What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures

Alan S. Cowen[a,1,2], Xia Fang[b,c,1], Disa Sauter[b], and Dacher Keltner[a]

[a]Department of Psychology, University of California, Berkeley, CA 94720; [b]Department of Psychology, University of Amsterdam, 1001 NK Amsterdam, The Netherlands; and [c]Department of Psychology, York University, Toronto, ON M3J 1P3, Canada

Edited by Dale Purves, Duke University, Durham, NC, and approved December 9, 2019 (received for review June 25, 2019)

What is the nature of the feelings evoked by music? We investigated how people represent the subjective experiences associated with Western and Chinese music and the form in which these representational processes are preserved across different cultural groups. US (n = 1,591) and Chinese (n = 1,258) participants listened to 2,168 music samples and reported on the specific feelings (e.g., "angry," "dreamy") or broad affective features (e.g., valence, arousal) that they made individuals feel. Using large-scale statistical tools, we uncovered 13 distinct types of subjective experience associated with music in both cultures. Specific feelings such as "triumphant" were better preserved across the 2 cultures than levels of valence and arousal, contrasting with theoretical claims that valence and arousal are building blocks of subjective experience. This held true even for music selected on the basis of its valence and arousal levels and for traditional Chinese music. Furthermore, the feelings associated with music were found to occupy continuous gradients, contradicting discrete emotion theories. Our findings, visualized within an interactive map (https://www.ocf.berkeley.edu/~acowen/music.html) reveal a complex, high-dimensional space of subjective experience associated with music in multiple cultures. These findings can inform inquiries ranging from the etiology of affective disorders to the neurological basis of emotion.

emotion | affect | music | culture | semantic space

Central to the meaning of music are the subjective experiences that it evokes (1–3). Performers across cultures (4, 5) can reliably convey intense feelings with songs and instruments of different kinds and often do so by relying on acoustic features and associated percepts—pitch, loudness, pace—characteristic of the human vocal expression of emotion (6) and of speech (7). Music also modulates activity in brain regions implicated in emotion-related processing (8).

What is not well understood is how music evokes feelings in listeners. What is the taxonomic structure of the subjective experiences that music evokes? How do people represent their experiences when hearing music? To what extent are these processes consistent or variable across cultural groups? The present investigation addresses these questions, guided by a theoretical approach to subjective experience (9) and empirical and quantitative approaches (10).

A Semantic Space Approach to Subjective Experience

Music can evoke intense subjective experiences that people represent with concepts such as "awe," "joy," and "anger" (1–3). Within any realm—social interactions, the appreciation of art, or the daily fluctuations of mood, for example—subjective experiences are varied, blended, and complex (11). To account for such complexity, we have offered a semantic space approach that structures subjective experiences in terms of 3 properties: their conceptualization, dimensionality, and distribution.

The conceptualization of subjective experience concerns the nature of the concepts that most precisely characterize our feelings. This question is central to affective science (11). When people hear a moving or ebullient piece of music, do particular feelings—e.g., "sad," "fearful," "dreamy"—constitute the foundation of their experience? Or are such feelings constructed from more general affective features such as valence (pleasant vs. unpleasant) and arousal (9, 11–15)? The answers to these questions not only illuminate the nature of how music elicits subjective experiences, but also can inform claims about how the brain represents our feelings (8, 15), how infants learn to recognize feelings in themselves and others (16), the extent to which representations of subjective experience are universal across cultures (4–7), and the etiology of affective disorders (17, 18).

The dimensionality of subjective experience concerns the range of feelings that people experience. How many distinct feelings do we experience in response to music? Which of the feelings associated with music are dependent on the cultural background of the listener, and which are consistent across cultural groups?

Finally, there is the question of distribution: How are the feelings that music evokes distributed: Are they discrete clusters of states, or can they be blended together (11, 19)?

Together, the conceptualization, dimensionality, and distribution of subjective experience capture what we refer to as a *semantic space* of subjective experience. Studies drawing upon this conceptual framework have renewed our understanding of the feelings elicited by videos of real-world events (12) and attributed to facial expressions and vocal utterances (10, 20, 21).

Significance

Do our subjective experiences when listening to music show evidence of universality? And if so, what is the nature of these experiences? With data-driven methodological and statistical approaches, we examined the feelings evoked by 2,168 music excerpts in the United States and China. We uncovered 13 distinct types of experiences that people across 2 different cultures report in listening to music of different kinds. Categories such as "awe" drive the experience of music more so than broad affective features like valence. However, emotions that scientists have long treated as discrete can be blended together. Our results provide answers to long-standing questions about the nature of the subjective experiences associated with music.

Author contributions: A.S.C., D.S., and D.K. designed research; A.S.C. and X.F. performed research; A.S.C. contributed new reagents/analytic tools; A.S.C. analyzed data; and A.S.C., X.F., D.S., and D.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]A.S.C. and X.F. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: alan.cowen@berkeley.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1910704117/-/DCSupplemental.

First published January 6, 2020.

A Methodological Approach to Subjective Experience

Characterizing the semantic space of any response modality requires a departure from past approaches. The vast majority of research on subjective experience has focused on a small number of emotions (often just 4 to 6) (11). These experiences are typically elicited using a limited number of stimuli—often only 1 or 2—preselected to elicit more prototypical emotional responses, which are measured by asking participants to choose among 4 to 6 emotion categories or to report ratings of valence and arousal (and on occasion other appraisals) (11, 18). These methods preclude advances in understanding how people conceptualize subjective experience in terms of both categories and affective features (conceptualization), the range of feelings people experience (dimensionality), and the nature of the boundaries between specific feelings (distribution) (9–12).

With respect to the conceptualization of subjective experience, the use of small stimulus sets precludes comparing competing models of experience—for example, whether it is driven by specific feelings such as "amusement" or broad affective features such as valence. This is because with few stimuli that vary in many ways, many alternative conceptualizations become equivalent (that is, competing model features become collinear) (22). For example, categories such as "amusing" and more general features such as "valence" become statistically inseparable when studying just 1 or 2 positive emotion stimuli.

With respect to the dimensionality of subjective experience, the focus on a narrow range of states—usually 6—precludes examining how subjective responses may vary along a much wider array of dimensions (9–12). Does music only evoke the 6 emotions that are most widely studied? Or does music evoke a more complex array of feelings?

With respect to the distribution of subjective experience, the study of only prototypical exemplars of particular emotions makes it difficult to account for potential blends in subjective experience (9, 11, 12), which are likely to be associated with music (1–3, 10).

Understanding the nature of subjective experience therefore requires collecting responses to a vast array of rich and diverse stimuli, using multiple-response formats, targeting a wide variety of potentially distinct feelings, and including more than just prototypical elicitors or exemplars associated with each feeling. A recent investigation guided by these considerations analyzed US participants' responses to thousands of evocative videos of real-world scenes including landscapes, artwork, sports, animals, and injuries. With data-driven statistical techniques, it was found that self-report captures over 2 dozen dimensions, or kinds of subjective experience evoked by videos (12). The taxonomy of subjective experience is far richer than anticipated by traditional theories of emotion (11). However, because videos impart detailed semantic information (e.g., images of baby faces, spiders, or nature scenes), the extent to which participants' responses represented distinct subjective experiences versus distinct classifications of semantic content is unclear. Moreover, the study of subjective experience within a single culture leaves open the question of the extent to which the conceptualization, dimensionality, and distribution of subjective experience is preserved across cultures. In the present study, we move beyond earlier work by examining the subjective experiences elicited by instrumental music, which is largely free of semantic content (although not necessarily of learned semantic associations, as we will discuss), and by comparing reports of subjective experience across the United States and China. We do so to understand more abstract representations of subjective experience and how they may be preserved across multiple cultures.

The Feelings Associated with Music

Although it is abstract and nonrepresentational, instrumental music evokes rich subjective experiences. In response to music, people report feeling chills, shivers down their spine, laughter, tears, goose bumps, lumps in their throat, and rushes of adrenaline (23–27). However, we have yet to fully understand the different varieties of subjective experience associated with different kinds of music [or the distinctions between experienced and perceived feelings in music, which are difficult to capture (1, 28); see *Discussion* for implications for the present study]. Well-replicated studies have found that people within a cultural group can, with 70% or greater accuracy, label the feelings associated with small assortments of musical selections with 6 or fewer emotion categories—most notably, "anger," "happiness," "fear," "sadness," "surprise," and "tenderness" (1)—or along scales of valence and arousal (29).

Less is understood, however, about the broader structure of subjective experiences associated with music. Many studies have contrasted models that sort the feelings associated with music into 6 discrete emotion categories with models that represent these feelings as levels of valence and arousal (29, 30). However, because these are 2 among many possible ways that the feelings associated with music may be structured (11), the focus on these 2 models has precluded a richer understanding of the conceptualization, dimensionality, and distribution of the subjective experiences associated with music.

One exception to the focus on 6 emotions or valence and arousal is a study by Zentner et al. (3), in which factor analysis was applied to participants' self-reports of how often they had felt 89 feelings in their past experiences of listening to music from their preferred genre. This study identified 9 potentially distinct dimensions of subjective experience. However, the methods used in this study were retrospective and consequently subject to memory biases—for example, the potential confusion between the feelings associated with music and those associated with the contexts in which the music was heard. It was also subject to the limitations of factor analysis, which does not consider the reliability with which responses are associated with individual stimuli (*SI Appendix*, *SI Discussion* and Movie S1). This raises the question of how many kinds of subjective experience can reliably be associated with actual music excerpts in controlled or randomized contexts.

Moreover, a limited number of studies have examined the feelings associated with music in multiple cultures (5), which is critical to endeavors to understand potential universals, and their likely neurophysiological underpinnings, in responses to music (31). One study focused on 3 emotion categories (happiness, sadness, and fear) and found that they were associated with similar music in Westerners and participants from a remote African population, the Mafa of the Mandara Mountains in Cameroon (32). In contrast, more recent studies have uncovered cultural differences in feelings of pleasantness versus unpleasantness associated with music (33, 34). For example, the Tsimané of the Amazon rainforest do not perceive dissonance as unpleasant (34).

Overall, then, we have a limited understanding of the structure that people across cultures rely on to represent the feelings associated with music. To what extent do categories or broader affective features capture cultural universals in the feelings associated with music? What is the range of feelings that music conveys? Are the boundaries between the feelings associated with music discrete or continuous? Together, the conceptualization, dimensionality, and distribution of feelings associated with music comprise a taxonomy of its experiential effects, the foundation for investigations of how and why music moves us.

Experiment 1. Uncovering the Feelings Associated with Music

In our primary investigation, we derive a semantic space of subjective experience reliably associated with a broad range of Western music across 2 rather different cultures. We do so by

examining several hundred thousand judgments from participants in the United States (n = 1,011) and China (n = 895) (for discussion of cultural differences, see *SI Appendix*, *SI Discussion*, and refs. 35 and 36). Each participant heard subsets of 1,841 nonlyrical music samples that were gathered in an open-ended format. We applied data-driven statistical modeling techniques to these judgments to characterize the semantic space of subjective experience associated with music in each culture. By comparing the feelings associated with music in participants from China and the United States, we can ascertain the extent to which the conceptualization, dimensionality, and distribution of subjective experiences are preserved across these 2 different cultures.

At stake in the study of the semantic space of subjective experience are answers to questions of central theoretical import (9, 11). In terms of the conceptualization of subjective experience, with what consistency across cultures are responses to music represented using specific categories of subjective experience, such as awe and fear, and information about broad affective features, such as valence and arousal? With respect to the dimensionality of subjective experience, how many distinct varieties of feelings associated with music are preserved across cultures? In terms of the distribution of subjective experiences across cultures, do feelings occupy discrete clusters—families of related states such as anxiety and fear—or do they lie along continuous gradients?

To answer these questions, we collected judgments from participants in the United States and China employing the richest stimulus set of Western music samples ever studied, both in terms of the variety of different feeling states and affective features conveyed and in terms of the diversity of exemplars associated with each feeling, ideal for deriving a semantic space of subjective experience.

Participants from the United States were recruited on Amazon Mechanical Turk. Participants from China were recruited from a multi-institutional participant pool by X.F. To collect a library of music, 111 US participants (41 women, mean age = 32) were each asked to contribute 5-s music samples conveying each of a wide range of subjective experiences drawn from studies of the feelings associated with music and the voice (1–3, 20). The categories were "amusing," "angry," "annoying," "anxious/tense," "awe-inspiring/amazing," "beautiful," "bittersweet," "calm/relaxing/serene," "compassionate/sympathetic," "dreamy," "eerie/mysterious," "energizing/pump-up," "entrancing," "erotic/desirous," "euphoric/ecstatic," "exciting," "goose bumps," "indignant/defiant," "joyful/cheerful," "nauseating/revolting," "painful," "proud/strong," "romantic/loving," "sad/depressing," "scary/fearful," "tender/longing," "transcendent/mystical," and "triumphant/heroic." These methods enabled us to investigate a broad range of feelings associated with music. (See *SI Appendix*, *Methods S1*, for details and *SI Appendix*, Table S2, for references for each category and Chinese translations.)

After assembling a diverse library of 1,841 music samples, we recruited participants from the United States (n = 1,011, 527 women, mean age = 35.5) and China (n = 895, 556 women, mean age = 30.4) to judge the music samples in 1 of 2 response formats (see *SI Appendix*, Fig. S1 for depiction of surveys). One group of participants was asked to choose the feelings evoked by each music sample from the 28 categories listed above.

A second group of participants rated each music sample on 11 scales measuring broad affective features. The scales were culled from dimensional and componential theoretical accounts of appraisal processes proposed to underlie emotional experience and expression (13, 37), including arousal, attention, certainty, commitment, dominance, enjoyment, familiarity, identity, obstruction, safety, and valence. (Note that these labels are shorthand for the questions to which raters responded. See *SI Appendix*, *Methods S2* and Table S2, for wording of each question and corresponding references.) Participants judged each music sample on 9-point Likert scales (1 = negative levels/none of the quality, 5 = neutral/moderate levels, 9 = extremely high levels).

Based on past estimates of reliability of observer judgment, for each music sample we collected an average of 12 responses from separate participants in each of the 2 response formats in each culture. This amounted to a total of 375,230 judgments of all music samples (59,277 forced-choice categorical judgments and 315,953 9-point scale judgments; see *SI Appendix*, *Methods S2*).

Results

Overview. Guided by our semantic space framework (9), past validated methods, and a central design feature of this investigation—data gathered from the United States and China, nations with cultures differing in their norms and values [e.g., individualism vs. collectivism, low vs. high power distance, low vs. high long-term orientation, and high vs. low indulgence (35)]—our data analysis proceeded as follows. First, to explore issues of conceptualization, we examined which is better preserved across the 2 cultural groups—categories of subjective experience (e.g., "angry," dreamy") or affective features (valence, arousal, etc.)—and which is more potent in explaining variance in judgments from the other cultural group (10). To address the dimensionality of the subjective experiences associated with music, we relied on statistical techniques that uncover how many distinct dimensions are required to account for cross-cultural similarities in judgments of subjective experience. Finally, with visualization techniques, we explored the distribution of feelings associated with music within a high-dimensional space.

Mapping Cultural Similarities in the Subjective Experiences Associated with Music. To assess cross-cultural similarity in subjective experience, past studies have most typically ascertained whether participants' judgments match experimenter expectations. Instead, we operationalized the reliability of subjective experience in terms of interrater agreement, a data-driven approach (see *SI Appendix*, *SI Discussion*, for detailed rationale). We first analyzed the combined data from Chinese and US participants to verify that the 1,841 music samples were reliably associated with different feelings. In this analysis, we found that many different categories of subjective experience were associated with music with a moderate degree of reliability. All but one category of subjective experience ("entrancing") were evoked at significant rates by a number of music samples ranging from 21 ("painful") to 360 ("calm/relaxing/serene") ($q < 0.05$; Monte Carlo simulation using empirical base rates; see *SI Appendix*, *Methods S3*). Indeed, the vast majority (93.3%) of the 1,841 music samples evoked at least one category of subjective experience at a significant rate ($q < 0.05$; see *SI Appendix*, Fig. S1, for distribution of categorical and dimensional ratings). On average, 42.3% of raters from China and the United States chose the most agreed-upon category of subjective experience for each music sample (chance level = 23.1%, Monte Carlo simulation of all category judgments). While interrater agreement rates varied, some music samples were labeled very consistently with categories including "energizing/pump-up" (up to 89% of raters), "triumphant/heroic" (83%), "amusing" (79%), "annoying" (79%), "scary/fearful" (76%), "joyful/cheerful" (74%), "calm/relaxing/serene" (74%), "eerie/mysterious" (69%), and "romantic/loving" (67%). The affective features of the music samples also varied significantly, with the magnitude of Cohen's d for each of the affective scale ratings exceeding 1.2 ($q < 0.05$; Monte Carlo simulation; see *SI Appendix*, *Methods S3*) for between 12 (certainty) and 95 (valence) different music samples.

Is the Experience of Specific Feelings or Broad Affective Features Better Preserved across Cultures? Next, we compared the subjective experiences associated with music across the 2 cultures.

In doing so, we sought to ascertain whether judgments of specific categories of subjective experience or affective features were better preserved. In past studies, cross-cultural similarity has typically been ascertained by comparing the rates with which members of different cultures label a target stimulus—a facial expression, vocalization, or less typically, a musical segment—with the same term. This approach does not capture whether members of each culture also use the same concepts (either categories or scale ratings) to label nontarget stimuli in a similar fashion. For example, if 60% of US participants rate a music sample as "joyful" and 40% as "triumphant," the interrater agreement between US and Chinese raters could never exceed 60% (the interrater agreement among US raters alone) even if the same proportions of Chinese participants respond with each category. Hence, data incorporating the full distribution of responses in each culture are critical to fully understanding cultural similarities in the feelings associated with stimuli of different kinds.

Given this concern, we took a different analytic approach than is typical in the field: for each category of experience and affective scale, we correlated the mean judgments of US raters with those of Chinese raters across all 1,841 music samples. This analysis reveals the extent to which US and Chinese participants use the categories and judgment scales in a similar fashion when labeling the subjective experiences associated with the 1,841 music samples. To control for within-culture variations in the use of the categories and affective scales, we then divided this value by the within-culture explainable variances of these mean judgments (see *SI Appendix*, *Methods S4* and *S5*, for elaborated rationale and further details). Dividing by the explainable variances results in an estimate of the correlation that would emerge if we averaged an infinite number of ratings of each music sample in each culture. We refer to this estimate as a *signal correlation*: it captures the degree of similarity between cultures in the association of each specific feeling and affective feature with music (see *SI Appendix*, Fig. S3, for demonstration that these methods are effective using simulated data). The cross-cultural signal correlations in judgments of each category and affective scale are shown in Fig. 1*A*.

If feelings such as "amusement" are constructed in part from more basic ingredients of valence and arousal, one would expect greater convergence across cultures in how these broad affective features are associated with music than in how the feelings putatively constructed from these affective features are associated with music (13, 14, 38). Inspection of cross-cultural correlations for judgments of each affective feature and category of subjective experience reveals otherwise (Fig. 1*A*). With cross-cultural signal correlations exceeding 0.76, 18 categories of subjective experience were better preserved across China and the United States than valence ($r = 0.75$). Furthermore, cross-cultural signal correlations were significantly greater for joyful/cheerful and triumphant/heroic [$P \leq 0.002$, bootstrap test; $q < 0.05$, Benjamini–Hochberg false discovery rate (FDR) correction (39) (*SI Appendix*, *Methods S5*)] and for amusing, compassionate/sympathetic, energizing/pump-up, exciting, sad/depressing, and scary/fearful ($P \leq 0.024$; $q < 0.1$) than for valence. Correlations for 16 categories of subjective experiences were also better preserved across the 2 cultures than that for arousal ($r = 0.81$), although the differences were not statistically significant for individual categories. The finding that experiences of many categories of subjective experience are better preserved across the 2 cultures than that of valence—sometimes considered a "basic building block of emotional life" (38)—contrasts with the claim that experiences of specific feelings derive from broader affective features (13, 14, 38).

That reported experiences of many categories of subjective experiences were more consistent across the 2 cultures than those of broader affective features raises an intriguing question about the conceptualization of subjective experience: perhaps affective features such as valence and arousal are in fact psychologically constructed from more specific feelings. In other words, perhaps subjective experiences are best represented in terms of more specific feelings such as and "anger" and "triumph," and levels of valence, arousal, and other affective features emerge as higher-order evaluations of these basic

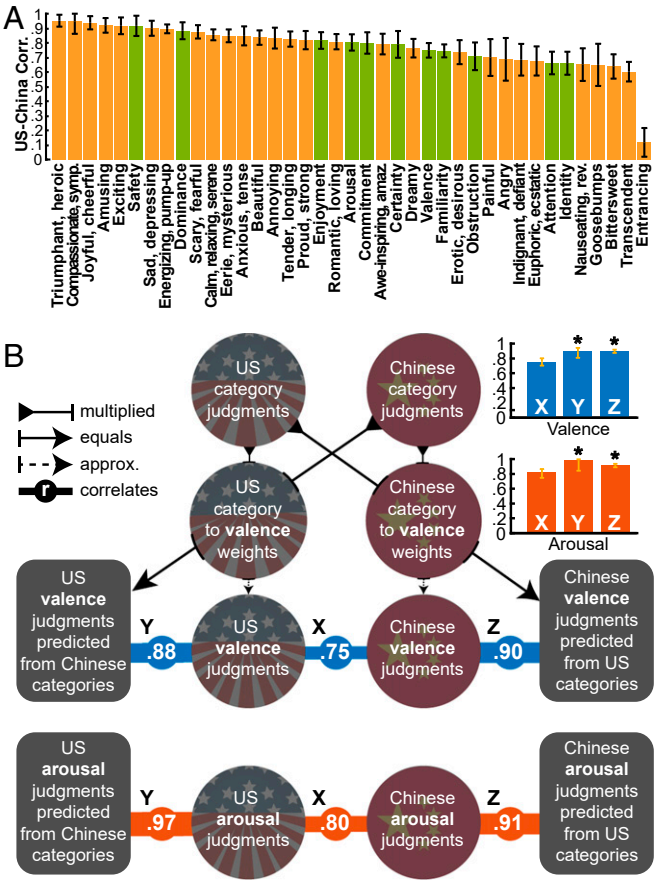

Fig. 1. (*A*) Correlations in feelings associated with music across 2 distinct cultures. The cross-cultural signal correlation for each category (orange bars) and affective scale (green bars) captures the degree to which each judgment is preserved across Chinese and US listeners across all 1,841 music samples. The cross-cultural signal correlation is calculated by correlating the mean responses by Chinese participants with the mean responses by US participants and dividing by the explainable variance in responses from each culture. Error bars represent SE. (See *SI Appendix*, *Methods S4* and *S5*, for details of explainable variance and SE estimation; *SI Appendix*, Fig. S3, for confirmation that these results accurately recover population-level correlations; and *SI Appendix*, Fig. S4, for similar results using Spearman correlations and/or binary affective scale ratings.) (*B*) Categories of subjective experience account for preservation of affective features across cultures. Category judgments are used to predict affective feature judgments within each culture using ordinary least squares regression. Then, regression weights are multiplied by the category judgments from the other culture to predict affective feature judgments in the first culture. Signal correlations between the Chinese and US participants predicted and actual valence/arousal judgments are given in the small red/blue lines and circles, signified by adjacent letters, and plotted on the *Upper Right*. Category judgments from each culture are better than valence/arousal judgments from each culture at predicting valence (*$P = 0.002$ [Y], 0.008 [Z]; bootstrap test, 2-tailed) and arousal (*$P = 0.012$ [Y], 0.024 [Z]) judgments from the other culture. These results are consistent with the hypothesis that specific feelings are elicited by music, and then subsequently used to construct valence/arousal judgments in a more culture-specific process of inference. (See *SI Appendix*, Fig. S5, for consistent results with other affective feature judgments.)

experiences. If this were the case, then the added processing required to infer valence and arousal could heighten differences across cultural groups. As a result, reports of specific feelings, as opposed to affective features, from a given cultural group would be more accurate in predicting reports of the same affective features from the other cultural group. Collecting many independent ratings of both categories of subjective experience and affective features such as valence and arousal across 1,841 music samples, as we have done, allowed for a rigorous test of this possibility.

To test this hypothesis concerning the primacy of specific feelings and broader affective features in subjective experiences, we used linear regression analyses to derive cross-cultural signal correlations in the mapping between category and affective scale judgments of the music samples (*SI Appendix*, *Methods S6*). These analyses ascertain whether ratings of specific feelings are stronger predictors of broad affective feature judgments across the 2 cultures or vice versa. In keeping with the idea that judgments of music in terms of broad affective features derive from the cross-culturally preserved experience of more specific feeling states, we find that category judgments consistently predict affective feature judgments from the other culture as robustly as, or more robustly than, affective feature judgments themselves. We can see this in Fig. 1*B*, which illustrates how category judgments from one culture better predict judgments of valence (blue bars labeled Y and Z) and arousal (red bars labeled Y and Z) from the other culture than the valence and arousal judgments themselves (blue and red bars labeled X). Results are similar for other affective features such as dominance, as shown in the bar graphs at the top of *SI Appendix*, Fig. S5. By contrast, the affective feature judgments from each country generally do a poorer job of predicting the category judgments from the other culture, as indicated in the bar graph at the bottom of *SI Appendix*, Fig. S5. Based on these results, it is more plausible that judgments of general affective features (valence, arousal, etc.) are psychologically constructed from experiences that are captured by specific feelings ("amusement," "fear," etc.) than vice versa in the representation of subjective experience conveyed by music.

Music Is Associated with at Least 13 Distinct Kinds of Subjective Experiences across 2 Cultures. Thus far we have documented that all but one of the categories of subjective experience that we examined were associated at above-chance levels of reliability with at least 21 music samples and that the feelings that participants reported upon hearing music were captured more reliably across the 2 cultures by these categories than by broad affective features (e.g., valence and arousal). These findings set the stage for our next question: How many distinct varieties of subjective experience were consistently evoked across the 2 cultures: That is, what is the dimensionality of the cross-cultural semantic space of subjective experience associated with music? For example, we have not yet ruled out whether some categories were redundant. For example, there may have been interrater reliability in labeling music with categories such as "awe" and "fear," but perhaps these categories were actually associated with precisely the same music samples. If we reduce the US and Chinese judgments to a more limited number of dimensions, what is the minimum number necessary to account for commonalities in reported subjective experience in response to music across cultures? To address this question, we use a recently developed analytic approach.

Namely, to compute the total number of distinct varieties of subjective experience that were significantly preserved across the US and Chinese judgments of the music samples, we used a method that we call principal preserved component analysis (PPCA) (10). PPCA extracts linear combinations of attributes (here, judgments of subjective experience) that maximally covary across 2 datasets measuring the same attributes (US and Chinese judgment data). The resulting components are ordered in terms of their level of positive covariance across the 2 datasets (cultures). (See *SI Appendix*, *Methods S7*, for further details and discussion.)

Given that we found that categories of subjective experience explained the cross-cultural preservation of the affective features, we applied PPCA to the US and Chinese category judgments of the 28 categories of subjective experience associated with the 1,841 music samples to determine the number of independent dimensions, or kinds, of subjective experience that were experienced in both cultures. We applied PPCA in a leave-one-rater-out fashion to determine the statistical significance of each component. (See *SI Appendix*, *Methods S7*, for further details.)

As we show in Fig. 2, PPCA revealed that 13 distinct semantic dimensions of subjective experiences were associated with music and significantly preserved across the US and China participants' judgments [$q < 0.001$ across all held-out raters, $q < 0.05$ across

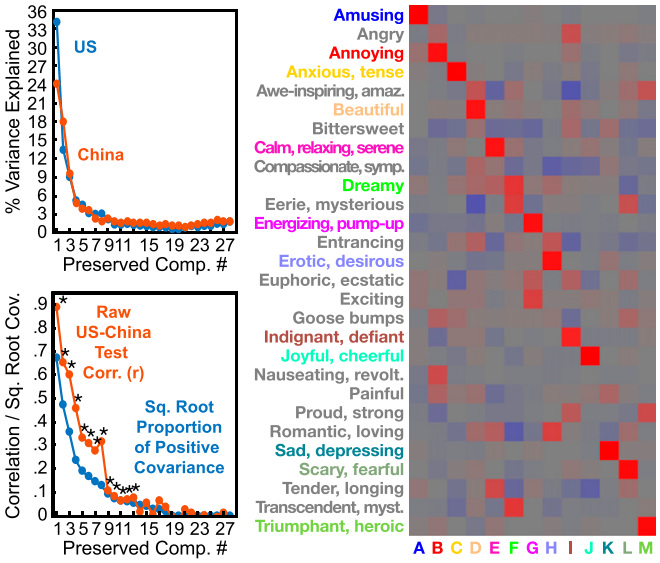

Fig. 2. Thirteen distinct feelings associated with music are preserved across cultures. (*Top Left*) Proportion of variance explained within the US and Chinese ratings by the 29 PPCs of the mean categorical ratings of 28 emotions across cultures. (*Bottom Left*) PPCA finds dimensions of maximal covariance across 2 cultures. Here, the covariance of each PPC across cultures, scaled by total positive covariance, is plotted alongside the cross-cultural test correlation derived from a cross-validation analysis (see *SI Appendix, Methods S7*, for details). We can see that the PPCs are ordered in terms of descending within-sample covariance explained. While it is useful to order the components in terms of the proportion of the shared variance across cultures that they each explain, we are ultimately interested in the out-of-sample test correlation, which measures how similarly the 2 cultures judged music samples along each dimension. The test correlation was significant for 13 PPCs [$q < 0.001$ across all held-out raters, $q < 0.05$ across held-out raters from each country individually, ForwardStop sequential FDR-corrected (40) one-tailed Wilcoxon signed-rank test (41)]. Given the complexity of the cross-validation analysis, these correlations are not adjusted for explainable variance, and it is safe to assume that the population-level correlations are substantially higher. (*Right*) By applying factor rotation (varimax) to the 13 significant PPCs, we found 13 preserved varieties of subjective experience that each load maximally on a distinct category: "amusing," "annoying," "anxious/tense," "beautiful," "calm/relaxing/serene," "dreamy," "energizing/pump-up," "erotic/desirous," "indignant/defiant," "joyful/cheerful," "sad/depressing," "scary/fearful," and "triumphant/heroic." In subsequent figures, we will refer to each component using the category of subjective experience that it loads on maximally. Each component is also assigned a letter, shown on the x-axis.

held-out raters from each country individually, ForwardStop sequential FDR-corrected (40) one-tailed Wilcoxon signed-rank test (41)]. We thus find that music conveys 13 distinct varieties of subjective experience across the 2 cultures. (In Fig. 2, the *Upper Left* chart shows the proportion of variance explained by each dimension [PPC] uncovered by PPCA in data from each culture; the *Bottom Left* chart reveals the proportion of preserved covariance for each dimension, as well as the corresponding correlation and its significance.)

What Are the Feelings Associated with Music? To find the 13 dimensions, or patterns, of subjective experience within the categorical judgments that were preserved across participants from the United States and China, we applied factor rotation (varimax) to the 13 significant components extracted using PPCA. Here, factor rotation extracts a simplified representation of the space by attempting to find dimensions constituted of only a few categories each. After factor rotation, we find that each of the 13 resulting dimensions (PPCs) loaded maximally on a distinct feeling reported in response to the music samples (Fig. 2). Thus, the dimensions can be interpreted as 13 distinct varieties of subjective experience: "amusing," "annoying," "anxious/tense," "beautiful," "calm/relaxing/serene," "dreamy," "energizing/pump-up," "erotic/desirous," "indignant/defiant," "joyful/cheerful," "sad/depressing," "scary/fearful," and "triumphant/heroic." We can infer that these 13 dimensions correspond to qualities of music that evoke similar feelings in China and the United States. Note that some dimensions also involve secondary or tertiary categories—for example, the "indignant/defiant" dimension also loads on "anger"—indicating that these categories were used similarly to label music samples. These findings replicate past studies' findings that several dimensions of subjective experience can be evoked across cultures by music ["calmness," "energy," "fear," "joy," "sadness," and "tension" (29)], but also uncover other dimensions that had not previously been demonstrated to be associated with distinct qualities of music, including "amusement," "annoyance," "defiance," "desire," "dreaminess," and "triumph."

Are the Feelings Associated with Music Distributed along Gradients of Subjective Experience? Having thus far examined the dimensionality and conceptualization of the semantic space of subjective experience associated with music, we sought to establish how the feelings associated with music are distributed. Do they lie within discrete clusters, as predicted by basic emotion theories (42), or along continuous gradients between categories of subjective experience, as recently documented in our investigation of the feelings associated with videos (12)? As can be seen in Fig. 3, the feelings associated with music lie along continuous gradients between categories of subjective experience. These gradients between different categories of subjective experience are evident when we visualize smooth variations in the categorical judgment profiles of the 1,841 music samples using a method called t-distributed stochastic neighbor embedding (t-SNE) (42). t-SNE projects data into a 2-dimensional (2D) space that largely preserves the local distances between data points. In Fig. 3, t-SNE is used to visualize the smooth gradients between feelings associated with music, represented in different colors, and the extent to which they are preserved across cultures (see *SI Appendix, Methods S8*, for details). To allow further exploration of the feelings associated with music and the smooth gradients between them, we also provide an online, interactive version of Fig. 3 in which each music sample can be played while viewing its categorical and affective scale ratings: https://www.ocf.berkeley.edu/~acowen/music.html.

Understanding whether subjective experiences occupy discrete clusters or continuous gradients informs the classification of stimuli and response, preliminary to studying the role of subjective experience in dysfunction, neurophysiology, and other important areas of research. For example, if subjective experiences could be sorted into discrete categories, then diagnoses of mood disorders might rely on how much time is spent experiencing each feeling, one at a time; however, if subjective experiences form continuous gradients, it could be more effective to parameterize feelings along multiple continuous scales. It is crucial, then, to understand whether the smooth gradients that we observe in the map actually correspond to smooth variations in the feelings associated with the music samples in individual listeners.

To answer this question, we analyzed the degree to which reports of valence, arousal, and other affective features steadily

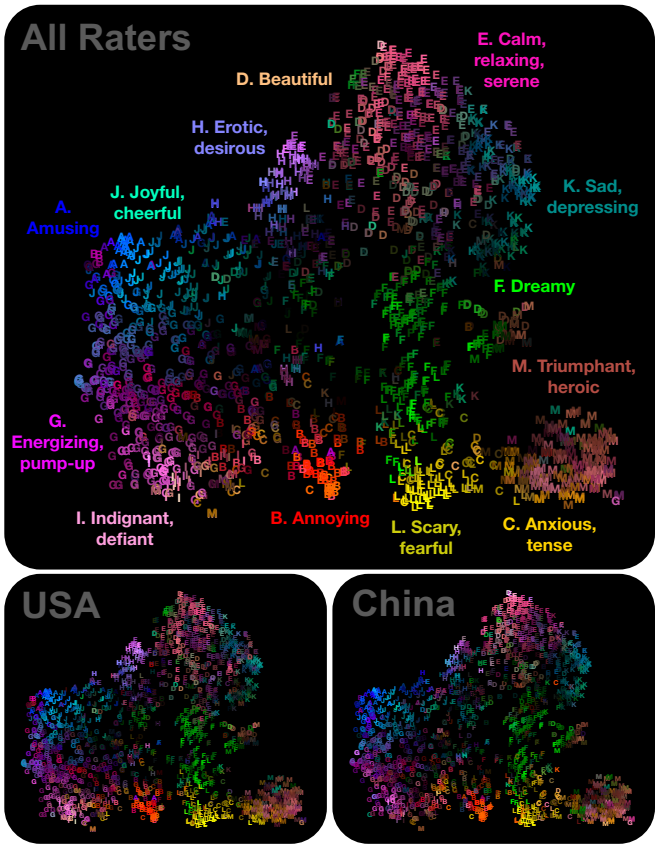

Fig. 3. Visualizing the 13-dimensional structure of subjective experience evoked by music reveals smooth gradients between specific feelings in both cultures. To visualize the feelings associated with music, maps of average category judgments of the 1,841 music samples within a 13-dimensional categorical space of subjective experience were generated. t-SNE, a data visualization method that accurately preserves local distances between data points while separating more distinct data points by longer, more approximate, distances (42), was applied to the concatenated US and Chinese scores of the 1,841 music samples on the 13 categorical judgment dimensions, generating coordinates of each music sample on 2 axes (this does not mean the data are in fact 2D; see *SI Appendix, SI Discussion*). The individual music samples are plotted along these axes as letters that correspond to their highest-loading judgment dimension (with ties broken alphabetically) and are colored using a weighted average of colors corresponding to their scores on the 13 dimensions (see *SI Appendix, Methods S8*, for details). The resulting map reveals gradients in subjective experience, such as a gradient from anxious to triumphant music. For an interactive version of this map, visit https://www.ocf.berkeley.edu/~acowen/musicval.html. Maps colored using projections of the mean US judgments only (*Top Right*) and Chinese judgments only (*Bottom Right*) on the same 13 dimensions demonstrate that reported experiences of the 13 distinct feelings associated with music and smooth gradients between them is largely preserved across the 2 cultures.

transitioned along gradients between categories of subjective experience. (Here, we collapsed data across cultures, given the similarities revealed in Fig. 3.) To do so, we first determined if continuous variations in the proportions of participants who assigned each category to each music sample more accurately predicted each sample's affective features than the one category that was assigned most frequently. This was determined to be true: the 13 dimensions explained 81.0% of the variance in judgments of the affective features; a model comprising 13 binary features designating the dimension to which each music sample was most strongly assigned explained 57.9% of the variance in judgments of the affective features; and a similar model that included magnitude (the score on the most strongly assigned dimension) explained 64.0% of the variance in judgments of the affective features. The 13-dimensional model performed significantly better than either of the mutually exclusive models ($P < 0.002$, bootstrap test; *SI Appendix*, *Methods S9*), consistent with the notion that valence, arousal, and other affective features steadily transition along gradients between categories of subjective experience. Moreover, to determine if these findings could be attributed to correlated ambiguity in judgments of the categories and affective features—for example, the same music being judged as scary and negative by some participants and triumphant and positive by others—we correlated the SD of the category judgments of each music sample with that of the affective feature judgments of each music sample. These had a mildly negative, rather than positive, correlation (Pearson's $r = -0.112$; Spearman's $\rho = -0.109$). Hence the smooth gradients between categories most likely cannot be explained by inter-subject ambiguity; rather, they point to intermediate blends of categories of subjective experience traditionally thought of as discrete.

Fig. 4 presents the number of music samples that loaded significantly on each of the 13 dimensions of subjective experience that we uncovered and on each combination of dimensions. This analysis reveals the many blended feelings that can be evoked by brief music segments. For example, inspection of Fig. 4 reveals that music traverses gradients from "anxious" to "energizing," from "dreamy" to "scary," from "indignant" to "triumphant," and from "calm" to "sad". However, not all categories of subjective experience can be blended together in a single music sample; for example, no music samples were labeled both sad and energizing. Thus, the feelings associated with music are neither entirely discrete nor entirely miscible, but rather are distributed along specific gradients of subjective experience.

Fig. 4. The 13 distinct dimensions of subjective experience associated with music can be blended together in a number of ways. Represented here are the number of music samples that loaded significantly on each dimension, or kind, of subjective experience (diagonal) and on pairwise combinations of dimensions [$q < 0.05$, Monte Carlo simulation using rates of each category judgment, Benjamini–Hochberg FDR corrected (39); see *SI Appendix*, *Methods S10*, for details]. Categories are often blended together, combining, for example, "amusement" with "joy," "indignance" with "triumph," "dreaminess" (or "eeriness") with "sadness," and "anxiety/tension" with "triumph."

Experiment 2. Examining the Primacy of Categories in Feelings Associated with Western and Traditional Chinese Music Targeting Valence and Arousal. In a second experiment, we sought to address 2 alternative explanations for the finding that specific feelings associated with music were better preserved across cultures than valence and arousal. One explanation is that music most reliably conveys whatever it is that it intended to convey, which would entail that the music samples in Experiment 1 reliably conveyed specific feelings rather than broad affective features across the 2 cultures because people provided music samples targeted toward specific feelings. A second alternative explanation is that the primacy of specific feelings is a property not of music in general but of Western music. To test these hypotheses, we examined the primacy of categories in responses to both Western and traditional Chinese music samples gathered on the basis of evoked valence and arousal rather than specific feelings.

To this end, we assembled 2 additional sets of music samples. One set was gathered by 22 US participants (6 women, mean age = 30.7) who were each asked to contribute 5-s music samples conveying 12 different levels of valence and arousal: very pleasant, somewhat pleasant, very unpleasant, somewhat unpleasant, very stimulating, somewhat stimulating, very subdued, somewhat subdued, pleasant and stimulating, pleasant and subdued, unpleasant and stimulating, and unpleasant and subdued. A second set of music samples was gathered by a group of 6 native Chinese participants familiar with Chinese traditional music (5 women, ages 30 to 50), who were asked to provide Chinese traditional music tracks that conveyed as many as possible of the 12 valence and arousal levels. From each Chinese traditional music track, 3 evenly spaced 5-s samples were then extracted. Using these procedures, we compiled a total of 138 modern Western music samples and 189 traditional Chinese music samples.

We recruited new participants from the United States (n = 580, 288 women, mean age = 35.9) and China (n = 363, 249 women, mean age = 23.2) to judge the new music samples in the same response formats used in Experiment 1. For each music sample we collected an average of 23.8 responses in each response format in each culture, amounting to a total of 151,367 individual judgments of the 327 newly gathered music samples (19,147 forced-choice categorical judgments and 132,220 9-point-scale judgments). Interrater agreement rates were similar to Experiment 1, with 37.8% of raters on average choosing the most agreed-upon category of subjective experience for each music sample (versus 42.3% in Experiment 1). Interestingly, Chinese raters converged slightly more in their ratings of traditional Chinese music samples than US raters (*SI Appendix*, Fig. S6), suggesting that ratings are likely also informed by culture-specific understandings of music (5).

Even so, as summarized in Fig. 5, our results regarding the primacy of categories of subjective experience were closely replicated when the models trained in Experiment 1 were validated on judgments of music samples selected based on their valence and arousal levels. Correlations across the 2 cultures in the projections of judgments of the new music samples onto the 13 varieties, or kinds of subjective experience that we uncovered, were uniformly positive [$P \leq 0.003$, q(FDR) ≤ 0.003, bootstrap test; because the dimensions were derived in Experiment 1, the data from Experiment 2 enable us to examine these correlations without "double-dipping"]. Correlations for 5 of the dimensions ("annoying," "beautiful," "energizing/pumped-up," "joyful/cheerful," and "scary/fearful") exceeded the correlation

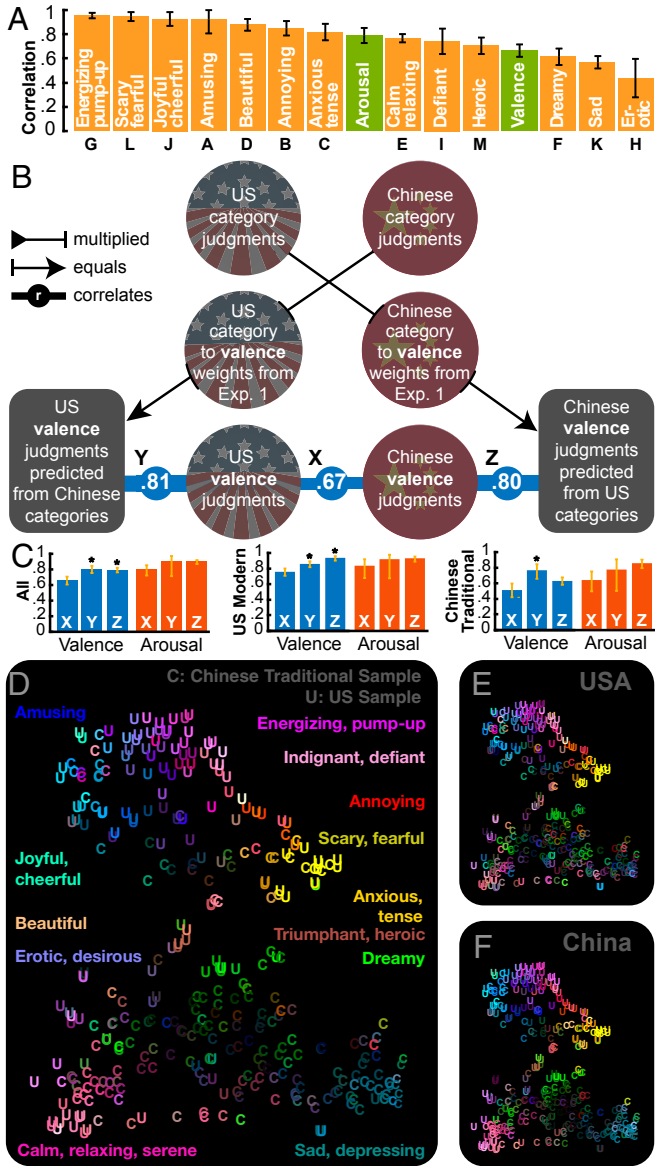

Fig. 5. The primacy of categories in the feelings associated with Western and traditional Chinese music targeting valence and arousal. (*A*) Correlations in the feelings associated with music across cultures. The cross-cultural signal correlation for each dimension derived in Experiment 1 (orange bars) and valence/arousal (green bars) captures the degree to which each attribute is preserved across Chinese and US listeners in the 327 music samples gathered on the basis of valence and arousal. Many feelings were better preserved than valence or arousal. For 5 categories of subjective experience, the difference from valence was significant [$P \leq 0.002$, q(FDR) ≤ 0.05, bootstrap test, 2-tailed]. Error bars represent SE. See *SI Appendix*, Fig. S7, for separate correlations across US and Chinese traditional samples. (*B*) Specific feelings account for the preserved experience of valence and arousal across cultures, even in response to valence and arousal targeted music samples. Category judgments were used to predict affective feature judgments using the model trained in Experiment 1. (*C*) Category judgments from each culture were significantly better than valence/arousal judgments from each culture at predicting valence (*Left*; $*P < 0.002$, 0.01; bootstrap test) judgments from the other culture and nominally better at predicting arousal judgments. This held true for the modern US music samples (*Middle*; $*P = 0.002$, 0.02) and the traditional Chinese music samples (*Right*; $*P < 0.002$) separately. These results are consistent with the hypothesis that specific feelings are elicited by music and then subsequently used to construct valence/arousal judgments in a more culture-specific process of inference. (*D*) Visualizing the 13-dimensional structure of subjective experience incidentally evoked by music samples gathered on the basis of only valence and arousal. We applied t-SNE (42) to the concatenated US and Chinese scores of the 327 new music samples on the 13 categorical judgment dimensions derived in Experiment 1. The resulting map reveals a highly similar structure to that revealed in Experiment 1, with similar gradients in subjective experience, such as a gradient from anxious to triumphant music. Again, reported experiences of these 13 feelings and smooth gradients between them were largely preserved across the 2 cultures (*E* and *F*). To navigate an interactive version of this map and explore the valence and arousal targeted music samples, visit https://www.ocf.berkeley.edu/~acowen/musicval.html.

for valence [$P \leq 0.002$, q(FDR) ≤ 0.05, bootstrap test, 2-tailed comparison].

These results further support the conclusion that judgments of music in terms of affective features derive from the cross-culturally preserved experience of specific feelings. In particular, we found, using the model trained in Experiment 1, that category judgments again predicted valence and arousal judgments from the other culture as robustly as, or more robustly than, valence and arousal judgments, even though the music samples were selected on the basis of their valence and arousal levels (Fig. 5 *B*–*E*). This finding rules out the possibility that music reliably conveys whatever it is selected to target and instead supports the notion of the primacy of specific feelings associated with music. Furthermore, we found that this holds true both for Western music samples (Fig. 5*D*) and for traditional Chinese music samples contributed by native Chinese participants (Fig. 5*E*), indicating that the primacy of the induction of specific feelings is not merely a property of Western music.

Given the surprising range of feelings associated with music samples selected based only on their valence and arousal levels, we measured the number of distinct dimensions of subjective experience associated with these music samples across cultures. We did so by computing partial correlations in judgments projected onto each of the 13 dimensions. Even after controlling for every other dimension, the partial correlation across cultures for each of the 13 dimensions was significantly greater than zero [raw $r \geq 0.13$, $P \leq 0.007$, q(FDR) ≤ 0.05, bootstrap test], indicating that all 13 varieties of subjective experience uncovered in Experiment 1 were also conveyed by distinct music samples across cultures in Experiment 2. (Here we measured dimensionality using a confirmatory approach rather than PPCA, given that Experiment 2 was designed to further investigate the cultural universality of the dimensions of subjective experience that we had already uncovered in a larger sample, not to discover new dimensions.) When considering the Western and Chinese music samples separately, 12 distinct dimensions were preserved across cultures in the 138 Western samples alone (all except "anxious/tense") and 9 were preserved in the 189 Chinese samples alone (all except "amusing," "annoying," "erotic/desirous," and "indignant/defiant") (*SI Appendix*, Fig. S7). It speaks to the richness of the subjective experiences that are routinely associated with music that, even across just a few hundred music samples not targeting specific feelings, so many distinct feelings were reliably distinguished across cultures.

Finally, inspection of the distribution of the feelings associated with music samples gathered on the basis of evoked valence and arousal (Fig. 5 *D*–*F*) reveals that they were structured similarly to those gathered on the basis of specific feelings (Fig. 3). In particular, we still find gradients of music both "calm" and "sad," "sad and "dreamy," and so on through "triumphant," "scary," "anxious," "annoying," "indignant," "energizing," "amusing," "joyful," "beautiful," and "erotic." Again, this was found despite the music samples having been gathered without targeting any of these feelings. This structure was preserved among both US (Fig. 5*E*) and Chinese (Fig. 5*F*) perceivers. These findings validate the high-dimensional, continuous structure of the feelings conveyed by music across cultures.

Discussion

Music is a universal and vital part of social life across cultures (4, 5). Central to its meaning, and likely to its cultural significance, is its well-documented power to evoke rich subjective experiences (1–3, 23–27, 33). What is less well understood is the taxonomy of feelings evoked by music—that is, how these feelings are arranged within a semantic space—and the extent to which such a taxonomy of subjective experience is preserved across cultures.

With quantitative and conceptual approaches (9, 11, 12), we examined the shared semantic space of subjective experience associated with Western and Chinese music by participants from the United States and China. Our focus was to test hypotheses related to 3 properties of this semantic space: its conceptualization, focusing on how specific feelings and broad affective features contribute to the feelings associated with music; its dimensionality, or number of distinct kinds of feelings associated with music; and its distribution of states, focusing on the nature of the boundaries between categories of subjective experience.

Guided by this conceptual framework, 2,777 US and Chinese participants judged over 2,000 samples from modern and classical Western music as well as traditional Chinese music. Judgments were made either in terms of which category of subjective experience, from a list of 28, was evoked, or along scales that captured 11 affective features proposed by appraisal and componential theories to underlie emotional experience. Applying large-scale statistical inference techniques, we compared the preservation of the recognition of 28 categories and 11 scales of affect across 2 different cultures, modeled the latent space that captured the shared variance in judgments between cultures, and interrogated the boundaries between the 13 categories that were found to underlie this latent space.

With respect to the dimensionality of the semantic space of subjective experience, many studies have targeted 6 or fewer categories of emotion and relied on either interrater agreement rates (1, 29, 43) or factor analysis (3, 14, 37) to characterize subjective experience (11). With data-driven statistical techniques applied to a vast array of stimuli, we uncovered 13 distinct dimensions of subjective experience that were associated with music in China and the United States: "amusing," "annoying," "anxious/tense," "beautiful," "calm/relaxing/serene," "dreamy," "energizing/pump-up," "erotic/desirous," "indignant/defiant," "joyful/cheerful," "sad/depressing," "scary/fearful," and "triumphant/heroic" (Fig. 3*B*). It is noteworthy that the subjective experiences reliably associated with music not only include commonly studied emotional experiences ("fear," "sadness," "joy"), but also less commonly studied subjective experiences, including traditionally understudied positive emotions ("amusement," "awe," "calmness," "triumph") (11, 16). It will be important to examine whether these feelings associated with music emerge in cultures beyond the United States and China. Germane to the more general study of aesthetic experience will be the study of whether the feelings elicited by other kinds of aesthetic stimuli—painting, poetry, drama, dance, sculpture—are distributed in semantic spaces similar to the one revealed for music. Understanding the rich variety of feelings associated with music may be particularly useful for studies of the physiological and neural representations of distinct subjective experiences and the emerging literature of neuroaesthetics (8, 15).

Inspection of the music that evokes each dimension (https://www.ocf.berkeley.edu/~acowen/music.html) seems to point to distinctive acoustic and perceptual underpinnings. For example, "indignant/defiant" music tends to involve a "growl-like" timbre, consistent with observed similarities between aggressive music and threatening vocal expressions in various species (44). "Scary/fearful" music tends to involve harmonic dissonance, an example of how some feelings may be associated with specific harmonic (45) and melodic intervals, potentially in a manner that mirrors emotional vocalization (46). Further study is required to determine how the rich array of subjective experiences uncovered here map onto a wide array of acoustic and perceptual features of music, perhaps in complex, nonlinear fashion.

Our second interest was to examine the distribution of feelings within a semantic space of subjective experience—in particular, boundaries between categories of subjective experience. In contrast to discrete theories of emotion (47), we found that the feelings associated with music were not characterized by discrete clusters of states, but by smooth gradients between categories of subjective experience. Music occupies gradients from "amusement" to "joy," "sadness" to "calmness," and "fear" to "tension" to "triumph," among many others (Figs. 4 and 5). These findings support a shift from the predominant scientific focus on how discrete patterns of expression, physiology, and neural activation distinguish discrete emotional states (15, 22) toward the investigation of how the variability and blending together of subjective experience may be represented in continuously varying patterns of expression, physiology, and neural activation.

Our final focus was to examine a critical issue related to the conceptualization of subjective experience: whether the feelings associated with music are explained at a more basic, cross-cultural level by categorical labels or scales of affect. It has been posited that, in subjective experience, valence and arousal can be thought of as core, low-level interpretive processes from which more specific feelings are constructed (13, 14, 38). In contrast to this claim, we found that associations of music with specific feelings such as amusement and triumph were better preserved across cultures than associations with valence and arousal (Fig. 1). Furthermore, judgments of these affective features were better predicted by category judgments from the other culture than by the same affective feature judgments from the other culture (Fig. 1*B*), even in music samples gathered on the basis of valence and arousal associations (Fig. 5). These results suggest that categories of subjective experience such as "amusement" and "desire" may be more directly associated with music, whereas judgments of broader affective features may be inferred in a more culture-specific manner from these categories. These findings contrast with the claim that valence and arousal are basic building blocks of subjective experience (13, 14, 17, 38), instead suggesting that they are higher-order inferences (11).

One might argue, however, that people have less conscious access to valence and arousal levels than to more specific categories of feeling states, causing reports of valence and arousal to be more subject to bias and therefore less predictive. However, given how well valence and arousal judgments within each culture were predicted by category judgments (Figs. 1 and 5), it seems that the valence and arousal judgments were reliable and quite precise—even if the meanings that they were precise in capturing were culture-specific. Perhaps, then, participants' valence and arousal ratings capture something distinct from "real" valence and arousal levels to which people do not have conscious access. But falling back on this hypothesis seems to undermine the original motivation behind the constructs of valence and arousal, which emerged in self-report data (14).

In considering the evidence for universals in music-related feeling, it is likely that the cultural similarities that we observed were influenced by participants' past exposure to music. For example, the association of music with an "erotic/desirous" dimension may derive from previous exposure in both cultures to music played during romantic scenes in films. Such semantic associations may have influenced our findings. However, this is less likely to be the case for Westerners' responses to Chinese traditional music, a genre with which most Westerners have limited familiarity, yet for which valence and arousal judgments were still best predicted by category judgments from the other culture. These findings render it unlikely that the primacy of specific feelings over broad affective features could be fully

explained by learned semantic associations. Nevertheless, given the influence of globalization, we avoid inferences about biologically ingrained universals.

In introducing quantitative approaches to study US and Chinese experiences of music, we hope to set the stage for a broader project documenting shared and culture-specific understandings of music around the world. Different experiences may emerge when studying musical traditions from other regions, such as Africa and South America. It will be particularly informative to study the semantic space of subjective experience associated with music in small-scale cultures with limited Western contact.

We also recognize that, although participants in the United States and China differ culturally on average, neither group is homogeneous. In deriving a semantic space that captures the similarities across countries in how participants represent their subjective experiences, we set aside variation in responses within each country, an important topic for further study. We also did not delve deeply into what past work suggests may be rich culture-specific understandings of music. For example, it has been found that when listening to Hindustani ragas known to have specific emotional connotations, Westerners can reliably identify the intended emotions, but do so with less granularity than Indian listeners (5). While some evidence was observed for culture-specific understandings of traditional Chinese music (*SI Appendix*, Fig. S6), the precise nature of these understandings—or, for that matter, understandings of music from more specific subcultures—is an interesting topic for further study.

Importantly, the present study did not focus on the distinctions between experienced and perceived feelings in music. Recent studies have found these constructs to be highly correlated, but not identical (48–50). It would be interesting for future studies to interrogate the possible differences in structure between perceived and evoked feelings in music, perhaps incorporating physiological measures. Moreover, it could be argued that the brief, 5-s duration of our stimuli makes them unlikely to evoke certain feelings—to cause people to become fixated and entranced, evoke memories that give rise to nostalgia, engage physiological changes that evoke empowerment (51), or enable a deeper appreciation of their historical and artistic context, for example. Future studies could examine the structure of such responses using longer music samples.

Although the present study cannot determine the full range of subjective experiences that can be evoked by musical compositions or whether these responses are innate, it places a lower bound on the richness of the feelings reliably associated with music in 2 rather different cultures. Even brief, 5-s music samples are reliably associated with at least 13 distinct dimensions of subjective experience by participants within the United States and China. To understand how and why music evokes these feelings, future studies will need to move beyond the pervasive focus on simple models of subjective experience composed of 4 to 6 emotion categories or levels of valence and arousal (1, 29), given that these models confound many distinct variables (11). Studies of the human response to music—its universality (32–34), acoustic underpinnings (1), and neural representation (8, 15)—should therefore incorporate a rich taxonomic structure of subjective experience. Similarly, technological endeavors to build machines that curate, compose, or alter music should incorporate the wide range of feelings uncovered here if they aim to reproduce the full complexity of human musical composition.

Finally, it is worth noting how the present findings dovetail with the implications of recent research on subjective responses to strongly evocative videos. Cowen and Keltner (12) found that, when people report their feelings in response to a wide range of evocative videos, they reliably distinguish among at least 27 varieties of subjective experience. In that study, specific feelings such as "awe" also emerged as more primary than broad affective features such as valence and arousal in determining the structure of subjective experience and were also found to be organized along continuous gradients. However, those results emerged within a single culture and may have been influenced by the semantic content of the videos. The present study examined subjective responses to a largely nonrepresentational stimulus modality across 2 different cultures. Together, these results converge on a taxonomy of subjective experience consisting of a rich array of distinct categories bridged by smooth gradients.

Questions regarding the structure of subjective experience bear upon central theoretical claims in affective science. They guide studies in fields ranging from neuroscience to machine learning. Our method of interrogating how subjective experiences associated with music are situated within a semantic space reveals a more complex taxonomy of feelings than is typical in existing accounts of how subjective experience is organized. By revealing that the feelings evoked by music span at least 13 distinct dimensions and are distributed along continuous gradients, we provide a road map for the study of how subjective experiences are associated with physiological responses, are represented in the brain, and can be predicted using machine learning.

Methods

Judgments were obtained using Amazon Mechanical Turk and Qualtrics. A total of 2,777 participants (1,591 from the United States, 1,258 from China) took part in the study. The experimental procedures were approved by the Institutional Review Board at the University of California, Berkeley. All participants gave their informed consent. Data and analysis code can be requested at https://forms.gle/73Diih84zksjkASe9. See *SI Appendix* for details.

1. P. N. Juslin, P. Laukka, Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *J. New Music Res.* **33**, 217–238 (2004).
2. P. N. Juslin, S. Liljeström, D. Västfjäll, G. Barradas, A. Silva, An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion* **8**, 668–683 (2008).
3. M. Zentner, D. Grandjean, K. R. Scherer, Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion* **8**, 494–521 (2008).
4. G. Rouget, M. Buckner, Musical efficacy: Musicking to survive—The case of the pygmies. *Yearbook Tradit. Music* **43**, 89–121 (2011).
5. W. F. Thompson, L.-L. Balkwill, "Cross-cultural similarities and differences" in *Handbook of Music and Emotion: Theory, Research, Applications*, P. N. Juslin, J. A. Sloboda, Eds. (Oxford University Press, 2010), pp. 755–788.
6. P. N. Juslin, P. Laukka, Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychol. Bull.* **129**, 770–814 (2003).
7. D. A. Schwartz, C. Q. Howe, D. Purves, The statistical structure of human speech sounds predicts musical universals. *J. Neurosci.* **23**, 7160–7168 (2003).
8. S. Koelsch, Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* **15**, 170–180 (2014).
9. A. S. Cowen, D. Keltner, Clarifying the conceptualization, dimensionality, and structure of emotion: Response to Barrett and colleagues. *Trends Cogn. Sci.* **22**, 274–276 (2018).
10. A. S. Cowen, P. Laukka, H. A. Elfenbein, R. Liu, D. Keltner, The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nat. Hum. Behav.* **3**, 369–382 (2019).
11. A. Cowen, D. Sauter, J. L. Tracy, D. Keltner, Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression. *Psychol. Sci. Public Interest* **20**, 69–90 (2019).
12. A. S. Cowen, D. Keltner, Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7900–E7909 (2017).
13. L. F. Barrett, Z. Khan, J. Dy, D. Brooks, Nature of emotion categories: Comment on Cowen and Keltner. *Trends Cogn. Sci.* **22**, 97–99 (2018).
14. J. A. Russell, Core affect and the psychological construction of emotion. *Psychol. Rev.* **110**, 145–172 (2003).
15. P. A. Kragel, K. S. LaBar, Decoding the nature of emotion in the brain. *Trends Cogn. Sci.* **20**, 444–455 (2016).
16. Y. Wu, P. Muentener, L. E. Schulz, One- to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11896–11901 (2017).
17. J. Posner, J. A. Russell, B. S. Peterson, The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* **17**, 715–734 (2005).
18. K. Oatley, J. M. Jenkins, Human emotions: Function and dysfunction. *Annu. Rev. Psychol.* **43**, 55–85 (1992).
19. D. Watson, K. Stanton, Emotion blends and mixed emotions in the hierarchical structure of affect. *Emot. Rev.* **9**, 99–104 (2017).

20. A. S. Cowen, H. A. Elfenbein, P. Laukka, D. Keltner, Mapping 24 emotions conveyed by brief human vocalization. *Am. Psychol.* **74**, 698–712 (2019).
21. A. S. Cowen, D. Keltner, What the face displays: Mapping 28 emotions conveyed by naturalistic expression. *Am Psychol.*, 10.1037/amp0000488 (2019).
22. F. E. Harrell, *Regression Modeling Strategies* (Springer, 2001).
23. A. Goldstein, Thrills in response to music and other stimuli. *Physiol. Psychol.* **8**, 126–129 (1980).
24. J. A. Sloboda, Musical structure and emotional response: Some empirical findings. *Psychol. Music* **19**, 110–120 (1991).
25. W. Tsitsos, Rules of rebellion: Slamdancing, moshing, and the American alternative scene. *Pop. Music* **18**, 397–414 (1999).
26. D. A. Hodges, *Bodily Responses to Music* (The Oxford Handbook of Music Psychology, 2008).
27. P. J. Silvia, K. Fayn, E. C. Nusbaum, R. E. Beaty, Openness to experience and awe in response to nature and music: Personality and profound aesthetic experiences. *Psychol. Aesth. Creat. Arts* **9**, 376–384 (2015).
28. A. Gabrielsson, Emotion perceived and emotion felt: Same or different? *Music. Sci.* **5**, 123–147 (2001).
29. J. K. Vuoskoski, T. Eerola, Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Music. Sci.* **15**, 159–173 (2011).
30. P. A. Kragel, K. S. Labar, Multivariate pattern classification reveals autonomic and experiential representations of discrete emotions. *Emotion* **13**, 681–690 (2013).
31. A. Norenzayan, S. J. Heine, Psychological universals: What are they and how can we know? *Psychol. Bull.* **131**, 763–784 (2005).
32. T. Fritz *et al.*, Universal recognition of three basic emotions in music. *Curr. Biol.* **19**, 573–576 (2009).
33. H. Egermann, N. Fernando, L. Chuen, S. McAdams, Music induces universal emotion-related psychophysiological responses: Comparing Canadian listeners to Congolese pygmies. *Front. Psychol.* **5**, 1341 (2015).
34. J. H. McDermott, A. F. Schultz, E. A. Undurraga, R. A. Godoy, Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature* **535**, 547–550 (2016).
35. G. Hofstede, Dimensionalizing cultures: The hofstede model in context. *Online Read. Psychol. Cult.* **2**, 1–26 (2011).
36. T. Hamamura, S. J. Heine, D. L. Paulhus, Cultural differences in response styles: The role of dialectical thinking. *Pers. Individ. Dif.* **44**, 932–942 (2008).
37. C. A. Smith, P. C. Ellsworth, Patterns of cognitive appraisal in emotion. *J. Pers. Soc. Psychol.* **48**, 813–838 (1985).
38. L. F. Barrett, Valence is a basic building block of emotional life. *J. Res. Pers.* **40**, 35–55 (2006).
39. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
40. M. G. G'Sell, S. Wager, A. Chouldechova, R. Tibshirani, Sequential selection procedures and false discovery rate control. *J. R. Stat. Soc. Series B Stat. Methodol.* **78**, 423–444 (2016).
41. F. Wilcoxon, Individual comparisons by ranking methods. *Biom. Bull.* **1**, 80 (1945).
42. L. V. D. Maaten, G. Hinton, Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
43. H. C. Lench, S. A. Flores, S. W. Bench, Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychol. Bull.* **137**, 834–855 (2011).
44. C. G. Tsai *et al.*, Aggressiveness of the growl-like timbre: Acoustic characteristics, musical implications, and biomechanical mechanisms. *Music Percept.* **27**, 209–221 (2010).
45. M. Costa, P. E. Ricci Bitti, L. Bonfiglioli, Psychological connotations of harmonic musical intervals. *Psychol. Music* **28**, 4–22 (2000).
46. D. L. Bowling, J. Sundararajan, S. Han, D. Purves, Expression of emotion in Eastern and Western music mirrors vocalization. *PLoS One* **7**, e31942 (2012).
47. P. Ekman, "All emotions are basic" in *The Nature of Emotion*, D. Ekman, Ed. (Oxford University Press, New York, 1992), pp. 15–19.
48. L. O. Lundqvist, F. Carlsson, P. Hilmersson, P. N. Juslin, Emotional responses to music: Experience, expression, and physiology. *Psychol. Music* **37**, 61–90 (2009).
49. P. G. Hunter, E. G. Schellenberg, U. Schimmack, Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions. *Psychol. Aesthet. Creat. Arts* **4**, 47–56 (2010).
50. M. Komosinski, A. Mensfelt, Emotions perceived and emotions experienced in response to computer-generated music. *Music Percept.* **33**, 432–445 (2016).
51. W. F. Thompson, K. N. Olsen, On the enjoyment of violence and aggression in music. Comment on "An integrative review of the enjoyment of sadness associated with music" by Tuomas Eerola et al. *Phys. Life Rev.* **25**, 128–130 (2018).