Quantifying music and music perception is a difficult problem, in part because there are many dimensions to music as it is an artistic and communicative medium that unfolds over time. Historically, experiments addressed this question from a rigid univariate perspective and often used simplified stimuli and strict controls. The resultant work is useful in identifying small-scale, specific elements of auditory perception, but does not generalize to larger scale works. Additionally, as research pivots to larger questions of universality, these studies have little to offer. To address this gap in the literature, we designed a bilingual experiment in which participants in either France or the United States rated 15 novel 30-second stimuli either on quantifiable musical qualities (e.g., tempo, meter, harmony) or using adjectives (e.g., bright, dark, soft, aggressive) .

A total of 30 stimuli were composed for this study (each participant rated half of the stimuli). The musical qualities survey asked participants to select from predetermined options and the other survey asked participants to select any adjectives that they felt described the excerpt. We then used exploratory multivariate analyses including hierarchical cluster analysis, distance analysis and correspondence analysis, in conjunction with bootstrapping and permutation tests to evaluate a) whether there was an identifiable cognitive listening space, b) whether this space was consistent across listeners of different nationalities or cultures, and c) whether this space was consistent between the two experiments (i.e ., whether music is described the same or not based on adjectives or expert evaluation of concrete musical qualities).

A correspondence analysis (CA) of the musical qualities survey revealed that the three most important dimensions explained only about 40% of the variance—a pattern suggesting that many different musical qualities contribute to judgements of difference between musical excerpts. Additionally, this analysis revealed that excerpts that reflect more modern musical movements (e.g., minimalism) greatly affect the overall factor space.

Another CA of the contingency table collecting the frequency of association between adjectives and excerpts provided a space that reflected the valence/arousal emotional model. The space defined by the descriptors is much more robust to musically diverse stimuli. A distance analysis on the participants’ ratings showed that the average American and French participants described the stimuli differently. The American participants also exhibited more variance than French participants in how they described the stimuli. The exact reason for this is not clear and merits further study.

An additional analysis correlating the results of both surveys suggests that multivariate correlations between the adjectives and qualities models are able to identify stimuli that anchor the ‘corners’ of the valence arousal model, namely: Positive valence, High arousal; Negative valence, High arousal; Negative valence, Low arousal; and Positive valence, low arousal.