KEER2026, EUGENE, OR, USA| AUGUST 11-12, 2026
INTERNATIONAL CONFERENCE ON KANSEI ENGINEERING AND EMOTION RESEARCH 2026

A MULTI-LAYER PSYCHOPHYSICAL MODEL OF PREFERENCE FOR FREQUENCY RESPONSE (EQ) IN MUSIC LISTENING

Taiki SHIMOZONO[1a], Yasuo SHIOZAWA[a] , Yuma YANO[b], Kiichi NAKA[c], Yoichi YAMAZAKI[d] , Noriko NAGATA[b]

[a] *Yamaha Corporation, Japan, taiki.shimozono@music.yamaha.com*
[b] *Kwansei Gakuin University, Japan, y-yuma1518@kwansei.ac.jp, nagata@kwansei.ac.jp,*
[c] *University of Nagasaki, Japan, naka.kiichi.research@gmail.com*
[d] *Aichi Prefectural University, Japan, yamazaki@ist.aichi-pu.ac.jp*

ABSTRACT

Frequency response (EQ) characteristics strongly affect listening satisfaction in music playback, but why a given EQ is preferred may differ depending on a listener's evaluation perspective. This study models the EQ preference process as a multi-layer psychophysical model—acoustic features → impressions → value/emotion—using both qualitative and quantitative approaches. First, we extracted individual evaluation structures for "preference" using the Evaluation Grid Method (EGM) and screened the resulting terms into a set of 43 rating words (16 value/emotion terms and 27 impression terms). Factor analyses organized impressions into four factors (Power, Relaxation, Clarity, and Coherence) and value/emotion into three factors (Comfortable, Overwhelmed, and Brightened). We also derived principal components from acoustic features and estimated structural paths linking them to impressions and value/emotion, yielding a model with good fit and high explanatory power. Listeners were further classified into three user types—Emotion, Sound, and Environment—and the results suggest that pathways of value formation may differ across types.

Keywords: *EQ preference, multi-layer model, impression, affect/value, user segmentation*

[1] Corresponding author. West Tower, Yokohama Symphostage 5-1-2 Minatomirai, Nishi-ku, Yokohama, Kanagawa, Japan. taiki.shimozono@music.yamaha.com

1 INTRODUCTION

Listening satisfaction in music is influenced not only by musical content but also by the frequency response characteristics of playback systems, such as the frequency response of loudspeakers/headphones and equalization (EQ) adjustments. EQ is a common means of shaping perceived balance to achieve experiences described as “easy to listen to,” “powerful,” or “calming.” However, the relationship between EQ-induced acoustic changes and subjective preference or value judgments is not straightforward. Even under the same EQ manipulation, listeners may prioritize different evaluation perspectives, leading to divergent preference judgments.

Previous studies have often addressed EQ and preference statistically, focusing on which responses tend to be preferred. In contrast, less attention has been paid to why they are preferred—namely, the evaluative process through which acoustic changes are translated into impressions and ultimately into value or emotion. To address this gap, the present paper (1) extracts evaluation structures of preference using the Evaluation Grid Method (EGM), (2) constructs a multi-layer psychophysical model based on rating experiments and acoustic-feature analysis, and (3) compares the model across user types to clarify individual differences and their underlying evaluative pathways.

2 RELATED WORK

2.1 Frequency Response and Preference

Equalization (EQ) is one of the primary factors shaping the perceived sound quality of headphones and loudspeakers. Prior work has reported both a commonly preferred tendency (often described as a target response) and substantial individual differences around that tendency. For example, Olive et al. (2014) conducted double-blind listening tests across multiple countries and demonstrated how frequency-response tuning influences sound-quality ratings and preference, as well as how preference varies with listeners’ age, experience, and cultural background. In addition, Olive (2019) applied clustering to preference data and showed that listeners can be segmented based on their preferred frequency-response profiles, with different segments exhibiting distinct response-shape tendencies (e.g., emphasis in low- and/or high-frequency regions).

2.2 Measuring Impressions and Emotion in Audio/Music (SD and Affect Grid)

Subjective evaluation of sound and music has been studied using two complementary frameworks: representing verbal impressions with a small number of underlying dimensions, and describing affective states in a low-dimensional space. A representative approach for the former is the Semantic Differential (SD) method, whose major dimensions—evaluation, potency, and activity—are widely recognized as a basic interpretive reference for impression ratings (Katahira et al., 2016). For the latter, affect is often characterized along two dimensions, valence and arousal, positioned as core components of emotional experience (core affect) (Ono et al., 2019). In this study, we use both perspectives—organizing impressions with reference to evaluation/potency/activity and representing affect in terms of valence and arousal—as a conceptual basis for describing the evaluative process from EQ-induced changes to preference judgments.

2.3 Uses of Music and Individual Differences

Music listening is not only a matter of "like/dislike"; its meaning also depends on *how* and *why* people listen, such as the intended purpose and situation of listening. Chamorro-Premuzic and Furnham (2007) investigated everyday functions of music listening and showed that uses of music can be organized into dimensions such as emotional use, cognitive use, and background use. This perspective suggests that preference for EQ may likewise reflect not only acoustic differences per se but also evaluation viewpoints tied to listeners' situations and listening purposes. Accordingly, we position the "uses of music" perspective as an interpretive framework for understanding evaluation structures and for defining user types in the present study.

3 RESEARCH FRAMEWORK: DEFINITION OF THE MULTI-LAYER PSYCHOPHYSICAL MODEL

Building on prior work showing that preferred frequency-response profiles vary across listeners, and drawing on frameworks for impression and affective evaluation, this study focuses on the internal evaluative pathways through which acoustic changes are interpreted as impressions and value/emotional responses. The proposed three-layer model consists of (i) an **acoustic feature layer** representing signal properties altered by EQ, (ii) an **impression layer** representing listeners' subjective impressions, and (iii) a **value layer** representing value judgments and affective responses elicited during listening. The impression layer refers to the major dimensions discussed in the Semantic Differential framework (evaluation, potency, and activity), and the value layer refers to the two-dimensional core-affect space (valence and arousal). We primarily assume a mediated pathway in which acoustic features influence value judgments through impressions, while not excluding the possibility of a direct influence from the acoustic feature layer to the value layer. Moreover, the mapping between layers—i.e., how physical differences are interpreted as particular impressions and values—may depend on personal constructs formed through experience (Kelly, 1955). Accordingly, we assume that the pathways between layers exhibit individual differences, as illustrated in the Figure1. A similar hierarchical approach to structuring impression layers has been demonstrated in perceptual evaluations across domains, including music liveness (Shimozono et al., 2025) and 3D shapes (Miyai et al., 2019).

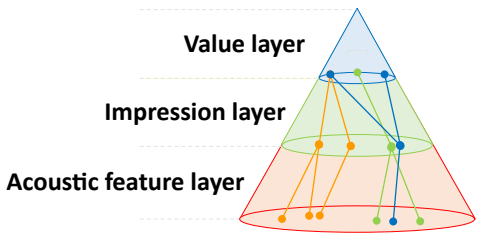


Figure 1. The multi-layer psychophysical model assumed in this study

4 EVALUATION STRUCTURE ANALYSIS

4.1 Aim

This chapter extracts individual evaluation structures for EQ preference through semi-structured interviews using the Evaluation Grid Method (EGM).

4.2 Stimuli

4.2.1 Source music

We selected the musical pieces from the Music Genre Database of the RWC Music Database (RWC-MDB-G), a copyright-cleared database designed for shared use in research (Goto et al., 2003) (Table 1).

Table 1. Audio stimuli used

Gener	ID	Title	Composer	SPL(dBA)
Classic	RWC-MDB-G-051	Prelude to Act I from "Die Meistersinger von Nürnberg"	Wagner, R.	70.4
Dance	RWC-MDB-G-021	Stf	Yamamoto, Ichi	67.6
Jazz	RWC-MDB-G-033	The Ramp	Haji, Hiroshi	66.6
Pop	RWC-MDB-G-003	In Your Arms	Kawagoe, Yu	75.3
Rock	RWC-MDB-G-008	Living in a World of Destruction	Kamaki, Shigeki	74.1

4.2.2 EQ processing

We prepared six EQ conditions: Flat (no EQ), low-frequency boost, low-frequency cut, mid-frequency boost, mid-frequency cut, and high-frequency cut. The low-frequency conditions were implemented as a low-shelf filter at 250 Hz (Q/slope = 1.0) with a gain of ±9 dB. The mid-frequency conditions were implemented as a peaking filter at 1 kHz (Q = 0.5) with a gain of ±9 dB. The high-frequency condition was implemented as a high-shelf filter at 4 kHz (Q/slope = 1.0) with a gain of −9 dB (Figure 2).

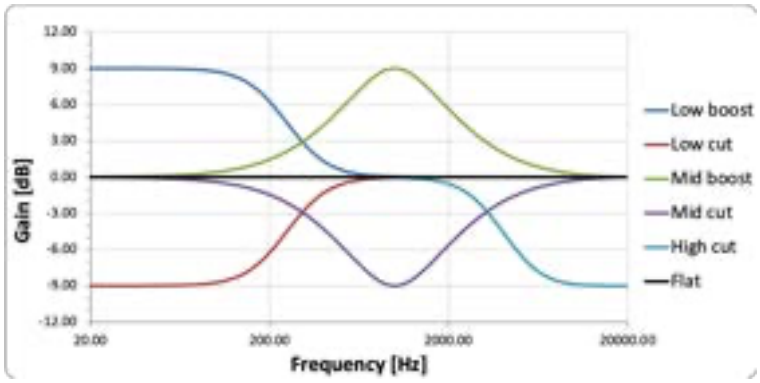


Figure 2. Frequency response curves of the six EQ conditions

4.2.3 Loudness normalization

Because level differences can strongly influence preference judgments in subjective evaluation, stimulus levels must be controlled. Stimuli were normalized using loudness estimation based on ISO 532-1 (Zwicker method) to minimize inter-stimulus level differences (ISO, 2017).

4.3 Experimental Design

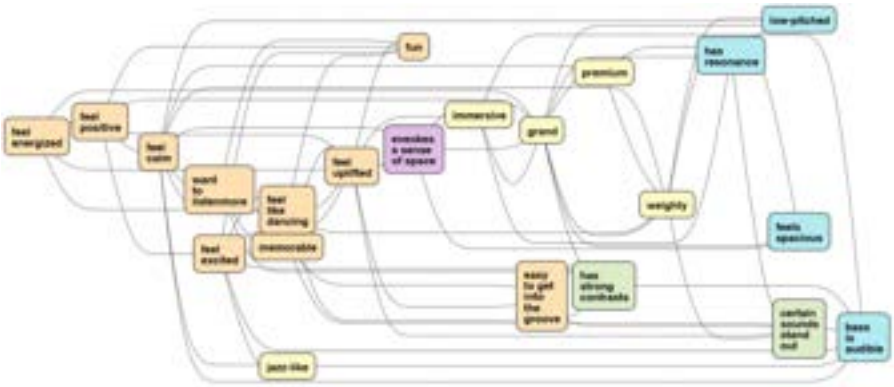
Participants were 20 university and graduate students (11 male, 9 female; age range = 19–23 years, M = 20.9, SD = 1.0) with normal hearing. All participants were right-handed. Their musical backgrounds varied: 7 were classified as musicians and 13 as non-musicians, and their daily music listening time ranged from 0.5 to 5.0 h (M = 2.77, SD = 1.18). We prepared 30 stimuli in total (5 genres × 6 EQ conditions). The experiment was conducted in a quiet environment, with

participants seated and listening via headphones. The presentation level was set to approximately 70 dBA. The study was exploratory, aiming to elicit evaluation viewpoints under controlled listening conditions and examine the feasibility of the proposed multi-layer framework. The sample size should therefore be considered a limitation when generalizing the user types and path models.

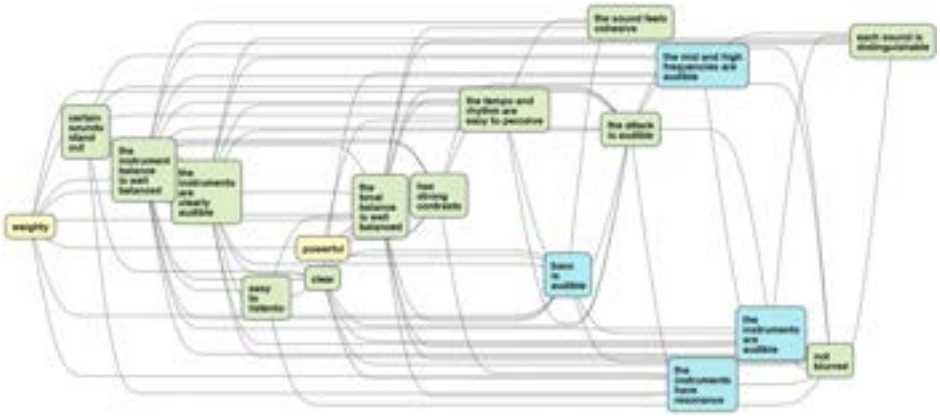
Within EGM, comparisons were made within each genre, i.e., participants compared only the six EQ conditions of the same genre (e.g., the six “Classic” stimuli were compared with each other). For each genre, participants listened to the six conditions and first selected the condition they found most preferable. They then assigned that condition a preference score of 100, and rated the remaining five conditions on a 0–100 scale relative to it. After the ratings, we conducted EGM interviews focusing on preference within the same genre. The interviews used laddering (ladder-up/ladder-down) to explore higher-level concepts and lower-level concrete reasons starting from participants’ own expressions (Sanui, 1996; Reynolds & Gutman, 1988; Sugimoto et al., 2023).

4.4 Results

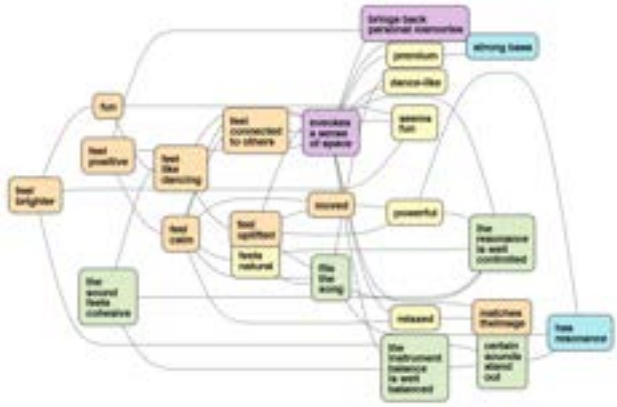
From the EGM interview records, we extracted utterances related to preference judgments as evaluation terms, yielding a total of 3,756 terms. Rather than directly referring to frequency bands or EQ operations, many terms described perceptual attributes and musical elements (e.g., how instruments are heard, contours, texture, and balance). Based on the evaluation-term data and individual evaluation structures, we found that listeners relied on different dominant viewpoints even under the same EQ conditions. To quantify these differences, we modeled each participant’s evaluation structure as a network and used Katz centrality to extract central terms (Katz, 1953). We then assigned each participant to a dominant viewpoint type, defined as the viewpoint category (emotion, sound, or environment) that accounted for the largest share of the extracted central terms. Consequently, participants were classified into 6 Emotion-type, 11 Sound-type, and 3 Environment-type listeners; the corresponding evaluation-structure diagrams are shown in Figure 3. These types may be interpreted in relation to prior frameworks on uses/functions of music listening (Chamorro-Premuzic & Furnham, 2007).



(a) Emotion type



(b) Sound type



(c) Environment type

Figure 3. Representative evaluation-structure networks for the three user types

5 QUALITATIVE MODEL DEVELOPMENT

Because the 3,756 EGM-derived evaluation terms were too numerous for subsequent rating experiments and modeling, this chapter constructs a representative rating-term set through (1) extracting central terms based on Katz centrality (Katz, 1953), (2) screening terms for usability in expressing EQ-induced differences, and (3) grouping and merging semantically similar terms. The final set comprised 43 terms (16 value/emotion terms and 27 impression terms).

Table 2. Selected evaluation terms for subsequent rating experiments

Value	Arousal	Comfort	Interest	Positivity	Moved	Naturalness/Ease	
preferred	uplifted	comfortable	memorable	brightened	cool	natural	
good	excited	pleasant	has substance	energized	overwhelmed	easy to listen to	
satisfied		reassured					
Clarity	**Presence**	**Audibility**	**Spatiality**	**Stability**	**Power**	**Coherence**	**Warmth**
clear	feels exciting	instruments clearly audible	has depth	stable	weighty	well-balanced	warm
each sound is distinguishable	immersive	bass is audible	has resonance		powerful	cohesive	
Imagery	**Ease of listening**	**Contrast**	**Sustain**	**Familiarity**	**Tempo /rhythm**	**Luxury**	
fits the song	easy to hear	has strong contrasts	good sustain	approachable	tempo and rhythm are perceivable	luxurious	
evokes space	good sound quality	penetrating		familiar	relaxed	premium	

6 QUANTITATIVE MODEL DEVELOPMENT

6.1 Aim

This chapter estimates a multi-layer structural model linking acoustic features to impressions and value/emotion responses.

6.2 Rating Experiment for Value/Emotion and Impressions

To quantify value/emotion and impressions, we conducted a headphone-based rating experiment. The stimuli were the same 30 audio excerpts as in Section 4.3 (5 genres × 6 EQ conditions). For each stimulus, participants rated both value/emotion and impressions using 43 evaluation terms identified in the previous chapter (16 value/emotion terms and 27 impression terms). Ratings were made on a 7-point scale for each stimulus–term combination.

6.3 Factor Analysis of Impression Structure and Value/Emotion Structure

6.3.1 Impression Structure (27 impression terms)

We examined several criteria to determine the number of factors and selected a solution based on interpretability. Factors were extracted using maximum likelihood estimation and rotated with Promax rotation. The impression ratings were summarized into four factors: Power (e.g., “powerful,” “weighty,” “immersive”), Relaxation (e.g., “relaxed,” “warm,” “friendly”), Clarity (e.g., “clear,” “each sound is discernible”), and Coherence (e.g., “fits the music,” “cohesive,” “well-balanced”).

6.3.2 Value/Emotion Structure (16 value/emotion terms)

The value/emotion ratings were analyzed in the same manner (maximum likelihood extraction with Promax rotation). The results were summarized into three factors: Comfortable (positive valence / low arousal), Overwhelmed (high arousal), and Brightened (positive valence / high arousal). The results are summarized in Table 3.

Table 3. Factor loadings for impression (27 terms; four factors) and value/emotion ratings (16 terms; three factors) (maximum likelihood, Promax rotation).

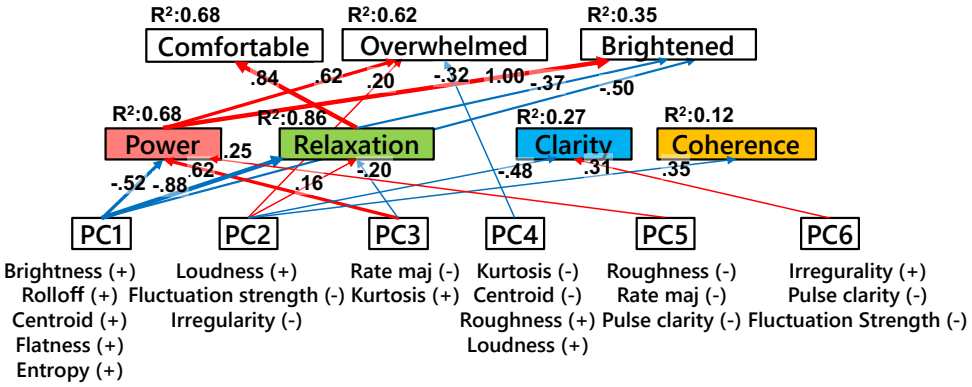


6.4 Acoustic Feature Extraction

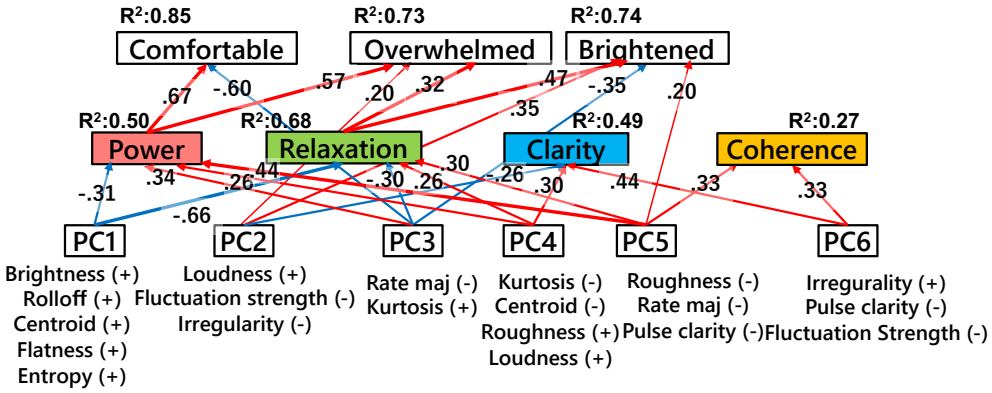
As explanatory variables capturing physical changes underlying impressions and values, we computed acoustic features from each stimulus using the MIR toolbox in MATLAB. We calculated 15 commonly used descriptors, mainly related to spectral information (e.g., loudness, sharpness, roughness, fluctuation strength; spectral statistics such as spectral centroid and flatness; pulse clarity; and major ratio). Because these descriptors are mutually correlated, we applied principal component analysis for dimensionality reduction. Based on cumulative explained variance, the first six principal components accounted for 99.5% of the variance; therefore, PC1–PC6 were used as representative acoustic variables in subsequent modeling.

7 USER-TYPE DIFFERENCES IN VALUE-FORMATION PATHWAYS

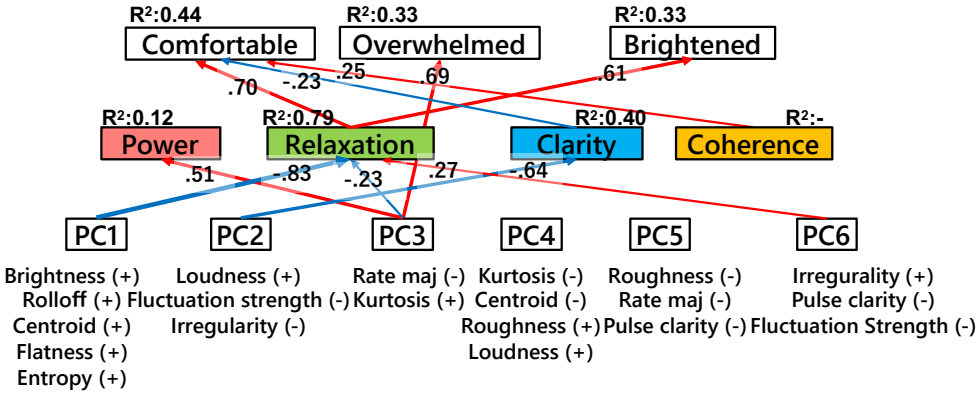
This chapter examines how value-formation pathways differ across the three user types (Emotion, Sound, and Environment) by estimating type-specific structural models. Figure 4 shows the value-structure models estimated for each user type. Model fit was generally good for all types: Emotion type (GFI = 0.97, AGFI = 0.92), Sound type (GFI = 0.97, AGFI = 0.88), and Environment type (GFI = 0.94, AGFI = 0.86).



(a) Emotion type



(b) Sound type



(c) Environment type

Figure 4. Structural models for each user type (emotion, sound, environment).

8 DISCUSSION

8.1 Significance of a Multi-layer Model for Representing EQ Preferences

In this study, EQ preference was modeled as a multi-layer process—acoustic features (principal components) → impressions → values—and we showed that the same modeling framework can be established for all three user types. This framework allows preference to be described not merely as a stronger or weaker liking, but as which impression cues are referenced and how they are translated into value responses (valence and arousal).

8.2 User-type-specific Approaches

Across types, the degree to which acoustic features explained impressions and values tended to be highest for the Sound type, followed by the Emotion type, and lowest for the Environment type. This implies that, particularly for the Environment type, contributions from factors that are difficult to capture by acoustic features alone—such as imagined situations and listening contexts—may be larger. In addition, the pathway forming the key value factor “Comfortable” differed by type: for the Sound type, the dominant pathway was mediated by Power (potency-related impressions), whereas for the Emotion and Environment types it was primarily mediated by Relaxation (low-activity impressions). From a practical standpoint, this suggests that EQ tuning strategies may be adapted by type: emphasizing a “power-oriented” direction for the Sound type, and a “relaxation-oriented” direction for the Emotion and Environment types. More broadly, the findings suggest that audio design can benefit from considering not only an average target response but also the subjective pathway through which users experience sound. If such evaluation viewpoints can be estimated or reflected in the design process, personalized or adaptive audio systems could support desired experiences—such as comfort, excitement, or ease of listening—through different acoustic strategies. Thus, the proposed model provides an interpretable basis for linking acoustic control parameters with user-perceived impressions and value responses.

9 CONCLUSION

This study showed that EQ preference can be modeled as a multi-layer process linking acoustic features, impressions, and value/emotion, with generally acceptable model fit. Categorizing participants into three user types (Emotion, Sound, and Environment) further revealed that value-formation pathways—such as the drivers of "Comfortable"—differ across types, supporting the usefulness of user-type-aware understanding and prediction. Future work should expand the sample size and experimental conditions, and integrate contextual factors such as use cases and listening environments.

REFERENCES

1. Olive, S. E., Welti, T., & McMullin, E. (2014). The influence of listeners' experience, age, and culture on headphone sound quality preferences. Audio Engineering Society Convention Paper.
2. Olive, S. E. (2019). Segmentation of listeners based on their preferred headphone sound quality profiles. Audio Engineering Society Convention Paper.
3. Katahira, K., Muto, K., Lee, N., Tobitani, K., Shiraiwa, A., Nakajima, K., Nagata, N., Kishino, F., Yamamoto, M., Kawasaki, K., Nikata, K., & Asano, T. (2016). Major factors in Kansei evaluation of 3D objects. *Transactions of Japan Society of Kansei Engineering*, *15*(4), 563–570. https://doi.org/10.5057/jjske.TJSKE-D-15-00087
4. Ono, Y., Aoyagi, S., Sugimoto, M., Yamazaki, Y., Yamamoto, M., & Nagata, N. (2019). Application of classification method of emotional expression type based on Laban movement analysis to design creation. In M. Kurosu (Ed.), *Human-Computer Interaction: Recognition and Interaction Technologies* (*Lecture Notes in Computer Science*, Vol. 11567, pp. 143–154). Springer. https://doi.org/10.1007/978-3-030-22643-5_11
5. Chamorro-Premuzic, T., & Furnham, A. (2007). Personality and music: Can traits explain how people use music in everyday life? *British Journal of Psychology*, *98*(2), 175–185. https://doi.org/10.1348/000712606X111177
6. Kelly, G. A. (1955). *The psychology of personal constructs*. W. W. Norton.
7. Goto, M., Hashiguchi, H., Nishimura, T., & Oka, R. (2003). RWC music database: Music genre database and musical instrument sound database. *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003)*.
8. International Organization for Standardization. (2017). ISO 532-1:2017 Acoustics—Methods for calculating loudness—Part 1: Zwicker method.
9. Sanui, J. (1996). Visualization of users' requirements: Introduction of the Evaluation Grid Method. *Proceedings of the 3rd Design & Decision Support Systems in Architecture & Urban Planning Conference*, 365–374.
10. Reynolds, T. J., & Gutman, J. (1988). Laddering theory, method, analysis, and interpretation. *Journal of Advertising Research*, *28*(1), 11–31. https://doi.org/10.1080/00218499.1988.12467766
11. Miyai, S., Katahira, K., Sugimoto, M., Nagata, N., Nikata, K., & Kawasaki, K. (2019). Hierarchical structuring of the impressions of 3D shapes targeting for art and non-art university students. In C. Stephanidis (Ed.), *HCI International 2019 – Posters Part I* (*Communications in Computer and Information Science*, Vol. 1032, pp. 385–393). Springer. https://doi.org/10.1007/978-3-030-23522-2_50
12. Shimozono, T., Sasai, D., Shiozawa, Y., Yamazaki, Y., Hashimoto, S., & Nagata, N. (2025). Understanding the structure of "liveness" and its value variations through large-scale affective evaluation. *Proceedings of the 18th International Conference on Music Perception and Cognition (ICMPC 18)*. Síntese Eventos.

Confidential

13. Sugimoto, M., Yagi, Y., & Nagata, N. (2023). How different tourist sites evoke different emotions: Investigation focusing on the urban and rural sites in Japan. In M. Kurosu & A. Hashizume (Eds.), *Human-Computer Interaction* (*Lecture Notes in Computer Science*, Vol. 14012). Springer. https://doi.org/10.1007/978-3-031-35599-8_21

14. Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, *18*(1), 39–43. https://doi.org/10.1007/BF02289026