 14*(2), 129-142.
Krugman, H. E. (1943). Affective response to music as a function of familiarity. *Journal of Abnormal and Social Psychology, 38*, 388-392.
Landers, D. M. (1980). The arousal-performance relationship revisited. *Research Quarterly for Exercise and Sport, 51*, 77-90.
Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgements. *British Journal of Psychology, 95*(4), 489-508.
Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode-developments and challenges in empirical aesthetics. *British Journal of Psychology, 105*(4), 443-464.
Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession*. New York: Dutton.
Lieberman, L. R., & Walters, W. M. J. (1968). Effects of repeated listening on connotative meaning of serious music. *Perceptual and Motor Skills, 26*, 891-895.
Martindale, C. (1984). The pleasures of thought: A theory of cognitive hedonics. *Journal of Mind and Behavior, 5*, 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: Praeger Publishers.
Martindale, C., & Moore, K. (1988). Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance, 14*(4), 661-670.
Martindale, C., & Moore, K. (1989). Relationship of musical preference to collative, ecological, and psychophysical variables. *Music Perception: An Interdisciplinary Journal, 6*(4), 431-445.
Martindale, C., Moore, K., & Borkum, J. (1990). Aesthetic preference: Anomalous findings for Berlyne's psychobiological theory. *The American Journal of Psychology, 103*(1), 53-80.
Martindale, C., Moore, K., & West, A. (1988). Relationship of preference judgments to typicality, novelty, and mere exposure. *Empirical Studies of the Arts, 6*(1), 79-96.
Mashour, G. A., & Alkire, M. T. (2013). Evolution of consciousness: Phylogeny, ontogeny, and emergence from general anesthesia. *Proceedings of the National Academy of Sciences, 110*(Supplement 2), 10357-10364.

McMullen, P. T. (1974). Influence of number of different pitches and melodic redundancy on preference responses. *Journal of Research in Music Education, 22*, 198-204.
McMullen, P. T., & Arnold, M. J. (1976). Preference and interest as a function of distributional redundancy in rhythmic sequences. *Journal of Research in Music Education, 24*, 22-31.
Meyer, M. (1903). Experimental studies in the psychology of music. *The American Journal of Psychology, 14*(3/4), 192-214.
Mull, H. K. (1957). The effect of repetition upon the enjoyment of modern music. *The Journal of Psychology, 43*(1), 155-162.
Neiss, R. (1988). Reconceptualizing arousal: Psychobiological states in motor performance. *Psychological Bulletin, 103*(3), 345-366.
North, A. C., & Hargreaves, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology, 14*, 77-93.
North, A. C., & Hargreaves, D. J. (1996a). Responses to music in a dining area. *Journal of Applied Social Psychology, 26*(6), 491-501.
North, A. C., & Hargreaves, D. J. (1996b). Responses to music in aerobic exercise and yogic relaxation classes. *British Journal of Psychology, 87*(4), 535-547.
North, A. C., & Hargreaves, D. J. (1997a). Liking for musical styles. *Musicæ Scientiæ, 1*(1), 109-128.
North, A. C., & Hargreaves, D. J. (1997b). Liking, arousal potential, and the emotions expressed by music. *Scandinavian Journal of Psychology, 38*, 45-53.
North, A. C., & Hargreaves, D. J. (2000). Musical preferences during and after relaxation and exercise. *The American Journal of Psychology, 113*, 43-67.
North, A. C., & Hargreaves, D. J. (2001). Complexity, prototypicality, familiarity, and the perception of musical quality. *Psychomusicology, 17*, 77-80.
Nunes, J. C., Ordanini, A., & Valsesia, F. (2015). The power of repetition: Repetitive lyrics in a song increase processing fluency and drive market success. *Journal of Consumer Psychology, 25*(2), 187-199.
Orr, M. G., & Ohlsson, S. (2001). The relationship between musical complexity and liking in jazz and bluegrass. *Psychology of Music, 29*, 108-127.
Pereira, C. S., Teixeira, J., Figueireido, P., Xavier, J., & Brattico, E. (2011). Music and emotions in the brain: Familiarity matters. *PLoS ONE, 6*(11), e27241. doi:10.1371/journal.pone.0027241
Peretz, I., Gaudreau, D., & Bonnel, A. M. (1998). Exposure effects on music preference and recognition. *Memory and Cognition, 26*(5), 884-902.
Radocy, R. E. (1982). Preference for classical music : A test for the hedgehog. *Psychology of Music*, Special Issue: Proceedings of the Ninth International Seminar on Research in Music Education, 91-95.
Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*(4), 364.
Russell, P. A. (1982). Relationships between judgements of the complexity, pleasingess and interestingness of music. *Current Psychological Research, 2*, 195-202.
Russell, P. A. (1986). Experimental aesthetics of popular music recordings: Pleasingness, familiarity and chart performance. *Psychology of Music, 14*, 33-43.

Russell, P. A. (1987). Effects of repetition on the familiarity and likeability of popular music recordings. *Psychology of Music, 15*, 187-197.
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.
Schubert, E. (2010). Affective, evaluative, and collative responses to hated and loved music. *Psychology of Aesthetics, Creativity, and the Arts, 4*(1), 36-46.
Schubert, E., Hargreaves, D. J., & North, A. C. (2014). A dynamically minimalist cognitive explanation of musical preference: Is familiarity everything? *Frontiers in Psychology, 5*(38).
Schuckert, R. F., & McDonald, R. L. (1968). An attempt to modify the musical preferences of preschool children. *Journal of Research in Music Education, 16*, 39-44.
Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews, 95*(3), 853-951.
Shehan, P. K. (1985). Transfer of preference from taught to untaught pieces of non-Western music genres. *Journal of Research in Music Education, 33*(3), 149-158.
Siebenaler, D. J. (1999). Student song preference in the elementary music class. *Journal of Research in Music Education, 47*(3), 213-223.
Silvia, P. J. (2005). What is interesting? Exploring the appraisal structure of interest. *Emotion, 5*(1), 89.
Smith, J. D., & Melara, J. M. (1990). Aesthetic preference and syntactic prototypicality in music: 'Tis the gift to be simple. *Cognition, 34*, 279-298.
Smith, K. C., & Cuddy, L. L. (1986). The pleasingness of melodic sequences: Contrasting effects of repetition and rule-familiarity. *Psychology of Music, 14*, 17-32.
Steck, L., & Machotka, P. (1975). Preference for musical complexity: Effects of context. *Journal of Experimental Psychology: Human Perception and Performance, 104*(2), 170-174.
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.
Tan, S. L., Spackman, M. P., & Peaslee, C. L. (2006). The effects of repeated exposure on liking and judgements of musical unity of intact and patchwork compositions. *Music Perception: An Interdisciplinary Journal, 23*(5), 407-421.
Teo, T., Hargreaves, D. J., & Lee, J. (2008). Musical preference, identification, and familiarity: A multicultural comparison of secondary students from Singapore and the United Kingdom. *Journal of Research in Music Education, 56*(1), 18-32.
Verveer, E. M., Barry, H., & Bousfield, W. A. (1933). Change in affectivity with repetition. *American Journal of Psychology, 45*, 130-134.
Vitz, P. C. (1966). Affect as a function of stimulus variation. *Journal of Experimental Psychology, 71*, 74-79.

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: Academic Press.

Washburn, M. F., Child, M. S., & Abel, T. M. (1927). The effect of immediate repetition on the pleasantness or unpleasantness of music. In M. Schoen (Ed.), *The effects of music: A series of essays* (pp. 199-210). New York: Books For Libraries Press.

Wiebe, G. (1940). The effect of radio plugging on student's opinions of popular songs. *Journal of Applied Psychology, 24*(6), 721-727.

Wundt, W. M. (1874). *Principles of physiological psychology* (E. B. Titchener, Trans.). Leipzig, Germany: Engelmann.

Yager, L. M., Garcia, A. F., Wunsch, A. M., & Ferguson, S. M. (2015). The ins and outs of the striatum: Role in drug addiction. *Neuroscience, 301*, 529-541.

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

Zajonc, R. B., Crandall, R., Kail, R. V. J., & Swap, W. (1974). Effect of extreme exposure frequencies on different affective ratings of stimuli. *Perceptual and Motor Skills, 38*, 667-678.

Table 1.
Categorization of reported preference results in relation to one or more collative variables.

Author(s) and year	Stimulus	Preference result i ii iii iv v	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants vi	Analysis type vii viii ix and Significance x xi	Comments
Bartlett (1973)	Excerpts of classical and popular music	↗ ↘	Familiarity (exposures)	ANOVA and Dunn's multiple comparison procedure	17	Liking	149	Li*	All classical music stimuli ↗, all popular music stimuli ↘
Bradley (1971)	Tonal, polytonal, atonal, and electronic music	↗	Familiarity (exposures)	T-test	3	Preference	14 classes (class size unspecified)	Li*	
Bragg & Crozier (1974)	Sequences of tones with 6 levels of uncertainty	∩	Uncertainty, and complexity	ANOVA	1	Exp. 1: Pleasingness. Exp. 2: binary choice between stimuli	72	Li*	
Brittin (1996)	Excerpts of Caribbean, African, Indian, and Papua New Guinean music	∩	Complexity (rated)	Two Pearson product-moment correlations	1	Preference	225	N-L*	A positive Pearson product-moment correlation up to an optimal point of complexity, and a negative correlation from this point
Burke & Gridley (1990)	Classical music, ranked in complexity by music professors	∩	Complexity	Visual inspection of means	1	Liking	40	n/a	

Author(s) and year	Stimulus	Preference result [i][ii][iii][iv][v]	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants [vi]	Analysis type [vii][viii][ix] and Significance [x][xi]	Comments
Crozier (1974)	Sequences of tones with 6 levels of uncertainty	∩	Uncertainty, and complexity	ANOVA with linear and quadratic trends, and "Duncan's new multiple range test"	1	Pleasingness	48	Li* and N-L*	Both dependent variables ∩
Cui, Collett, Troje, & Cuddy (2015)	Tone sequences varying in distinctiveness	↘	Familiarity (rated)	ANOVA	1	Preference	82	Li*	
Eerola & North (2000)	182 pieces by The Beatles	↘	Melodic complexity (as rated by the authors)	Pearson product-moment correlation	Music released 1962 - 1970	Popularity, measured by number of weeks and position in music charts	56	Li*	
Erdelyi (1940)	Popular music	↗	Familiarity (radio plugging)	Centroid of the rank-weighted week	13 week periods of radio plugging	Preference is measured by weekly sales rank	General Population	n/a	
Erdmonston (1969)	Asian Indian music	↗	Musical awareness (training and knowledge in the style of the stimuli, or a lack of it)	ANOVA and T-test	1	Musical appreciation and aesthetic evaluations	21	Li*	
Getz (1966)	Classical music	∩	Familiarity (exposures)	T-test	11	Preference	339	Li*	

Author(s) and year	Stimulus	Preference result i ii iii iv v	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants vi	Analysis type vii viii ix and Significance x xi	Comments
Gilliland & Moore (1924)	Classical and jazz music	↗ —	Familiarity (exposures)	Compared means for the first and last exposures. See Comments	25	Enjoyment value	35	Li* and Li (n.s.)	We performed a paired samples t-test using the data in this study. Tchaikovsky's *6th Symphony* ↗, while all other stimuli — (n.s.)
Gordon & Gridley (2013)	Jazz music, representing a range of complexity	∩	Complexity	Visual inspection of box plots	1	Liking	27	n/a	
Hamlen & Shuell (2006)	Classical music	↗	Familiarity (rated)	(See Comments)	1	Liking	127	Li*	We pooled the results from all three conditions and performed a Pearson product-moment correlation
Hargreaves (1984)	Classical, "easy-listening", popular, and avant-garde jazz music	↗ ∩ —	Familiarity (rated)	ANOVA	3 - 12	Liking	59; 40	Li*	Exp. 2 avant-garde stimulus — (statistically flat result), however this could also be interpreted as a small yet statistically significant increase
Hargreaves (1987)	Classical and popular music	↗	Familiarity (rated)	ANOVA and Pearson product-moment correlation	1	Liking, and quality	30	Li*	

Author(s) and year	**Stimulus**	**Preference result** i ii iii iv v	**Manipulated collative variable(s)**	**Inferential statistics**	**Exposures**	**Dependent variable**	**Number of Participants** vi	**Analysis type** vii viii ix **and Significance** x xi	**Comments**
Hargreaves & Castell (1987)	Carols, folk songs, nursery rhymes, and tone sequences noted as common or uncommon	↘ ∩	Age (proposed to hold a positive relationship with familiarity for well-known stimuli)	ANOVA	1	Liking	96	Li*	Familiar melodies ↘. Unfamiliar melodies also ↘, but ANOVA analysis only examined extreme ends of liking ratings. As such the final increase may not be significant. This code is not definitive
Heingartner & Hall (1974)	Pakistani folk music	↗	Familiarity (exposures)	ANOVA	1 - 8	Appealing	96; 54	Li*	Exp. 2: As the ANOVA analysis only reports significance between extreme points of ratings, it is not possible to determine whether the initial decrease in preference was significant. This code is not definitive
Heyduk (1975)	Self-composed examples of classical music, increasing in complexity	↗ ↘ ∩	Complexity (rated)	Chi-squared test reporting whether distribution of codes ↗, ↘ or ∩ versus code — (occurred by chance)	16	Liking	120	n/a	70.83% of *N* were observed following codes ↗, ↘, or ∩. The remaining *N* were classified as —, however these were omitted from the preference result column as the number reported was significantly less than would be expected from chance alone

Author(s) and year	Stimulus	Preference result i ii iii iv v	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants vi	Analysis type vii viii ix and Significance x xi	Comments
Hunter & Schellenberg (2011)	Excerpts of orchestral music, primarily from the Baroque, Classical, and Romantic periods	∩	Familiarity (exposures)	ANOVA with linear and quadratic trends	2, 8, or 32	Liking	79	Li (n.s.) and N-L*	
Johnson, Kim, & Risse (1985)	48 melodies from Korean pieces, performed in single notes on the piano	↗	Familiarity (exposures)	ANOVA	2, 6, or 11	Liking	24	Li*	Participants were either 1) alcoholic Korsakoff Syndrome patients; 2) alcoholic patients; 3) non-alcoholic patients
Johnston (2016)	10 music excerpts from the Romantic era	↗	Familiarity (exposures)	T-test	8	Preference	174	Li*	
Krugman (1943)	Classical and jazz music	∩	Familiarity (exposures)	Pearson product-moment correlation and visual inspection of means	8	Pleasantness	9	Li* and N-L	Only linear inferential statistics were performed; the N-L inspection was visual
Lieberman & Walters (1968)	Classical music	↗	Familiarity (exposures)	Chi-squared test	10	Pleasantness	32	Li*	Expected values for Chi-squared test are not reported. Analysis was conducted using counts of *N* whose ratings ↗, ↘, or —

Author(s) and year	Stimulus	Preference result [i][ii][iii][iv][v]	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants [vi]	Analysis type [vii][viii][ix] and Significance [x][xi]	Comments
Martindale & Moore (1989)	Artificial melodies and short musical themes	↗ ∩	Uncertainty, and complexity	ANOVA and Pearson product-moment correlation	1	Liking	34; 42	Li* and N-L (n.s.)	Exp. 1: ∩ with uncertainty. Exp. 2 ↗ with complexity
McMullen (1974)	Randomly generated melodies with varying numbers of pitches	↗ ↘	Melodic complexity, and melodic redundancy	Scheffé tests	1	Preference	82	Li*	Melodic complexity ↘, melodic redundancy ↗
McMullen & Arnold (1976)	Compound rhythmic sequences, varying in distributional redundancy	∩	Distributional redundancy	Friedman analysis of variance	1	Preference	35; 15	Li*	Distributional redundancy is an approximation of objective complexity, in which a higher amount of information is equivalent to a higher level of complexity
Meyer (1903)	A self-composed, micro-tonal instrumental piece performed on a reed organ	↗ ↘	Familiarity (exposures)	Descriptive statistics (comparison of individual preference ratings between the first and the last exposure)	12 - 15	Preference	14	n/a	Some participants returned for a second session
Mull (1957)	Classical music (works of Hindemith and Schoenberg)	↗	Familiarity (exposures)	Comparison of mean from first to last exposure, paired samples T-test	5	Liking	16	Li*	We performed a paired samples t-test using the data in this study.

Author(s) and year	Stimulus	Preference result [i ii iii iv v]	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants [vi]	Analysis type [vii viii ix] and Significance [x xi]	Comments
North & Hargreaves (1995)	Popular music	↗ ∩	Complexity (rated), and familiarity (rated)	Linear and quadratic regression analysis	1	Liking	75	Li* and N-L*	Complexity ∩, familiarity ↗
North & Hargreaves (1996a)	New age music varying in three levels of complexity	∩	Complexity (rated)	A one-way ANOVA and Tukey HSD tests	1	Liking	236	Li*	
North & Hargreaves (1996b)	Excerpts of new age and ambient house music, representing 5 levels of complexity	↗ ∩	Complexity (rated)	Linear and quadratic regression analysis	1	Liking	100	Aerobic group: Li* and N-L (n.s.). Yoga group: Li (n.s.) and N-L*	Aerobic group ↗, yoga group ∩
North & Hargreaves (1997a)	30 excerpts of well-known music, that was also identifiable as part of British music culture	↗	Familiarity (rated)	Linear and quadratic regression analysis	1	Liking	64	Li* and N-L (n.s.)	The result reported here is for Exp. 2. Exp. 1 did not meet the inclusion criteria for this review
North & Hargreaves (1997b)	Popular music	∩	Arousal (rated)	Linear and quadratic regression analysis	1	Liking	120	Li* and N-L*	

Author(s) and year	Stimulus	Preference result i ii iii iv v	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants vi	Analysis type vii viii ix and Significance x xi	Comments
North & Hargreaves (2001)	5 pieces by The Beatles, with each accompanied by 4 “cover versions” in other styles	↗ —	Complexity (rated), and familiarity (rated)	Partial correlations	1	Liking/quality	50	Li*	Complexity ↗, Familiarity —
Orr & Ohlsson (2001)	Jazz and bluegrass music	∩	Complexity (rated)	Linear and quadratic regression analysis	1	Liking	64; 151	Li* and N-L*	
Peretz, Gaudreu, & Bonnel (1998)	Melodic lines taken from “popular” and “unpopular” repertoire	↗	Familiarity (rated)	ANOVA	1 - 3	Liking	48	Li*	
Radocy (1982)	15 excerpts of classical music	↗ ∩	Complexity (rated), and familiarity (rated)	ANOVA with linear and quadratic trends	1	Preference	139	Li* and N-L*	Complexity ∩, familiarity ↗
Russell (1982)	A variety of post 1940 jazz styles	↘	Complexity (rated)	Partial correlation	1	Pleasingness	132	Li*	
Russell (1986)	Popular music	↗	Familiarity (rated)	Regression analysis and Pearson product-moment correlation	1	Pleasingness	428; 100	Li*	

Author(s) and year	Stimulus	Preference result [i ii iii iv v]	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants [vi]	Analysis type [vii viii ix] and Significance [x xi]	Comments
Russell (1987)	Popular music	—	Familiarity (see Comments)	Regression analysis, ANOVA	1	Pleasingness	97; 190	Li (n.s.)	Familiarity was rated, and reported as number of weeks spent in music charts. Both Exp. reported — (n.s.) results for pleasingness
Schellenberg, Peretz, & Vieillard (2008)	Classical melodies, performed as MIDI arrangements	↗ ∩	Familiarity (rated, and also over exposures)	ANOVA with significant quadratic component, and post hoc tests	2 - 32	Liking	108	Incidental group: Li* and N-L (n.s.). Focused group: Li* and N-L*	Incidental group ↗, focused group ∩
Schubert (2007)	Classical music	↗	Familiarity (rated)	Linear regression analysis	1	Liking	65	Li*	
Schubert (2010)	Excerpts of classical, jazz, popular, and rock music	∩	Familiarity (rated)	Visual inspection of z-scores	1	Liking	25	n/a	
Schuckert & McDonald (1968)	Classical and jazz music	—	Familiarity (exposures)	McNemar test for the significance of changes	5	Preference	24	Li (n.s.)	Participants were children 4-6 years old
Shehan (1985)	Western popular and classical, Asian Indian, African, Hispanic, and Japanese music	↗ ↘	Familiarity (exposures)	ANOVA, and reported means from the first and last exposures	1 or 5	Liking	26	Li*	All non-Western styles ↗, all Western styles ↘

Author(s) and year	Stimulus	Preference result [i][ii][iii][iv][v]	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants [vi]	Analysis type [vii][viii][ix] and Significance [x][xi]	Comments
Siebenaler (1999)	Popular music	↗	Familiarity (rated)	Pearson product-moment correlation	10	Liking	160	Li*	
Smith & Melara (1990)	76 chord progressions varying in complexity, amongst other variables	↘	Complexity (rated)	Linear and quadratic regression analysis	1 or 2	Pleasingness	69	Li* and N-L (n.s.)	
Smith & Cuddy (1986)	Sequences of tones	↗ ↘	Familiarity (exposures), and complexity (structure, corresponding with less complexity)	ANOVA	12 or 13	Pleasingness	36	Li*	Interaction: “with repetition, the point of optimal complexity shifted to the left on the structure scale.” (p. 29). Also complexity ↘. Exposure ↗ for 4 sets of stimuli, and ↘ for the final set
Steck & Machotka (1975)	Randomly generated tone sequences using a non-Western scale	↗ ↘ ∩	Complexity (note density and tempo)	Visual inspection of individual results	1	Liking	60	n/a	
Szpunar, Schellenberg, & Pliner (2004)	Monophonic tone sequences, and excerpts of orchestral music	↗ ∩ —	Familiarity (exposures)	ANOVA of linear and quadratic trends, with post hoc tests	2, 16, 32, or 64	Liking	50; 40	Exp. 1: IG Li* and N-L (n.s.), FG Li (n.s.) and N-L (n.s.); Exp. 2: IG Li* and N-L (n.s.), FG Li (n.s.) and N-L*	For Exp. 1, the incidental condition (IG) was ↗, and the focused condition (FG) was — (n.s). For Exp. 2 the IG was ↗, and the FG was ∩

Author(s) and year	Stimulus	Preference result [i ii iii iv v]	Manipulated collative variable(s)	Inferential statistics	Exposures	Dependent variable	Number of Participants [vi]	Analysis type [vii viii ix] and Significance [x xi]	Comments
Tan, Spackman, & Peaslee (2006)	Excerpts of classical piano solos - some unaltered and some created by linking unrelated excerpts together	↗ ↘	Familiarity (exposures)	ANOVA with linear trend	4	Liking	74	Li*	Overall results showed a linear increase in preference. When observed via stimulus group, patchwork stimuli ↗, intact stimuli ↘
Teo, Hargreaves, and Lee (2008)	Excerpts of Malay, Chinese, and Asian Indian music	↗	Familiarity (rated)	Pearson product-moment correlation	1	Liking	89	Li*	
Verveer, Barry, & Bousfield (1933)	Jazz music	∩	Familiarity (exposures)	Descriptive statistics (means)	8	Pleasantness	19	n/a	8 recorded ratings, exposure number not reported
Vitz (1966)	Sequences of tones, representing a range of complexity levels	∩	Complexity	Visual inspection of means	1	Pleasantness	36; 44	n/a	Exp. 2 and 3: all conditions ∩. Exp. 1 was not a study on preference
Washburn, Child, & Abel (1927)	Classical music and popular music from the 1920s	↗ ↘	Familiarity (exposures)	Descriptive analysis of ratios	5	Pleasantness	220	n/a	

Author(s) and year	**Stimulus**	**Preference result** i ii iii iv v	**Manipulated collative variable(s)**	**Inferential statistics**	**Exposures**	**Dependent variable**	**Number of Participants** vi	**Analysis type** vii viii ix **and Significance** x xi	**Comments**
Wiebe (1940)	Popular music	↗ ↘ —	Familiarity (radio plugging)	T-test	4 week exposure period of radio plugging	Liking	136	Li* and Li (n.s.)	Unplugged stimuli ↘, Less liked stimuli ↗. All other stimuli — (n.s.)

Note.

i Results are coded according to the categories reported in Section *Coding*. See also "Comments" column for further details. Preference is used here to mean the generic dependent variable related to liking, enjoyment, hedonic tone, etc. of the music stimuli. The specific term used to describe the preference variable in each study is indicated in the Dependent variable column.

ii ∩ inverted-U relationship for dependent variable as a function of the independent collative variable.

iii ↗ positive relationship.

iv ↘ negative relationship.

v — neither positive, negative nor inverted-U relationship found or reported ("Analysis type and Significance" column will indicate if inferential statistical analysis was performed and the result).

vi Number of participants in multiple experiments (Exp.) are separated by semicolon.

vii Li Linear analysis conducted (e.g. a single correlation analysis, a single *t*-test, an ordinary least squares regression analysis).

viii N-L Non-linear analysis conducted (any inferential statistical analysis that permits curvilinear interpretation of the data, such as polynomial regression, ANOVA with more than two levels of an independent variable, etc.).

ix n/a analysis was neither linear nor curvilinear. Alternatively, no relevant inferential statistics could be identified in the paper or deduced from the data presented in the paper, meaning that the results were usually based on visual inspection or descriptive statistics.

x * $p \leq .05$.

xi n.s. not statistically significant ($p > .05$).

Table 2.
Results of additional investigation using paired samples t-tests.

Gilliland and Moore (1924)						
Stimulus	1st exposure *M* (*SD*)	25th exposure *M* (*SD*)	*df*	*t*	*p*	*d*
Classical piece (A): Beethoven's *5th Symphony*	5.60 (1.69)	6.12 (1.78)	33	-1.673	.104	0.287
Classical piece (B): Tchaikovsky's *6th Symphony*	5.28 (2.17)	6.96 (1.81)	33	-4.24	<.001	0.733
Popular Piece (C): *That's it – a foxtrot*	4.85 (2.06)	4.91 (1.86)	33	-.129	.898	0.023
Popular piece (D): *Umbrellas to mend*	4.50 (2.11)	4.50 (2.03)	33	<.001	.999	<0.001
Mull (1957)						
Stimulus	1st exposure reported *M*	5th exposure reported *M*	*df*	*t*	*p*	*d*
Hindemith's *String Quartet IV, Op. 32*	-0.50	0.22	15	-2.008	.063	0.509
Schoenberg's *String Quartet III, Op. 31*	0.53	1.15	15	-3.796	.002	0.946

Note. *d* refers to Cohen's *d*.

The stimuli are listed here as reported by Mull (1957), however the original publication contains typographical errors in regards to the work numbers; see Footnote 3.

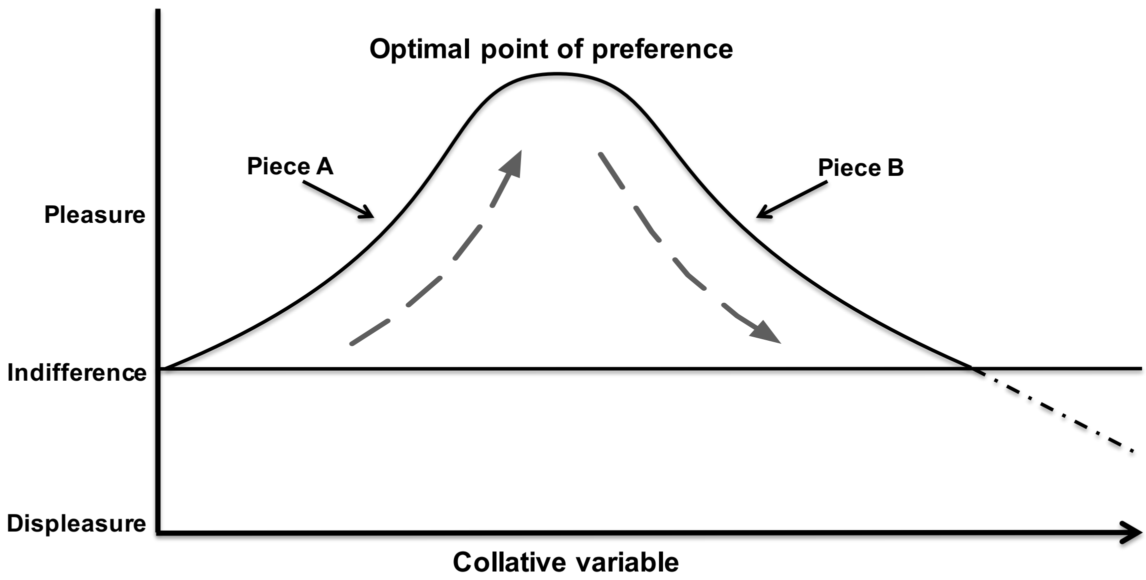

Figure 1. The Wundt curve, outlining Berlyne's proposed inverted-U relationship between preference and arousal. The first two "segments" of the curve are depicted with dashed arrows.

Note. This figure is based on one published by Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: Appleton-Century-Crofts (p. 89).

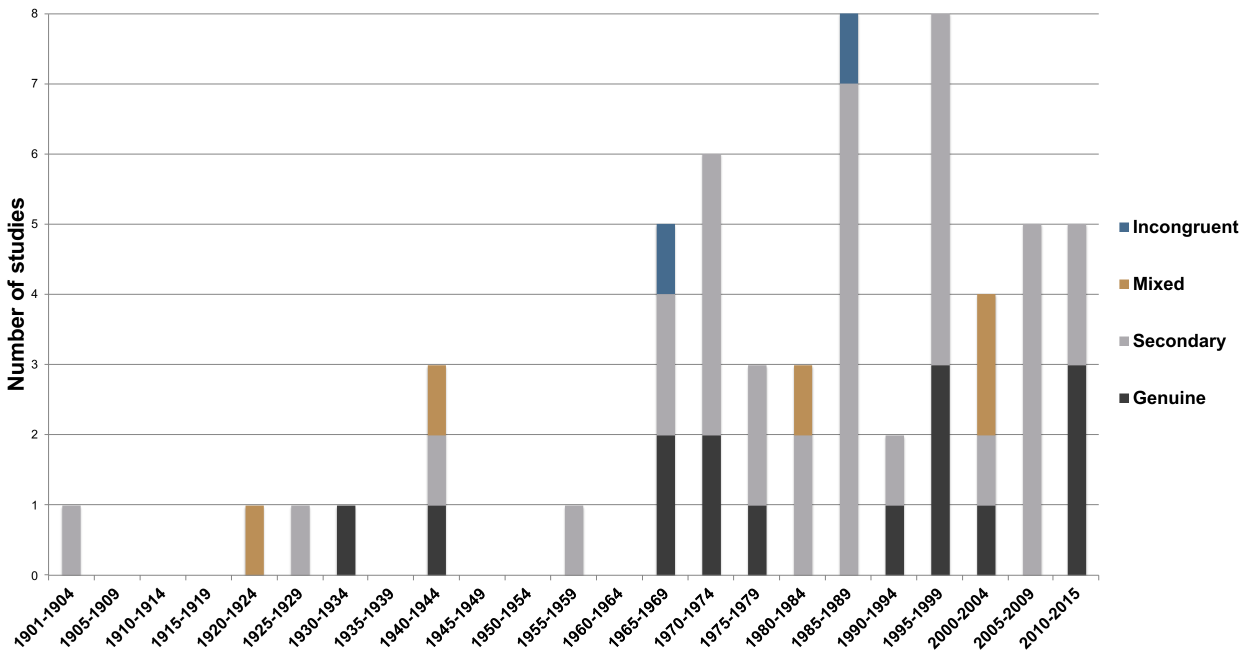

Figure 2. Count of studies by inverted-U evidence category and time period.