









How Music–Video metaphors build destination brand resonance: Dyadic affect, meaning access, and cultural cues

Joston Gary ^{a,b,c,*} , Yang Gu ^{d,e} , Hannah Wang ^f , Xixing Zhou ^g , Yan Feng ^h ,
António Carrizo Moreira ^{i,j,k} 

^a Division of Business Administration, Department of Management and Engineering, Linköping University, SE-581 83, Linköping, Sweden

^b Department of Economics, Management and Industrial Engineering and Tourism, University of Aveiro, 3810-193, Aveiro, Portugal

^c Faculty of Business, City University of Macau, 999078, Taipa, Macao Special Administrative Region of China

^d Business Research Unit, ISCTE — University Institute of Lisbon, 1649-026, Lisboa, Portugal

^e Department of Organization, Strategy and Entrepreneurship, School of Business and Economics, Maastricht University, 6211 LK, Maastricht, Netherlands

^f Department of Music, School of Arts, University of Évora, 7000-872, Évora, Portugal

^g Institute of Social Sciences, University of Minho, 4710-057, Braga, Portugal

^h Centre for Research and Studies in Sociology, ISCTE — University Institute of Lisbon, 1649-026, Lisboa, Portugal

ⁱ DCSG — Department of Social Sciences and Management, Universidade Aberta (Open University), 1269-001, Lisboa, Portugal

^j NECE — Research Center for Business Sciences, University of Beira Interior, 6201-001, Covilhã, Portugal

^k CICEE — Research Center in Business and Economics, Autonomous University of Lisbon, 1150-293, Lisboa, Portugal

ARTICLE INFO

Keywords:

Music receptivity

Destination brand resonance

Cognitive metaphor theory

Circumplex model of affect

S-O-R theory

ABSTRACT

Short-form destination videos often rely on music to carry cultural meaning. This paper links Cognitive Metaphor Theory with the circumplex dyad of pleasure and arousal to explain how music–image pairings build destination brand resonance (DBR). Three experiments show that pleasure is the stable route to DBR, arousal helps only under favorable tone, and their effects are additive. A Meaning-Access Prime (MAP) raises both emotions under identical clips and, in Bayesian structural models, also exerts a direct path to DBR, strongest when pleasant tone is low. DBR then predicts destination brand identification and destination consumption intention. We also show a useful state view: Resonant versus Emergent DBR. The framework provides design rules for co-tuning tone, activation, and cultural cues in creator-made clips that improve resonance, identification, and intention.

1. Introduction

Consumer response to destination brand communication inhabits a theoretical crossroads between cultural cognition and emotional processing. Existing frameworks for destination branding communication, such as the cognitive–affective–conative (CAC) destination image model and customer-based destination brand equity (CBBE), often inadequately explain how culturally embedded non-verbal cues, such as music, shape consumer–brand relationships, leaving a gap between what consumers think and what they feel about a destination. Music can evoke vivid mental imagery and strong emotions that reshape destination image; however, these influences depend on the individual traits of the consumer, such as fantasy proneness and self-brand congruence (Fan et al., 2023). Irregular musical structures enhance perceptions of brand innovativeness by creating and resolving cognitive disfluency (Zoghaib

et al., 2023). Consumers engaging in destination brand communication translate sound into meaning through embodied metaphors (Asano et al., 2022). Moreover, music's social infrastructure, live venues, festivals, and performance networks, help anchor collective place identity and brand positioning (Baiocchi et al., 2024). The empirical studies suggest that music operates simultaneously along emotional, cognitive, metaphorical, and social pathways. However, existing research still treats these strands in isolation. Bringing these literature streams together, integrating cognition and emotion, can provide deeper insights into how consumers actually process brand-related stimuli in specific cultural contexts. We propose that music receptivity activates metaphorical processing, which then triggers the transmission of metaphorical clues, ultimately shaping destination brand resonance and destination consumption intention (see Figs. 4 and 5).

The need for such theoretical integration is underscored by the

* Corresponding author. Division of Business Administration, Department of Management and Engineering, Linköping University, SE-581 83 Linköping, Sweden.
E-mail addresses: li.rui@liu.se (J. Gary), Yang_Gu@iscte-iul.pt (Y. Gu), yudan.wang@fa.uevora.pt (H. Wang), id11123@uminho.pt (X. Zhou), Feng_Yan@iscte-iul.pt (Y. Feng), antonio.moreira@uab.pt (A.C. Moreira).

<https://doi.org/10.1016/j.jretconser.2025.104637>

Received 15 September 2025; Received in revised form 1 November 2025; Accepted 17 November 2025

Available online 22 November 2025

0969-6989/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

globalized tourism market. Cultural context significantly shapes consumers' experiences of destination brand communication, which increasingly incorporates symbolic elements embedded within music, visuals, and narratives. These symbolic elements are decoded using culturally influenced cognitive frameworks, making the alignment between symbolic branding and cultural cognition essential for the effectiveness of destination brand communication (Xu, 2024). Although cultural tailoring can enhance brand relevance and deepen consumer engagement, it also complicates the emotional internalization of brand messages (Liu, 2025; Ting and Ahn, 2023). Therefore, understanding the interplay between culturally grounded symbolism and emotional processing becomes crucial for marketers aiming to navigate diverse cultural landscapes (Osemwegie, 2025).

Extensive research on cultural cognition predominantly focuses on explicit verbal messages and overt brand positioning, largely ignoring how implicit, culturally embedded non-verbal cues, particularly music, affect tourists. Prior work demonstrates that music has a profound impact on pleasure and arousal, yet its influence on destination brand resonance remains underexplored (Fan et al., 2023; Moreno-Lobato et al., 2023; Zhuang et al., 2023). Music shapes pleasure and arousal, yet its influence on destination brand resonance remains underexplored (Fan et al., 2023; Moreno-Lobato et al., 2023). More critically, the field still lacks a process-level explanation of why music-evoked emotions would translate into deeper psychological ties with a destination and an intention to spend there (Coronel and Irimiás, 2022; Zhuang et al., 2023). That missing link motivates the present focus on Music Receptivity (MR).

The research defines Music Receptivity as a situational state, induced by brief instructions or cues, that expands access to culturally resonant metaphors embodied in melody, rhythm, and paired imagery. Under Cognitive Metaphor Theory, listeners convert these sound-image patterns into abstract destination meanings through well-learned metaphorical structures. This conversion should be especially effective in China, where metaphor pervades everyday language and visual symbolism (Chen and Liu, 2024; Lan and Yin, 2020; Tian and Abdullah, 2024; Zhang et al., 2022). When metaphorical meaning becomes readily accessible, listeners tend to register both heightened pleasure and increased physiological activation, aligning their emotions with the destination story.

On this reasoning, we test whether arousal, conceptualized as activation on the circumplex grid, moderates the influence of implicit cultural cues in short destination videos. We argue that when arousal is high, viewers in a high-context environment can assemble coherent meaning from sparse verbal content and densely coded musical-visual motifs, even when hedonic tone is muted. To examine this interaction, we expose a high-context audience to clips that orthogonally vary pleasure, arousal, and the richness of cultural embedding, and then assess how changes in music receptivity translate into destination brand resonance. This design isolates the boundary role of arousal in high-context message processing, an issue that prior work has not addressed.

This research advances knowledge in two complementary ways. High-context orientation is introduced as a boundary condition that governs how pleasure and arousal translate into destination brand resonance, clarifying when emotional responses gain traction in consumer-place relationships. In addition, Hall's context dimension is linked with the circumplex affect model and Cognitive Metaphor Theory, demonstrating how symbolic music-video combinations can bind viewers to a location during the brief exposures typical of social video platforms.

Through this framework, the study resolves three issues. First, previous studies have not specified how Music Receptivity channels emotional reactions within culturally framed destination experiences. Second, empirical ties between music-induced emotions and downstream outcomes, specifically destination brand resonance and brand identification, have been limited in scope. Third, previous research lacks an integrated account of how cultural context steers metaphor-based

emotional routes that connect music-driven affect to resonance. The findings address these gaps by demonstrating when and how arousal heightens metaphor access, thereby aligning emotional and cultural mechanisms in destination branding.

The research organizes propositions into three studies that isolate each stage of the process model and then validate the full pathway longitudinally. Study 1 manipulates pleasure and arousal to show their additive effects on destination brand resonance. Study 2 primes music receptivity and demonstrates its impact on pleasure, arousal, and resonance. Study 3 conducts a full test on participants to examine whether the process from initial music exposure to the emergence of resonance can predict brand identification and consumption intention.

2. Literature review

2.1. Cognitive Metaphor Theory within S-O-R and CMA framing

Cognitive Metaphor Theory (CMT) provides a conceptual basis for understanding how individuals derive meaning from abstract concepts through metaphorical mappings grounded in bodily and affective experience (Bakri, 2024; Lakoff and Johnson, 1980). Rather than treating metaphor as decorative language, CMT posits that metaphor is a core cognitive mechanism that shapes perception, emotion, and evaluative judgment. In branding communication, CMT has been widely used to explain how symbolic messages are conveyed not only through language but also through multimodal elements such as music, visual design, and narrative structure (Ayrton, 2020; Forceville, 2008; Peh, 2025; Requejo, 2016).

Within the Stimulus-Organism-Response (S-O-R) framework, destination music videos act as the stimuli; music receptivity (MR) and the emotional responses elicited by metaphorically encoded cues represent the organism; and destination brand resonance (DBR) and destination consumption intention (DCI) function as the outcomes. Emotional states are modeled using the Circumplex Model of Affect (CMA), which defines affective experience along two continuous dimensions: valence (pleasure-displeasure) and arousal (activation-deactivation) (Posner et al., 2005; Russell, 1980). This model has been applied extensively in marketing and tourism contexts to explain how external cues influence approach-oriented responses via emotional mechanisms.

In this research, MR is conceptualized as a potential antecedent to these two affective dimensions, though not as a necessary or exclusive driver. It is treated as a neutral cognitive-affective orientation, capable of influencing emotional responses but not essential. Prior S-O-R applications show that while emotions often mediate the effects of stimuli, direct paths from pre-existing appraisals or dispositions to brand-related outcomes can also occur, with the relative influence of indirect and direct paths varying by context, stimulus type, and audience (Blanco-Encomienda and Rosillo-Díaz, 2025; Khairani et al., 2025).

Extensions of CMT into multimodal research confirm that metaphors can be distributed across sensory channels, auditory, visual, and narrative, and integrated as cohesive emotional cues. For example, upward melodic progression may symbolize growth or aspiration, while visuals like light piercing clouds can suggest hope or renewal. These metaphorical structures are not passive codes but emotionally potent stimuli that elicit affective reactions, namely, valence (positive emotional tone) and arousal (intensity of activation), as defined by the CMA framework (Noble, 2022; Golubkova, 2021). In S-O-R terms, metaphor-laden musical-visual stimuli trigger CMA-based emotional states, which in turn influence DBR and DCI. Research in marketing and tourism supports this pathway: external stimuli shape affective responses along valence and arousal dimensions, which then predict approach outcomes such as loyalty, satisfaction, and behavioral intent (Cao et al., 2024; Jiang & Garrod, 2025; Zhu et al., 2025).

The emotional impact of metaphor, however, is moderated by contextual factors such as individual sensitivity to metaphorical cues

and competing cognitive frames (Demjén and Semino, 2020; Ervas et al., 2021). In this framework, MR acts as such a moderator. It refers to a cognitive–affective readiness to perceive and respond to culturally embedded metaphors in sound and imagery. Under S–O–R and CMA, MR is positioned as an antecedent to valence and arousal, and also as a potential direct predictor of DBR and DCI, consistent with the idea that organismic traits may both influence emotional states and independently shape downstream responses.

Metaphor plays a particularly salient role in the Chinese cultural context, where communication often relies on symbolic expression, indirectness, and shared cultural imagery (Zhang, 2024). Compared to Western cultures that favor explicit verbal exchange (e.g., the U.S., Sweden), Chinese consumers more frequently engage in metaphorical thinking to convey meanings related to harmony (he), balance (ping-heng), prosperity (fanrong), and traditional values such as filial piety (xiao). Music embedded with culturally familiar metaphors offers an effective, implicit pathway to emotional and cognitive engagement. In such contexts, valence and arousal, as modeled by CMA, capture how culturally embedded stimuli shape emotional interpretation and, through it, brand equity and loyalty.

In destination-oriented music videos (TikTok), culturally embedded musical metaphors, such as pentatonic riffs evoking harmony or genre hybrids suggesting modernity, function as emotional triggers. When these cues are processed fluently, they generate elevated valence and arousal, which in turn enhance mental imagery and conceptual engagement. Music receptivity intensifies this process by increasing sensitivity to symbolic detail and emotional tone. As a result, higher receptivity facilitates stronger thematic alignment with destination meaning, reinforcing both brand resonance and consumption intention.

2.2. High-context culture and consumer meaning-making

Hall's (1976) high-versus low-context framework proposes that the balance between explicit language and implicit, situationally bound cues varies systematically across cultures. China sits toward the high-context pole, with metaphoric scripts and cultural archetypes woven into everyday speech, visual arts, and musical traditions.

Context orientation is best understood as an individual-level construct rather than a national proxy. While early studies relied on country dummies to infer high-versus low-context tendencies, distinctions often dissolve when person-level measures are included. Found that Belgian consumers, who scored higher in context sensitivity, interpreted metaphor-rich advertisements as more appealing and less cognitively demanding than Dutch consumers (Hornikx and Le Pair, 2017). However, this effect disappeared once individual context orientation was controlled, highlighting the explanatory power of personal orientation over nationality. Similarly, in the domain of communication channel preferences, Japanese students with a high-context orientation favored relational modes, such as face-to-face or phone conversations, while low-context U.S. students preferred email (Richardson and Smith, 2007a,b). This conceptual evolution—from context as a fixed cultural trait to a flexible individual tendency—is further substantiated by evidence from the digital sphere. An individual's context orientation, for instance, predicts their trade-offs between a social platform's technical features and its relational affordances, even within the same culture (Yang et al., 2021). This cognitive disposition is understood not as innate, but as an emergent outcome of ongoing cultural immersion and personal experience (Wu et al., 2023). Consequently, validated instruments have been developed to measure this individual-level construct, allowing for more nuanced cross-cultural comparisons than those based on national proxies (Warner-Söderholm, 2013). Adding further nuance, communication styles in East Asia reflect not just cultural norms, but also accumulated intercultural exposure, suggesting that context orientation evolves through both cultural socialization and life experience (Wu et al. (2023).

Across this body of evidence, two regularities stand out. First, high-

context audiences process metaphor and other implicit cues with less cognitive strain and often report more favorable attitudes (Kong and Lou, 2023). Second, the explanatory power of context orientation primarily operates at the individual level, cautioning against the use of simple national-culture proxies (Haoyue and Cho, 2024). These insights are crucial for destination branding because short-form videos condense verbal content while saturating musical and visual symbolism. If context orientation conditions how quickly viewers decode symbolic material, then high-context audiences should integrate that material into their brand judgments more readily than low-context audiences.

Metaphor occupies a central place in the Chinese aesthetic tradition. Classical poetry employs layered nature imagery, ink painting conveys moral virtues through landscape, and pentatonic musical structures evoke shared affective codes (Yu, 2023). Contemporary marketing employs these conventions by incorporating guzheng riffs, brushstroke animations, and cosmological references into video clips. Such cross-modal metaphors are likely to resonate with a high-context audience precisely because the cues align with well-practiced decoding scripts (Liu and Li, 2022).

Short-form video communication compresses language while amplifying symbolic density, placing greater weight on the viewer's ability to decode metaphor and visual–musical cues. High-context orientation plays a central role in this process, not as a cultural backdrop, but as an active moderator of how symbolic stimuli are interpreted and mapped onto brand meaning. In destination branding, this decoding advantage positions high-context viewers to extract emotional and destination brand value more efficiently, shaping how they form judgments even in brief, media-rich encounters.

3. Hypothesis and experimental model

3.1. Music receptivity and dyadic emotional valence

In the S–O–R framework, destination TikTok videos serve as the stimuli; pleasure and arousal represent the internal (organismic) states; and downstream judgments and behavioral intentions function as the responses (Elshazly et al., 2025; Liu, Jiang, & Muhammad, 2024). The CMA model identifies pleasure and arousal as the core emotional dimensions that mediate how environmental inputs translate into approach-oriented outcomes (Bourdeau et al., 2024; Jaeger et al., 2022; Liu, Y et al., 2023; Liu, 2025). This model has been widely applied in marketing and tourism research to examine how communication cues shape evaluation and behavior (Chen & Cheng, 2023; Zhang et al., 2022).

CMT helps explain how music–visual pairings influence emotional states (Elshazly, Yan & Ghaith, 2025; Liu et al., 2023). Through metaphorical mappings, listeners associate musical elements, such as motion, harmony, and timbre, with abstract meanings tied to place (Tham et al., 2023; Zhu et al., 2025). Multimodal studies show that images and sound together can convey these mappings effectively (Fong et al., 2024). Research also indicates that individual differences in sensitivity to metaphor framing can alter emotional interpretations (Liu et al., 2023). Tonal processing engages brain regions associated with affect and figurative meaning, and music linked to a destination has been shown to reshape its image by triggering emotion and visual imagination (Tham et al., 2023; Zhu et al., 2025).

In this study, music receptivity is treated as a temporary orientation shaped by the stimulus set and accompanying instructions (Fong et al., 2024). It primes participants to attend more closely to culturally embedded cues, thereby intensifying their emotional response, specifically in terms of pleasure and arousal, during video exposure (Wu et al., 2023). As such, the following hypotheses are put forward.

Hypothesis 1. Music receptivity, Meaning-Access Prime (MAP) versus a neutral condition, will increase the pleasure experienced during exposure to culturally embedded destination music videos.

Hypothesis 2. Music receptivity, MAP versus Neutral, will increase the arousal experienced during exposure to culturally embedded destination music videos.

3.2. Dyadic emotional valence and destination brand resonance

Dyadic emotional valence, comprising pleasure and arousal from the circumplex model, represents the internal (organismic) state linking symbolic stimuli to evaluative responses. Pleasure reflects positive hedonic appraisal, while arousal captures the level of emotional activation. They summarize the affective impact of short-form video exposure (Ma et al., 2023). Russell's circumplex model formally defines these two axes, and research in consumer psychology consistently shows that the effect in these dimensions predicts judgments and approach tendencies (Cheng & Huang, 2022).

CMT offers insight into how culturally embedded music–visual pairings influence this affective dyad (Coronel and Irimías, 2022). Metaphorical mappings connect sonic motion, timbral texture, and visual motifs to abstract place-related meanings (Fan et al., 2023). These connections shape how viewers emotionally interpret a destination (Bai et al., 2023). In marketing and tourism studies, emotions triggered by such symbolic cues are shown to influence brand attitudes, loyalty, and destination appeal.

In branding research, affective responses frequently mediate the effects of communication cues on higher-order judgments (Wu et al., 2023). Foundational studies linked advertising-induced pleasure and arousal to favorable brand perceptions, while subsequent work confirmed that affective activation enhances recall and evaluative positivity (Luo and Mattila, 2023). In destination branding, specifically, music and imagery adjust viewers' perceptions of a place by evoking emotional and mental imagery (Moreno-Lobato et al., 2023). This supports the idea that dyadic valence serves as the immediate emotional pathway to brand resonance.

Destination brand resonance (DBR) represents a personal, affective connection to a place, encompassing elements of attachment, identification, and perceived symbolic significance (Nikraftar et al., 2024). When a video communicates culturally familiar or resonant meanings, higher levels of pleasure and arousal are expected to deepen brand resonance (Zhou and Jiao, 2024). The affective dyad thus serves as an index of how strongly a stimulus engages the viewer and aligns with their internal image of the destination.

Empirical studies in services and tourism consistently demonstrate that emotional states, such as pleasure and arousal, predict brand equity, satisfaction, and behavioral intentions (Tang et al., 2024). This evidence supports their treatment as direct precursors to destination brand resonance in the current study. Thus, the following hypotheses are put forward.

Hypothesis 3. High levels of pleasure will lead to greater destination brand resonance compared to low levels of pleasure during exposure to culturally embedded destination music videos.

Hypothesis 4. High levels of arousal will lead to greater destination brand resonance compared to low levels of arousal during exposure to culturally embedded destination music videos.

3.3. Music receptivity and destination brand resonance

DBR refers to a consolidated psychological bond that encompasses attachment, identification, and perceived symbolic relevance to a place. Within the brand equity framework, DBR is considered the apex response, emerging when meaning construction has been successful (Jiménez-Marín et al., 2021; Qiu et al., 2024; Prasetyo and Ilimawan, 2024).

Metaphorical mappings connect sonic elements, such as movement, harmony, and timbre, to abstract representations of place. Multimodal research demonstrates that audiovisual elements in advertising jointly

encode these mappings (Zhuang et al., 2023; Forceville, 2008; Jia, 2024; Xu and Yang, 2025). When viewers are oriented toward recognizing such metaphorical structures, alignment with destination meanings becomes more likely, even in the absence of heightened emotional states (Malkawi, 2025; Johnson and Jyothula, 2025). Cultural psychology further supports this mechanism by showing that individuals interpret symbolic stimuli through culturally shared schemas; thus, heightened receptivity to culturally embedded cues facilitates faster cognitive congruence with destination identity (Landau et al., 2018; Mohamed et al., 2025; Shavitt & Cho, 2016; Zhong et al., 2025).

Accordingly, the MR in this study is expected to increase DBR not only indirectly through emotional valence (pleasure and arousal) but also through a direct cognitive route. MR functions as an appraisal-level orientation that enhances access to culturally structured metaphors, thereby promoting immediate semantic fit with destination meanings. Prior findings indicate that metaphor framing influences consumer evaluation, narrative-style processing strengthens self-brand connection, and emotionally resonant destination content enhances brand equity (Escalas, 2004; Lin et al., 2025). These insights support the existence of a direct pathway from music receptivity to DBR, operating in parallel with affective mediation. Therefore, the following hypothesis is put forward.

Hypothesis 5. Music receptivity, exposure to the MAP condition, compared to a neutral condition, positively influences destination brand resonance during viewing of culturally embedded destination music videos.

3.4. Destination brand awareness and outcome

DBR reflects more than momentary liking: it captures an enduring psychological bond with a place, shaped by attachment, identity, and perceived relevance. In branding models, resonance is not just a reaction, but the culmination of meaning-making that precedes loyalty behaviors, such as advocacy or purchase intention. In tourism research, similar patterns are observed: visitors who feel emotionally connected to a destination are more likely to return, recommend it to others, or form lasting brand attachments (Guleria et al., 2023). Viewed through this lens, DBR functions as a key inflection point, where symbolic engagement gives rise to behavioral momentum (Zhang & Zhu, 2024).

Culturally familiar sonic and visual cues, when mapped onto abstract destination meanings, can form structured associations that foster self-relevance and perceived alignment with the destination. These metaphorically grounded linkages create the psychological coherence necessary for brand identification and planned consumption. Rather than operating solely through emotion, DBR reflects a deeper integration of symbolic meaning and self-concept. Prior research demonstrates that narrative-like brand engagement strengthens identification, while attachment-like bonds increase commitment-related behaviors (Yang et al., 2021; Rather et al., 2020). Studies on short-form destination media further show that emotional resonance and perceived self-congruity with destination content enhance both identification and loyalty. These patterns establish DBR as a cognitive-emotional anchor that facilitates both destination brand identification (DBI) and destination consumption intention (DCI) (Cao et al., 2024; Le et al., 2024).

DBR is modeled as the bridge between the organismic state and action-oriented responses. Higher pleasure and arousal indicate stronger affective engagement with the stimulus; DBR captures whether that engagement consolidates into a bond that carries forward to identity alignment and planned spending. Evidence from branding and tourism studies shows that such consolidated bonds are related to identification, loyalty, and intention outcomes (Rather et al., 2020; Yen et al., 2020). As such, the following hypotheses are put forward.

Hypothesis 6. Destination brand identification will be higher under the Resonant condition of destination brand resonance than under the Emergent condition.

Hypothesis 7. Destination consumption intention will be higher under the Resonant condition of destination brand resonance than under the Emergent condition.

According to research hypotheses, we constructed the following experimental model show in Fig. 1.

4. Research study

4.1. Research design

The research program consists of three experiments that use short destination-promotion videos sourced from TikTok and YouTube creators. Each clip pairs culturally embedded music with aligned visuals, excluding slogans and voiceovers. The conceptual frame is S–O–R with dyadic emotional valence (pleasure, arousal) as the organismic state. Study 1 varies the emotional profile of the music and the cultural context of the video to test whether pleasure and arousal independently predict DBR. Study 2 induces a meaning-access state of music receptivity versus a neutral orientation and examines effects on pleasure, arousal, and DBR. Study 3 follows participants to test whether DBR measured predicts DCI.

All studies were run online with adult short-video users through the Tencent questionnaire platform, and participants were recruited on Credamo. Each participant received €2.5 (≈20 CNY). The same stimulus pool, exposure procedure, and measurement battery were used across studies, with only the focal manipulation differing by study, and assignments were randomized within each study. Because the research budget was fixed and payment per participant could not be raised, sample sizes were planned in G*Power 3.1 to reach 0.80 power at $\alpha = .05$ for small-to-moderate effects, and identical per-cell quotas were opened (200 per condition; 800 per study). This budget-driven choice explains why the three studies have very similar Ns and it keeps sampling variance comparable across studies. Attention checks and a 30-s minimum exposure were enforced at the platform level, and all items were set to forced response, so item-level missingness is zero. Under this design, the response rate (starters/invited) was 100 % in all studies, and

the analyzed-sample rates after these screens were 76.8 % (Study 1, 614/800), 75.8 % (Study 2, 606/800), and 77.5 % (Study 3, 620/800). Exposure-time distributions showed typical dwell well above 30 s and no heaping at the cutoff, and “cell-level exclusions” refer only to removals by experimental condition caused by the exposure-time rule and the attention checks, not to missing items. Data analysis involved ANOVA and multiple regression using SPSS 26.0, Bayesian structural equation modeling with Mplus 8.3, and response surface analysis (RSA) to explore interaction effects within the model. Demographic and background covariates (age, gender, music training, career, brand preference) were entered in each cell model and retained only if significant; a consolidated summary is reported in Supplementary C. In Study 1, this procedure involved keeping music training (Group 1) and career (Group 2) separate. In Study 2, it kept age, career, and brand preference in the low-pleasure cells. In Study 3 (Bayesian SEM), covariates were not included to maintain model stability. Because Study 3 estimated multi-group BSEM models with weakly informative priors, adding five demographic covariates to each group would have reduced the effective degrees of freedom and destabilized the PPP, so the covariates were not included in this analysis. A consolidated sample-flow diagram for the three studies is provided in Supplementary_C - Figs. S1–S3. Supplementary_C - Appendix B2 provides the results of the Covariate effects.

Stimuli and procedure. Each participant viewed one randomly assigned clip (30–90 s), then completed the measures. For Study 1, clips were edited to yield combinations of high versus low pleasure and high versus low arousal, crossed with high-context versus low-context cultural embedding. For Study 2, instructions and exemplars induced MAP versus Neutral prior to viewing; the clip set itself was held constant. Study 3 implemented a one-time experimental procedure to test the full structural model, using the same initial setup as Studies 1 and 2 while incorporating additional measures for DCI. A clip list with visual and musical feature descriptors, together with the exact MAP prompt shown for each representative video, is provided in Supplementary_E, as copyright and funding rules do not allow us to release the original files.

Measures. Pleasure and arousal were measured with 7-point items adapted from CMA. DBR captured psychological connection, identification with the destination, and perceived relevance. DCI reflected

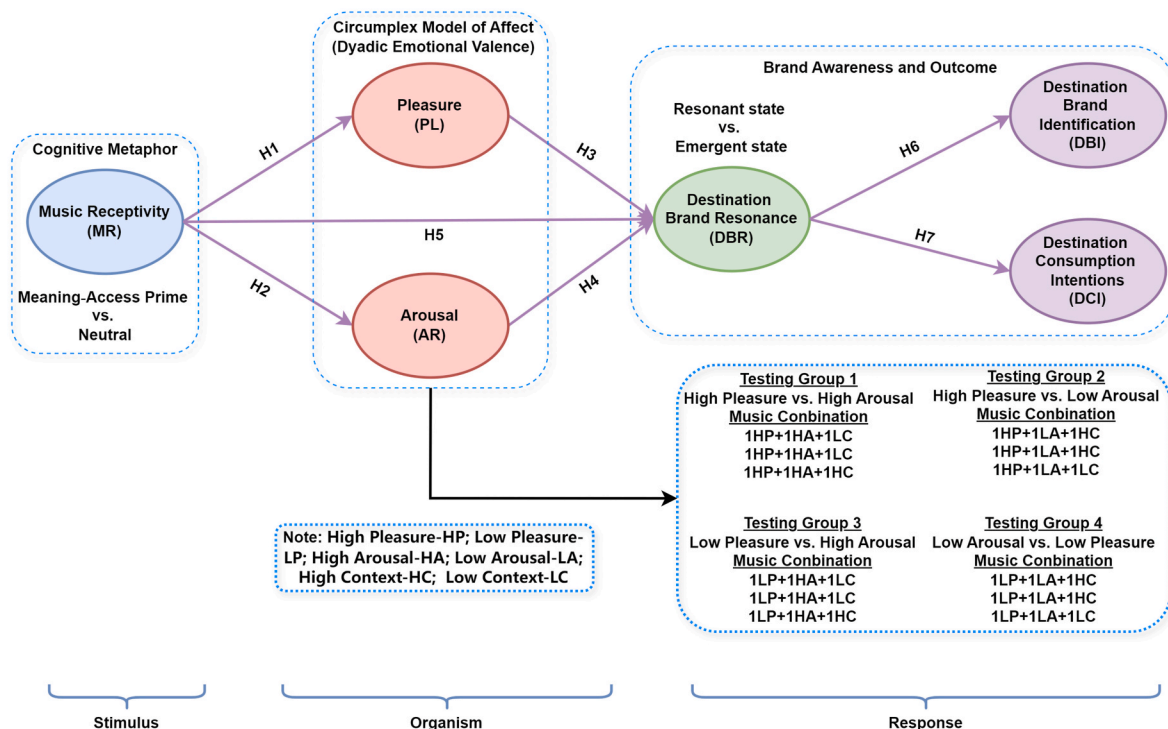


Fig. 1. Experimental model.

destination consumption intention, defined as the willingness to spend during a hypothetical visit within the next 12 months. A three-item appraisal check assessed perceived metaphorical richness and story imagery. Unless noted otherwise, all scales used a 1-to-7 point scale and showed acceptable internal consistency. For the reliability and validity of the scale, EFA/CFA, multicollinearity, and other numerical values, refer to Supplementary_C - Appendix A to B1; for the items of the original scale, please refer to Supplementary_D. In the same supplementary file, a multi-group measurement invariance test (configural, metric, and scalar) across the four experimental cells and MAP/Neutral conditions is reported, showing $\Delta CFI \leq 0.010$ at each step, which supports comparing structural paths across groups.

4.2. Manipulation check

The three-item appraisal check (“the music signaled symbolic meanings about the place,” “sound and visuals worked together to suggest abstract ideas,” “the video prompted story-like images”) formed a single factor. Across all participants pooled over the three studies ($N = 1840$), internal consistency was $\alpha = .86$, with standardized loadings from 0.72 to 0.84. The mean was 5.38 ($SD = 0.93$), well above the midpoint of 4, one-sample $t(1,839) = 63.65$, $p < .001$, Cohen’s $d = 1.48$; the 95% CI for the mean was [5.34, 5.42]. These results indicate that, on average, participants perceived the clips as conveying abstract, culturally grounded meanings.

Study-level checks aligned with the intended manipulations. In Study 1 ($N = 614$), clips designed to produce higher pleasure yielded higher reported pleasure than low-pleasure clips, $M_{\text{High-P}} = 5.62$ versus $M_{\text{Low-P}} = 4.58$, $SD_{\text{pooled}} = 1.02$, $t(612) = 12.63$, $p < .001$, $d = 1.02$. Clips designed to produce higher arousal exceeded low-arousal clips, $M_{\text{High-A}} = 5.33$ versus $M_{\text{Low-A}} = 4.43$, $SD_{\text{pooled}} = 1.08$, $t(612) = 10.32$, $p < .001$, $d = 0.83$. High-context edits scored higher on the appraisal check than low-context edits, $M_{\text{HC}} = 5.49$ versus $M_{\text{LC}} = 5.28$, $SD_{\text{pooled}} = 0.94$, $t(612) = 2.77$, $p = .006$, $d = 0.22$, confirming that cultural embedding was perceptible without introducing large differences in baseline liking.

In Study 2 ($N = 606$), the music-receptivity induction produced the intended state difference. Participants in the MAP condition reported higher appraisal-check (music-receptivity scale items) scores than those in the Neutral condition, $M_{\text{MAP}} = 5.56$ versus $M_{\text{Neutral}} = 5.20$, $SD_{\text{pooled}} = 0.95$, $t(604) = 4.66$, $p < .001$, $d = 0.36$. Pleasure and arousal means were above the scale midpoint overall, $M_{\text{P}} = 5.10$ ($SD = 1.07$) and $M_{\text{A}} = 4.92$ ($SD = 1.12$).

Scale reliability was strong across studies. DBR showed $\alpha = .91$ pooled; DCI showed $\alpha = .89$ pooled. Pleasure and arousal reliabilities were $\alpha = .88$ and $\alpha = .85$, respectively. In Study 3 ($N = 620$), a brief descriptive comparison was run to show the downstream relevance of DBR. DBR at Time 1 was split at its median to classify participants into “Resonant” and “Emergent” states, and the “Resonant” group reported higher DCI ($M_{\text{Resonant}} = 5.18$ versus $M_{\text{Emergent}} = 4.36$, $SD_{\text{pooled}} = 1.20$, $t(618) = 8.51$, $p < .001$, $d = 0.68$). This descriptive pattern is consistent with treating DBR as a meaningful driver, but the formal tests of this role rely on the continuous DBR scores in the structural models reported later in Section 4.5.

These checks support three conclusions relevant to the S–O–R specification. First, participants perceived the videos as carrying symbolic, culturally embedded meaning. Second, the emotional manipulation in Study 1 and the receptivity induction in Study 2 operated as intended. Third, DBR functioned as a proximal bridge to intention, which justifies its role as the response that precedes DCI in Study 3.

4.3. Study 1

This study tested the independent and interactive effects of pleasure and arousal on destination brand resonance (DBR) using a 2 (pleasure: high vs. low) \times 2 (arousal: high vs. low) between-subjects design. The

design isolated the causal roles of the two CMA dimensions while keeping destination identifiers absent. It was predicted that pleasure would have a positive influence on DBR across conditions and that arousal would contribute to DBR primarily when pleasure was high. Each condition paired a brief scene with a musical excerpt selected to elicit the target affective tone. Two music experts (with Chinese and Portuguese training) curated materials to embody culturally relevant metaphors. Traditional-style pieces with gentle pentatonic melodies signified balance and collective harmony; contemporary pieces fused traditional instruments (e.g., guzheng, erhu) with modern rhythms (e.g., ancient-style animation cues, rock), signaling dynamism and continuity. The same music pool was used in Studies 1 to 3 to ensure comparability and reduce testing bias. Participants were randomly assigned to one of the four affect conditions and viewed emotionally congruent destination music-video stimuli that had been pretested to target the intended affective profiles. This enabled an assessment of whether DBR is driven chiefly by hedonic tone (pleasure), by activation (arousal), or by their combination, establishing a baseline emotional mechanism for the subsequent studies.

4.3.1. Data description

A four-condition, between-subjects experiment was conducted. The full sample consisted of 614 participants recruited via Credamo and randomly assigned to high pleasure/high arousal ($n = 150$), high pleasure/low arousal ($n = 159$), low pleasure/high arousal ($n = 150$), or low pleasure/low arousal ($n = 155$). Each participant viewed a 30-s to 60-min destination-neutral video to target one of the four affective states. Narrative structure, destination identity cues, and other content confounds were held constant. After exposure, participants reported pleasure and arousal on seven-point semantic differentials adapted from Russell (1980) and then completed the DBR scale. Reliability exceeded $\alpha = .85$ in all four groups. Individual-level covariates were included only when significant in the model for that cell: self-taught music training in Group 1 and occupational category (“public employee”) in Group 2. No destination names or logos were shown.

4.3.2. Results

All analyses in this study were based on the following regression equation:

$$(a) \text{DBR}_i = \beta_0 + \beta_1 \cdot \text{Pleasure}_i + \beta_2 \cdot \text{Arousal}_i + \beta_3 \cdot (\text{Pleasure}_i \times \text{Arousal}_i) + \beta_4 \cdot \text{Covariates}_i + \varepsilon_i$$

where all continuous predictors were mean centered prior to computing the interaction term. The dependent variable was DBR, and covariates were included based on model-specific significance.

In the high pleasure/high arousal group ($M_{\text{pleasure}} = 5.1850$, $SD = 1.3528$; $M_{\text{arousal}} = 4.8233$, $SD = 1.2703$; $M_{\text{brand}} = 4.7267$, $SD = 1.4367$), arousal significantly predicted destination brand resonance ($B = 0.535$, $p < .001$), while pleasure did not reach significance ($B = 0.077$, $p = .584$). The interaction term was also non-significant ($B = -0.055$, $p = .397$). Interestingly, self-taught music training emerged as a significant covariate in this group ($B = 0.914$, $p = .025$), suggesting that individuals with informal musical expertise may be more sensitive to symbolic cues embedded in high-intensity emotional content. The model accounted for a substantial proportion of variance ($R^2 = 0.439$, Adjusted $R^2 = 0.326$), with an overall model fit of $F(25, 124) = 3.881$, $p < .001$, indicating that emotional arousal, when accompanied by high pleasure, robustly contributes to symbolic destination brand resonance.

A different pattern emerged in the high pleasure/low arousal group ($M_{\text{pleasure}} = 5.9066$, $SD = 0.6117$; $M_{\text{arousal}} = 5.4198$, $SD = 0.8698$; $M_{\text{brand}} = 5.0940$, $SD = 0.7314$). Here, both pleasure and arousal were significant predictors of destination brand resonance ($B = 0.505$ and 0.287 , respectively; $ps < 0.001$), while the interaction remained non-significant ($B = -0.058$, $p = .486$). The negative effect of occupational category (career = 2) was statistically significant ($B = -0.481$, p

= .040), implying that professional orientation may attenuate affective openness in more reflective, low-arousal environments. The model's explanatory power was strong ($R^2 = 0.495$, Adjusted $R^2 = 0.409$), with $F(25, 124) = 5.746$, $p < .001$, reinforcing the notion that both emotional tone and energy contribute additively to destination brand resonance when pleasure is high.

In the low pleasure/high arousal condition ($M_{\text{pleasure}} = 5.5250$, $SD = 0.9640$; $M_{\text{arousal}} = 5.0817$, $SD = 1.0974$; $M_{\text{brand}} = 5.7086$, $SD = 0.9738$), only pleasure predicted destination brand resonance ($B = 0.505$, $p < .001$). Neither arousal ($B = 0.039$, $p = .660$) nor the interaction term ($B = 0.006$, $p = .926$) was significant, suggesting that physiological activation alone, when not supported by a hedonic frame, lacks persuasive force in destination branding contexts. This interpretation is further supported in the low pleasure/low arousal group ($M_{\text{pleasure}} = 5.4000$, $SD = 0.8588$; $M_{\text{arousal}} = 4.8823$, $SD = 1.1940$; $M_{\text{brand}} = 5.4184$, $SD = 1.0150$), where pleasure again significantly predicted destination brand resonance ($B = 0.544$, $p < .001$), and both arousal and the interaction term were non-significant ($B = 0.030$ and 0.054 , $ps > 0.48$). Although variance explained was lower in the latter two models ($R^2 = 0.396$ and 0.331 ; Adjusted $R^2 = 0.280$ and 0.213), both reached conventional significance thresholds with $F(24, 125) = 3.418$, $p < .001$ and $F(25, 124) = 2.817$, $p < .001$, respectively, and yielded medium-to-large effect sizes ($\eta^2 > 0.14$ across all groups).

Across conditions, the absence of interaction effects indicates that pleasure and arousal do not combine multiplicatively to shape destination brand resonance. Instead, their effects appear to be largely additive and context-dependent. Arousal exerts a positive effect only when paired with high pleasure or when the overall emotional environment supports motivational intensity. In contrast, pleasure is a more robust and consistent effect, functioning as a primary driver of symbolic destination brand meaning even in emotionally muted conditions.

Fig. 2 summarizes the main effects by averaging the relevant cells (e.g., Groups 1 + 2 vs. 3 + 4 for pleasure; Groups 1 + 3 vs. 2 + 4 for arousal). The left panel shows the mean DBR difference between nominal high- and low-pleasure conditions; the right panel shows the smaller difference between nominal high- and low-arousal conditions. Error bars depict standard deviations (see Fig. 3).

Three-dimensional response-surface analyses (RSA) were estimated for each cell using second-order polynomials to visualize continuous relationships. The centered equations and balance/imbalance slices are reported in Equations (1)–(12). In Group 1, the balance slope was large and positive, and the surface rose with intensity. In Group 2, the balance

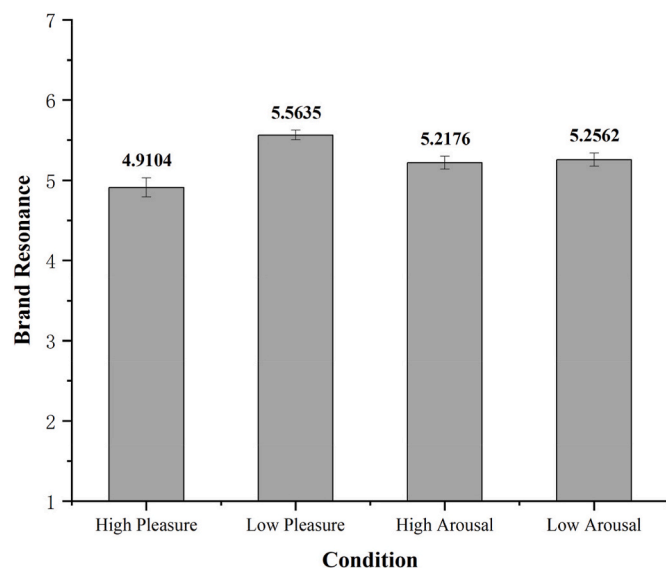


Fig. 2. Main effects of pleasure and arousal on destination brand resonance.

path climbed steadily while the imbalance slice showed a U-shape. In Group 3, imbalance curvature was positive, indicating a valley when arousal rose without corresponding pleasure. In Group 4, the surfaces were relatively flat. The overlaid slices indicated two regularities: under moderate or lower intensity (Groups 2–4), balanced pleasure–arousal profiles outperformed imbalanced profiles; once both cues were high (Group 1), intensity dominated. For ease of exposition, we quote the composite slope and curvature terms advocated by Kim, Steinhoff and Palmatier (2022):

$$\text{Slope}_{\text{bal}} = b_1 + b_2, \text{Curv}_{\text{bal}} = b_3 + b_4 + b_5, \text{Slope}_{\text{imb}} = b_1 - b_2, \text{Curv}_{\text{imb}} = b_3 - b_4 + b_5.$$

Centered regression equation:

$$\text{DBR}_t = 4.58020 + 0.01360 * u + 0.71560 * v - 0.43763 * u * v + 0.22070 * u^2 + 0.21792 * v^2 \quad (1) - \quad \text{Group 1}$$

$$\text{DBR}_t = 5.88842 + 0.50243 * u + 0.29474 * v - 0.16911 * u * v - 0.07782 * u^2 + 0.04294 * v^2 \quad (2) - \quad \text{Group 2}$$

$$\text{DBR}_t = 5.65266 + 0.52718 * u + 0.09411 * v - 0.10402 * u * v - 0.13347 * u^2 + 0.00357 * v^2 \quad (3) - \quad \text{Group 3}$$

$$\text{DBR}_t = 5.30949 + 0.50759 * u + 0.09707 * v - 0.00081 * u * v - 0.04003 * u^2 + 0.09725 * v^2 \quad (4) - \quad \text{Group 4}$$

Balance line equation ($PL = AR, u = v = t$):

$$\text{DBR}_t = 4.58020 + 0.00095 * t^2 + 0.7292 * t \quad (5) - \quad \text{Group 1}$$

$$\text{DBR}_t = 5.88842 - 0.04835 * t^2 + 0.79717 * t \quad (6) - \quad \text{Group 2}$$

$$\text{DBR}_t = 5.65266 - 0.03302 * t^2 + 0.62129 * t \quad (7) - \quad \text{Group 3}$$

$$\text{DBR}_t = 5.30949 - 0.05803 * t^2 + 0.60466 * t \quad (8) - \quad \text{Group 4}$$

Imbalance line equation ($PL = -AR, u = t, v = -t$):

$$\text{DBR}_t = 4.5802 + 0.87621 * t^2 - 1.0299 * t \quad (9) - \quad \text{Group 1}$$

$$\text{DBR}_t = 5.88842 + 0.28987 * t^2 + 0.20769 * t \quad (10) - \quad \text{Group 2}$$

$$\text{DBR}_t = 5.65266 + 0.24106 * t^2 + 0.43307 * t \quad (11) - \quad \text{Group 3}$$

$$\text{DBR}_t = 5.30949 + 0.05641 * t^2 + 0.41052 * t \quad (12) - \quad \text{Group 4}$$

In Group 1, the content is euphoric, and the score is correspondingly exuberant. The balance slope is large and positive ($a_1 = 0.80$, $p < .01$), but the imbalance curvature is likewise positive ($a_4 = 0.31$, $p < .05$); the surface therefore rises almost monotonically. Intensity, not symmetry, drives destination brand resonance: incremental coordination yields limited upside once both cues sit in the upper-right quadrant. The content already positions participants in a “celebratory” frame. The high-energy soundtrack, therefore, reinforces rather than creates the effect. Because the prime and the score are effectively consonant, additional symmetry ($PL = AR$) yields only marginal gains; intensity dominates. This is visible in the near-planar ridge along both balance and imbalance slices.

In Group 2, the music supplies positive valence. The balance path climbs steadily ($a_1 = 0.80$, $a_2 = 0.05$). In contrast, the imbalance slice is U-shaped with its trough near the stimulus point ($a_3 = -0.21$, $a_4 = 0.29$). Affective mis-alignment, high pleasure in a low-energy context, dampens destination brand resonance. Restoring arousal (or lowering pleasure cues) would move the consumer up the balance ridge. Restoring arousal (or lowering pleasure cues) would move the consumer up the balance ridge. The affective gap (low arousal in text, high pleasure in music) manifests asymmetry cost, precisely the dip along the imbalance slice.

In Group 3, the content is tense or urgent, but the music is dark rather than uplifting. Music reinforces arousal but offers little valence. The imbalance curvature is strongly positive ($a_4 = 0.24$, $p < .01$), producing

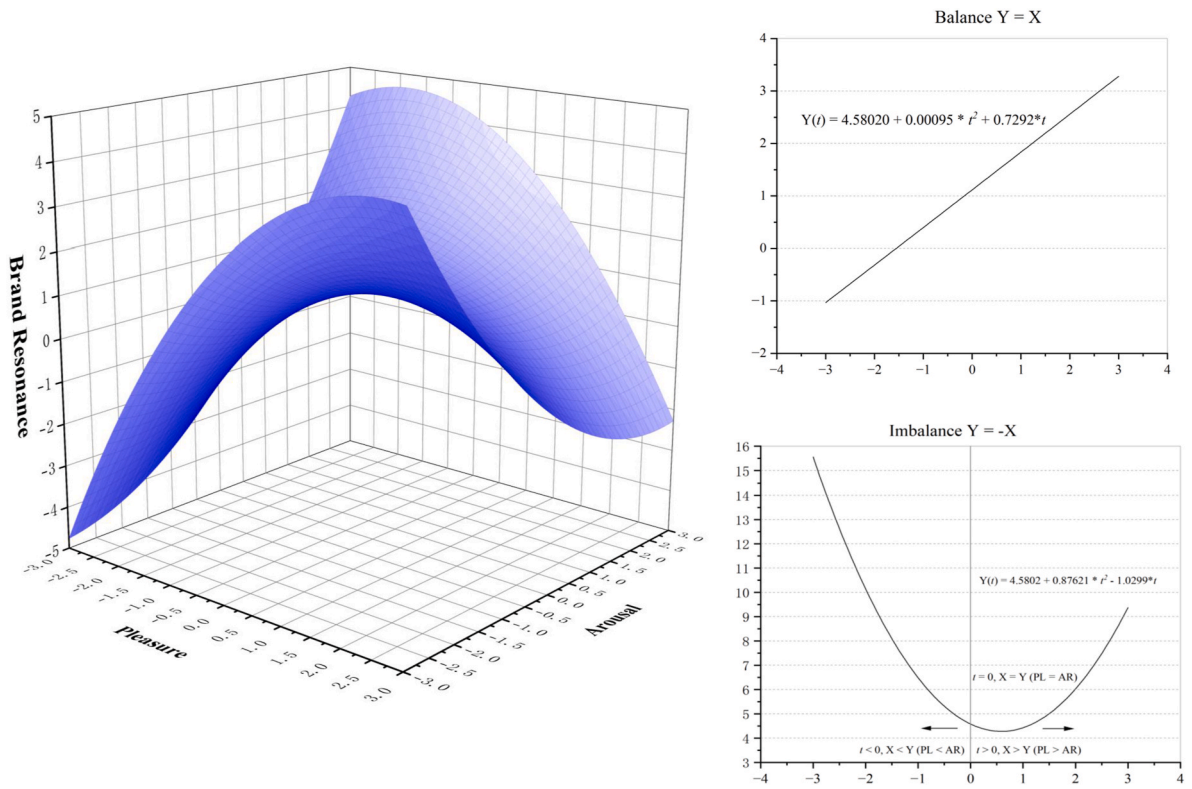


Fig. 3. Response surface (Group 1).

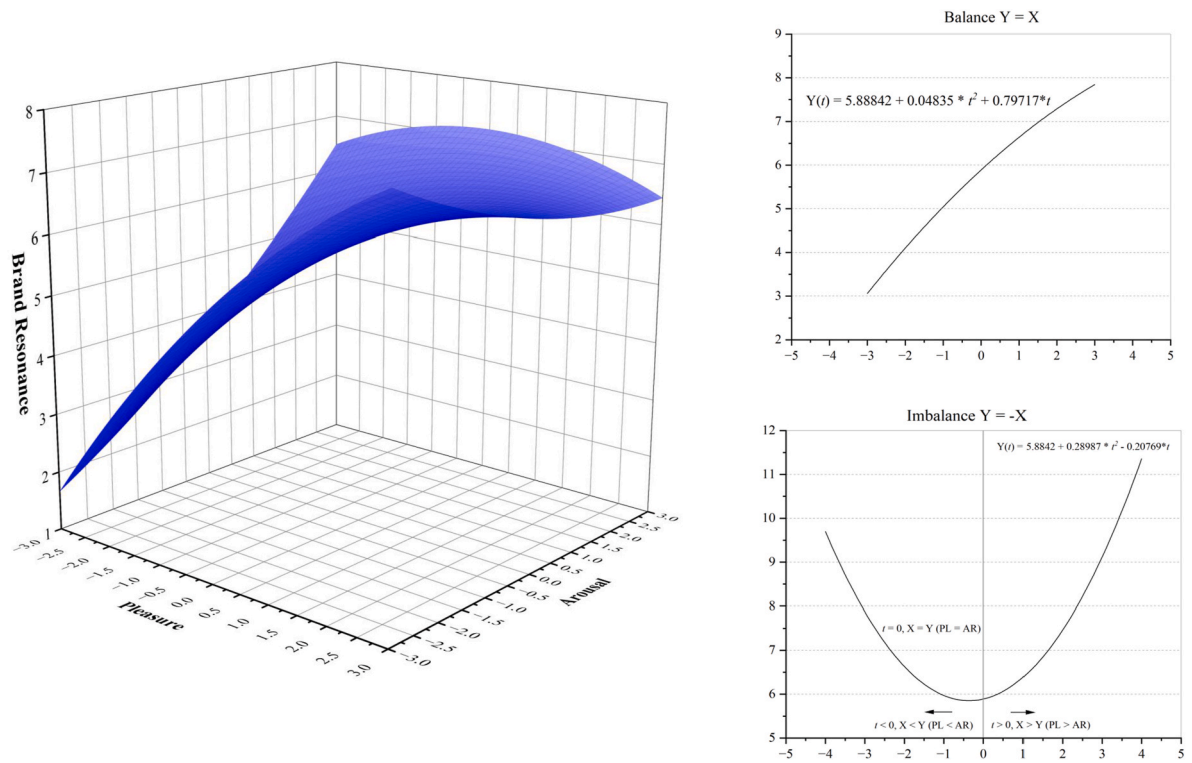


Fig. 4. Response surface (Group 2).

a deep valley at $t \approx -0.6$. High arousal without pleasure evokes anxiety, pushing resonance to its minimum. Moving towards balance by brightening the music or softening the storyline rapidly elevates performance. Participants experience “aroused displeasure”. The RSA captures this

tension as a deep concave segment on the imbalance slice. Hence, the steep ascent occurs when PL and AR become aligned.

Group 4, both cues are subdued. Slopes are small ($a_1 = 0.61$, $a_3 = 0.41$) and curvatures modest ($a_2 = 0.06$, $a_4 = 0.06$); the surface is nearly

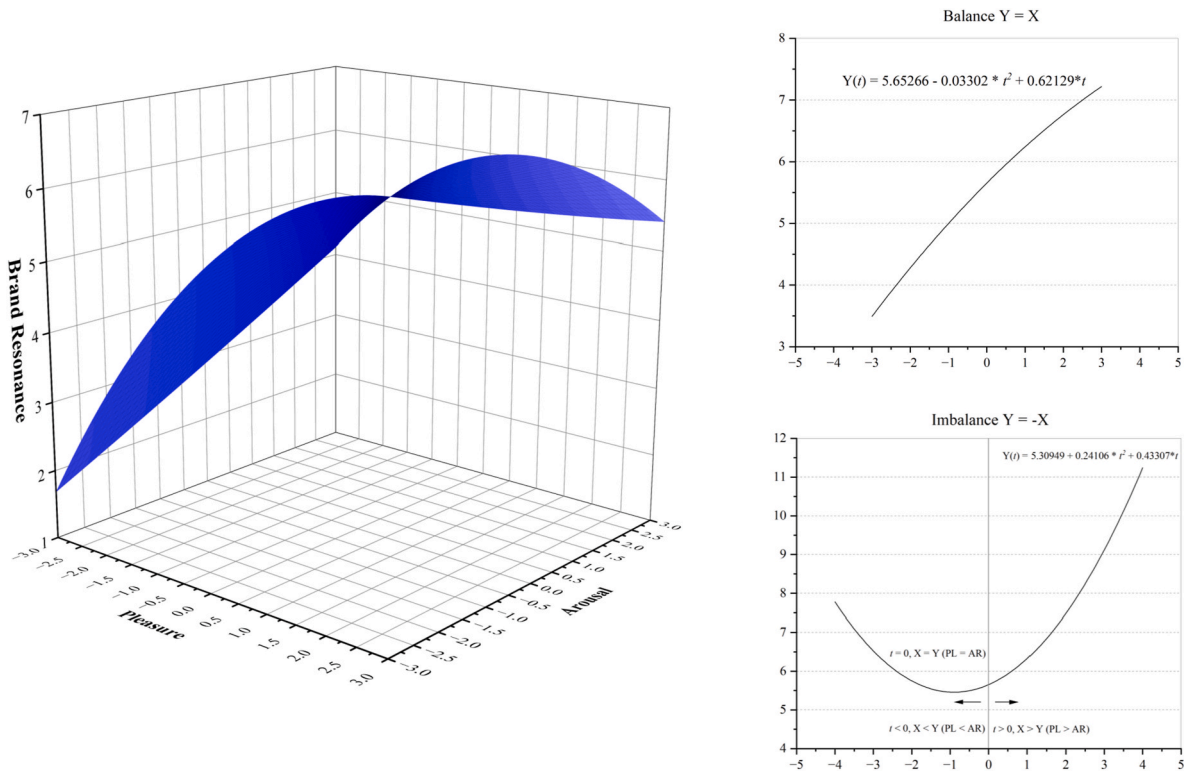


Fig. 5. Response surface (Group 3).

planar. When the baseline effect is inert, neither symmetry nor extra intensity shifts outcomes materially. The brand forfeits emotional leverage in this cell. The prime already fixes respondents in a low-activation, low-valence quadrant, leaving little variance for the soundtrack to exploit. The surface is therefore nearly planar; neither symmetry nor additional intensity shifts destination brand resonance materially. Strategically, the brand forfeits emotional leverage in this condition.

RSA overlays the four balance and imbalance curves. Two regularities emerge: Symmetry advantage under moderate or low intensity (Groups 2–4): balance slopes dominate imbalance slopes, and balance curvature exceeds imbalance curvature; Intensity advantage once both cues are high (Group 1): imbalance penalty disappears; magnitude alone sustains performance.

Managerially, according to Table 1, when pleasure and arousal are both high, added symmetry yields little extra gain, and overall intensity sustains DBR. When pleasure is high but arousal is low, or when arousal is high but pleasure is low, imbalance reduces DBR, and moving toward balance helps. When both cues are low, the surface is flat, and DBR is hard to shift. High-context edits showed a slight advantage on the appraisal check; keeping cultural cues while adjusting intensity toward balance is advisable.

4.3.3. Discussion

Study 1 shows that hedonic tone and activation shape destination brand resonance (DBR) through separate paths. Pleasure is the more stable path: across the four effect cells, a one-unit increase in pleasure is associated with roughly a half-unit increase in DBR, even when arousal is controlled. This pattern also appears when both cues are low, indicating that a positive tone is sufficient to support symbolic attachment when activation is muted. Arousal behaves differently. It adds influence only when pleasure is already high or when the scene and the soundtrack together elevate the affective baseline; when pleasure is low, the arousal coefficient is near zero. The response-surface plots are consistent with this reading. Along the imbalance slice, DBR falls when arousal exceeds pleasure and rises when pleasure exceeds arousal, while the surface is steepest on the pleasure axis and comparatively flat on the arousal axis unless pleasure is high. Thus, in this symbolic setting, the two dimensions combine additively rather than interactively, and arousal contributes conditionally.

Two covariates warrant brief notes. Informal musical training strengthened the arousal effect in the high-pleasure/high-arousal cell, suggesting greater sensitivity to high-energy scores. The occupational category reduced the pleasure effect in the high-pleasure/low-arousal cell, suggesting a more guarded response in that group. These patterns

Table 1
Cross-group synthesis.

Dyadic affect profile	Typical cell	RSA summary	Practical guidance	Cultural note
High pleasure – High arousal	G1	Intensity dominates; balance and imbalance slices both rise; interaction is not needed.	Keep high energy; minor symmetry changes add little.	Retain salient cultural cues; focus on maintaining intensity.
High pleasure – Low arousal	G2	Balance path climbs; imbalance slice shows a U-shape near the stimulus point.	Raise activation (tempo, dynamics) or soften pleasure cues to meet the scene tone.	High-context cues are effective when activation is lifted to match pleasure.
Low pleasure – High arousal	G3	Imbalance curvature positive; valley when arousal rises without pleasure.	Add positive valence in music or ease content (e.g. narrative) tension to approach balance.	Use culturally familiar bright timbres or motifs to supply valence.
Low pleasure – Low arousal	G4	Near-planar surface; small slopes and curvatures.	Redesign both scene and soundtrack; neither cue moves DBR reliably.	Increase embedded cultural symbolism only together with stronger affect.

indicate that listener attributes can bound affect-DBR links.

Fig. 2, which averages cells by nominal pleasure and arousal, aligns with the regression: the pleasure contrast is sizable, whereas the arousal contrast is small. Overall, Study 1 indicates that pleasure is the primary affective driver of DBR, arousal is secondary and contingent, and their effects are additive. The study does not address why people facing the same scene and soundtrack report different levels of affect and DBR. Study 2 introduces a MAP that induces a temporary receptivity state and tests whether that state elevates pleasure and arousal, raises DBR, and whether including MAP as a direct predictor attenuates the CMA paths.

4.4. Study 2

Receptivity was examined as an antecedent to dyadic emotional valence under the same content-music package. The study manipulated a receptivity state, MAP versus Neutral, to test whether this state causally raises pleasure and arousal and, in turn, DBR.

A between-subjects design assigned participants to MAP or Neutral across the four emotion-context cells used in Study 1. This permits a test of the causal effects of receptivity while holding the affective context constant. The prime adjusted openness to musical-symbolic content; participants then viewed the same destination videos as in Study 1. Covariates (age, career, music background, brand preference) were retained only when statistically justified (see Fig. 6).

4.4.1. Method and results

Participants (N = 606; per-cell Ns = 145-159) were randomly assigned to one of four cells: high pleasure/high arousal, high pleasure/low arousal, low pleasure/high arousal, or low pleasure/low arousal. After exposure, respondents reported pleasure and arousal on seven-point semantic differentials from the circumplex model and then completed DBR. The primary independent variable was the MAP versus Neutral manipulation. For regression plots (Figs. 7 and 8), receptivity was represented as a mean-centered continuous score.

Analyses were run separately by cell with pleasure and arousal as

dependent variables and MAP (1 = MAP, 0 = Neutral) as the predictor; covariates such as age, education, career status, music training background, and brand preference were included where significant, ensuring that observed effects were attributable to MR rather than demographic confounds.

In the high-pleasure/high-arousal cell, MAP (1) versus Neutral (0) produced large lifts on both dimensions. Relative to Neutral, MAP increased pleasure by about 1.49 points (B = 0.744, p < .001) and arousal by about 1.25 points (B = 0.659, p < .001) after controls. The models explained substantial variance (R² = 0.342 for pleasure; R² = 0.284 for arousal), showing that when cues already carry strong positive tone and activation, the MAP state materially amplifies affective appraisal.

In the high-pleasure/low-arousal cell, MAP > Neutral again raised both outcomes, with smaller magnitudes: pleasure rose by about 0.90 points (B = 0.431, p < .001) and arousal by about 0.68 points (B = 0.297, p = .008). Education, music-training background, and career status accounted for additional variance, but MAP remained the key predictor (R² = 0.322 for both outcomes). Thus, MAP still elevates the dyad when tone is positive but activation is restrained.

In the low-pleasure/high-arousal cell, MAP > Neutral yielded increases of about 1.08 points on pleasure (B = 0.493, p < .001) and 0.98 points on arousal (B = 0.423, p = .001). Age and music background shifted levels but did not alter the MAP slope, indicating a robust effect where activation is high but valence is limited.

In the low-pleasure/low-arousal cell, MAP > Neutral produced the most asymmetric response: arousal rose more than pleasure. MAP increased pleasure by about 0.90 points (B = 0.416, p = .001) and arousal by about 1.13 points (B = 0.596, p < .001). The arousal model fit was strongest here (R² = 0.366), consistent with MAP supplying activation when the stimulus is affect-sparse.

MAP versus Neutral reliably elevates dyadic emotional valence under identical audiovisual inputs. The lift is largest when both dimensions are already high; when both are low, MAP shifts arousal more than pleasure. Framed by Cognitive Metaphor Theory, the MAP state

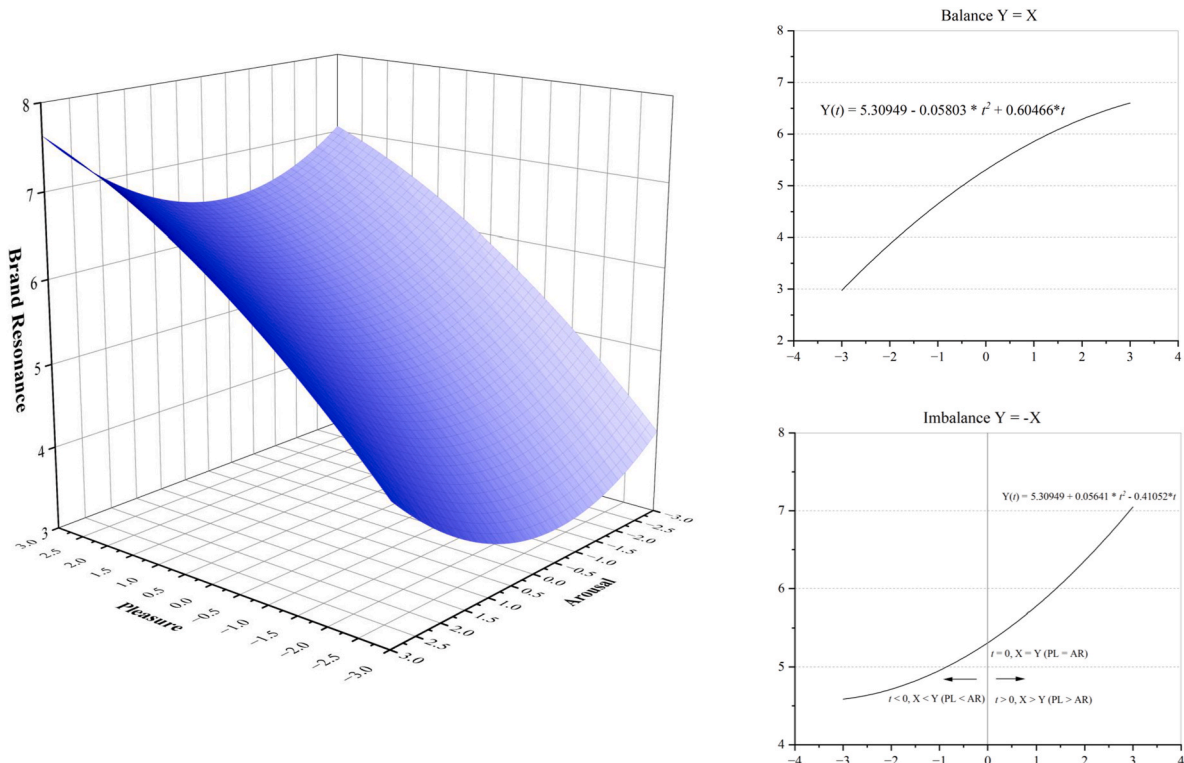


Fig. 6. Response surface (Group 4).

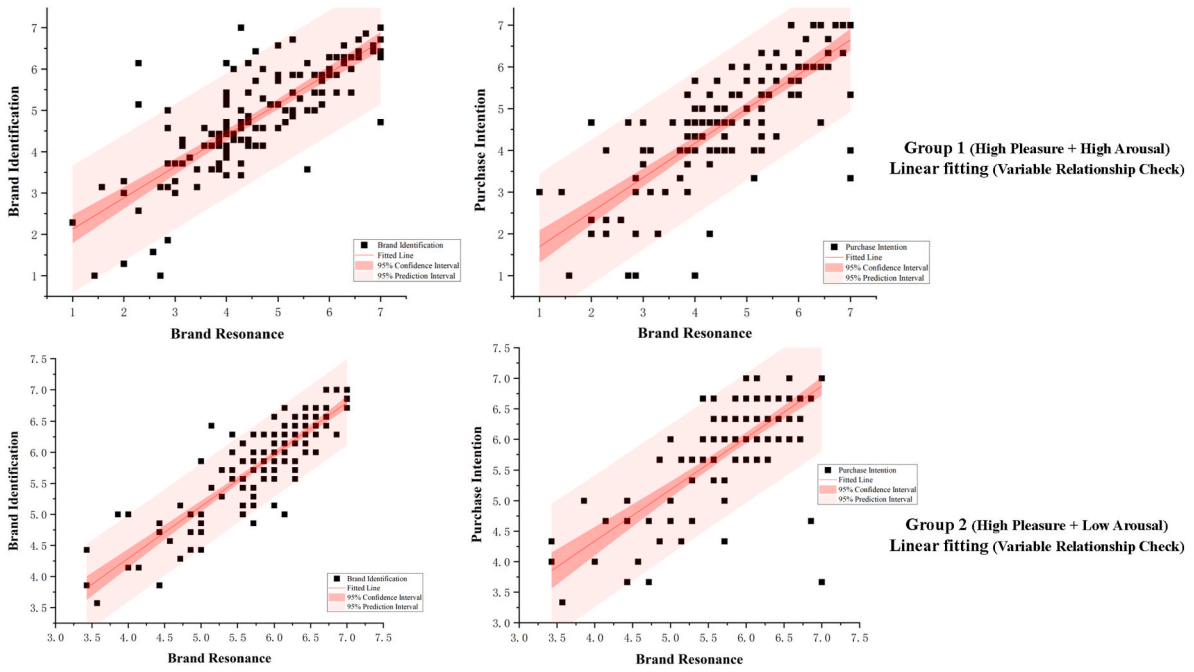


Fig. 7. Multiple regression (MR→PL/AR, Group 1&2).

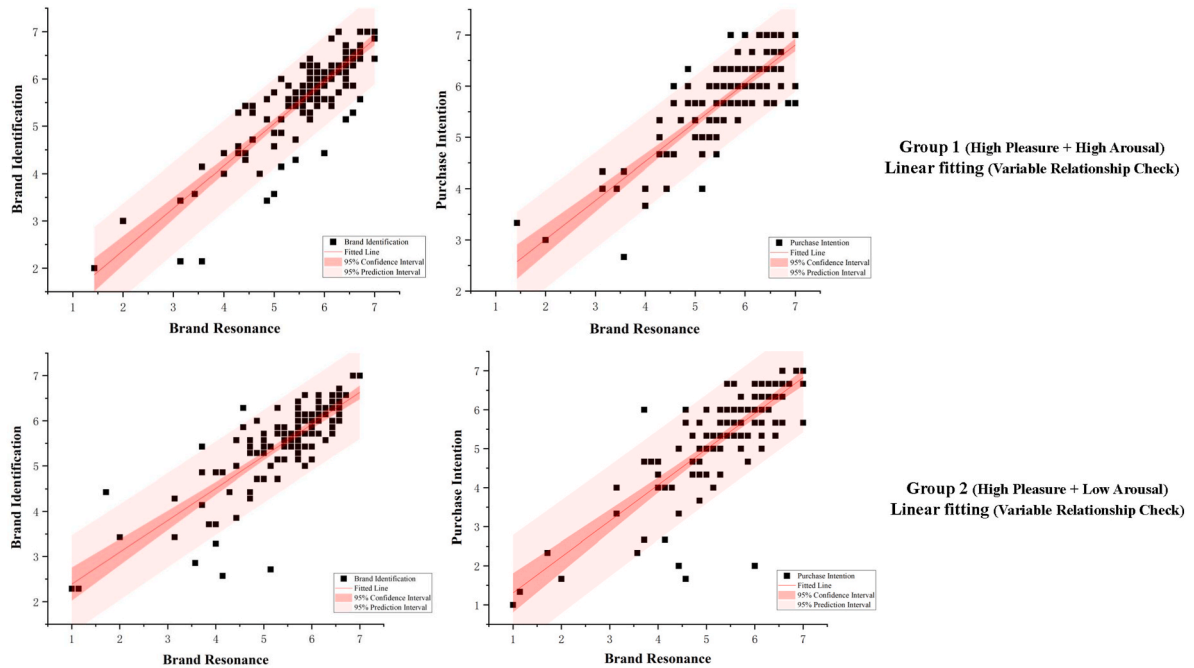


Fig. 8. Multiple regression (MR→PL/AR, Group 3&4).

heightens access to culturally patterned mappings in the music–visual pairings, which strengthens affective appraisal and sets the stage for higher destination brand resonance in the subsequent analyses.

4.4.2. Discussion

Study 2 establishes that a meaning-access prime causally increases dyadic emotional valence under identical audiovisual stimuli. Across all four cells, MAP raised pleasure and arousal relative to Neutral after accounting for demographic and background covariates. The size of the MAP effect varied with the affective context: effects were largest when both dimensions were high, and in low-affect contexts MAP had a

stronger influence on arousal than on pleasure. Interpreted through CMT, the prime enhanced access to culturally patterned mappings, producing stronger affective appraisal and, consequently, higher DBR in the subsequent tests. Study 3 tests the full causal pathway from music receptivity to brand outcomes, examining whether pleasure and arousal mediate its effect on brand resonance across emotional conditions.

4.5. Study 3

This study implements the complete model in a single experiment to verify that the mechanism operates when downstream variables are

included. DBR is modeled as the proximal bond that links appraisal to downstream outcomes. Two outcomes are assessed: DBI and DCI. CMT motivates the pathway in which MAP increases access to culturally patterned mappings in music-image pairings, which elevates pleasure and arousal and consolidates into DBR; DBR then carries forward to DBI and DCI. To probe boundary conditions, the design retains the effect cells from Study 1 (high/low pleasure × high/low arousal) and the cultural-embedding edit (high-context vs. low-context). DBR is treated as a latent continuous construct in the structural models, which provide the formal tests of all hypotheses. For illustration only, a median split of DBR (“Resonant” vs. “Emergent”) is used to display group means on DBI and DCI; these descriptive contrasts are consistent with, but not a substitute for, the continuous-model results. Two structural specifications are compared: a full-mediation model estimating MAP → pleasure and MAP → arousal, pleasure and arousal → DBR, and DBR → {DBI, DCI}; and an alternative that also includes the direct MAP → DBR path.

A partial-mediation model adds a direct MAP → DBR path to capture a non-affective route consistent with CMA-based meaning access. Model comparison focuses on overall fit and information criteria. The seven hypotheses tested in this study and in Studies 1–2 are summarized in Supplementary_D, Appendix K (“Hypotheses, tests, and outcomes”), where each hypothesis is matched to the study, model, and observed result to maintain systematic reporting without interrupting the results flow. Because Study 3 uses Bayesian SEM, a prior-sensitivity check for all four groups is also reported in Supplementary_C (Appendices E–H), and the fit and key paths remained stable across priors.

4.5.1. Method

Design and procedure. Study 3 retained the experimental structure established in Study 1, using the same four affect conditions and TikTok-sourced video stimuli. Participants (N = 620) were randomly assigned to affective cells and received either MAP or Neutral instructions prior to viewing. All variables were measured in a single post-exposure session.

Measures. DBR was assessed using the multi-item scale from Studies 1–2 ($\alpha > .80$). Destination brand identification (DBI) employed a four-item scale capturing perceived identity overlap with the destination ($\alpha > .80$). DCI was measured using three items reflecting willingness to spend within 12 months if visiting the location ($\alpha > .80$).

Analytic strategy. Bayesian structural equation modeling (BSEM)

was conducted using Mplus 8.3, chosen for its robustness in estimating complex models with moderate sample sizes and its use of weakly informative priors. Both full-mediation and partial-mediation models were estimated on the pooled sample, with affect condition and cultural embedding included as exogenous variables. Model evaluation was based on posterior predictive p-values (PPP), 95 % credibility intervals, and Bayesian information criteria (BIC). For descriptive purposes, DBR was also dichotomized into Resonant versus Emergent states to compare mean differences on DBI and DCI. All continuous predictors were mean-centered prior to analysis.

4.5.2. Results

Across the four affect cells, Bayesian SEM with the direct receptivity → DBR path fit better than the model that routed receptivity only through pleasure and arousal. The overall effect can be seen in Figs. 9–12. Posterior predictive p-values (PPP) rose when the direct path was added, Group 1: 0.006 → 0.073; Group 2: 0.047 → 0.116; Group 3: 0.001 → 0.249; Group 4: 0.001 → 0.347, indicating closer correspondence between the model-implied and observed covariance structures, with the largest gains in the low-pleasure cells (Groups 3–4). Deviance/Bayesian information criteria also declined in every group, confirming that the model with the direct effect captured variance not explained by the mediators.

The direct MR → DBR effect was statistically different from zero in all four groups, with 95 % credibility intervals excluding zero: $\beta = 0.239$ in Group 1, 0.149 in Group 2, 0.352 in Group 3, and 0.294 in Group 4. The strongest coefficient in Group 3 (high arousal/low pleasure) indicates that the receptivity state induced by MAP (vs. Neutral) supports formation of DBR even when pleasant tone is limited, consistent with CMT in which access to culturally patterned mappings can consolidate meaning beyond immediate hedonic tone.

Mediation shares varied with the effect setting. In Group 1, arousal accounted for 38.2 % of the total MR effect on DBR, whereas the pleasure pathway was not credible; intense stimuli appear to channel the receptivity effect mainly through activation. In Group 2 (high pleasure/low arousal), both routes contributed, pleasure 42.0 % and arousal 20.7 %, yielding a more balanced dyad under calming but pleasant content. In the low-pleasure cells (Groups 3–4), the pleasure route was modest (~30 %) and the arousal route was negligible; the direct MR → DBR path

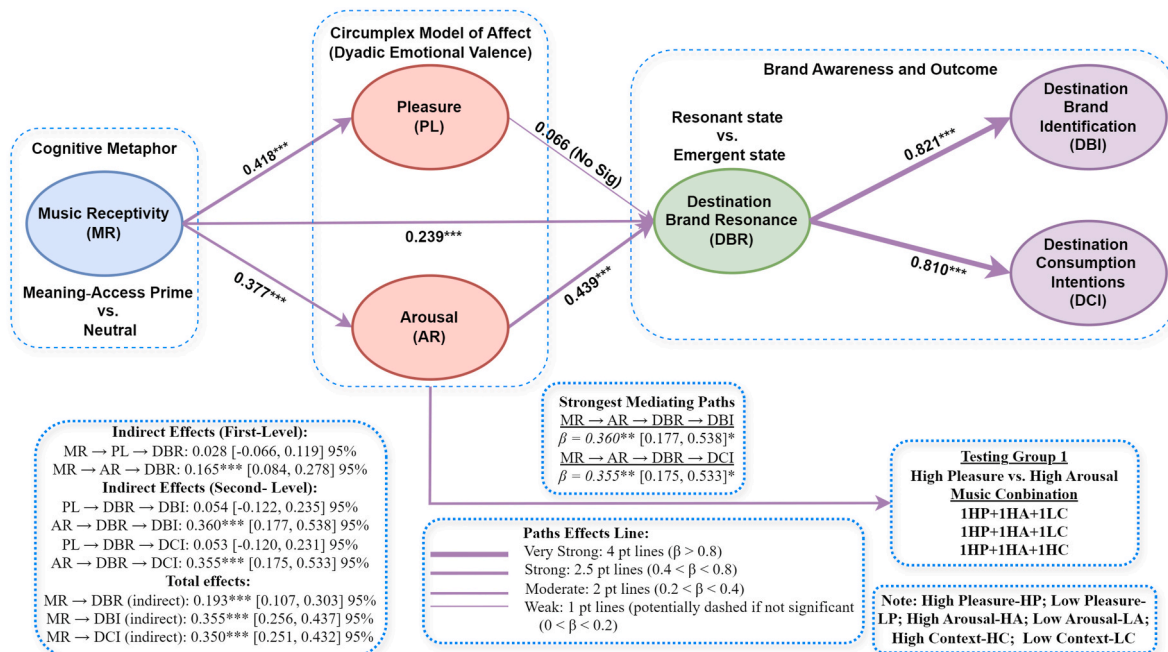


Fig. 9. Path result - Group 1.

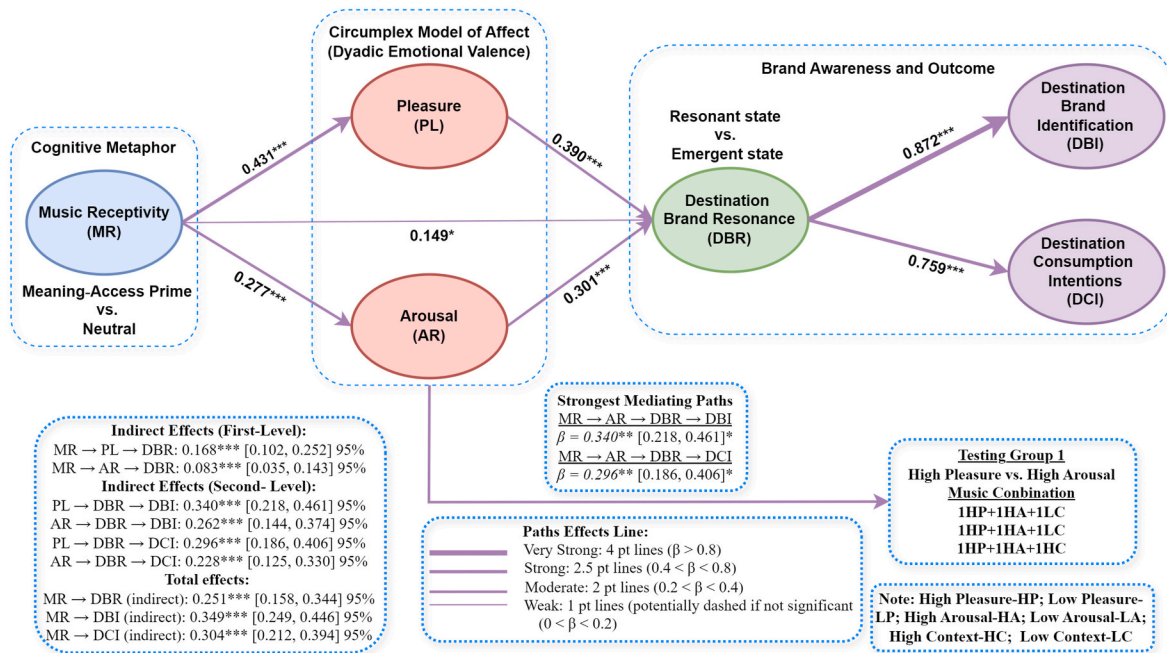


Fig. 10. Path result - Group 2.

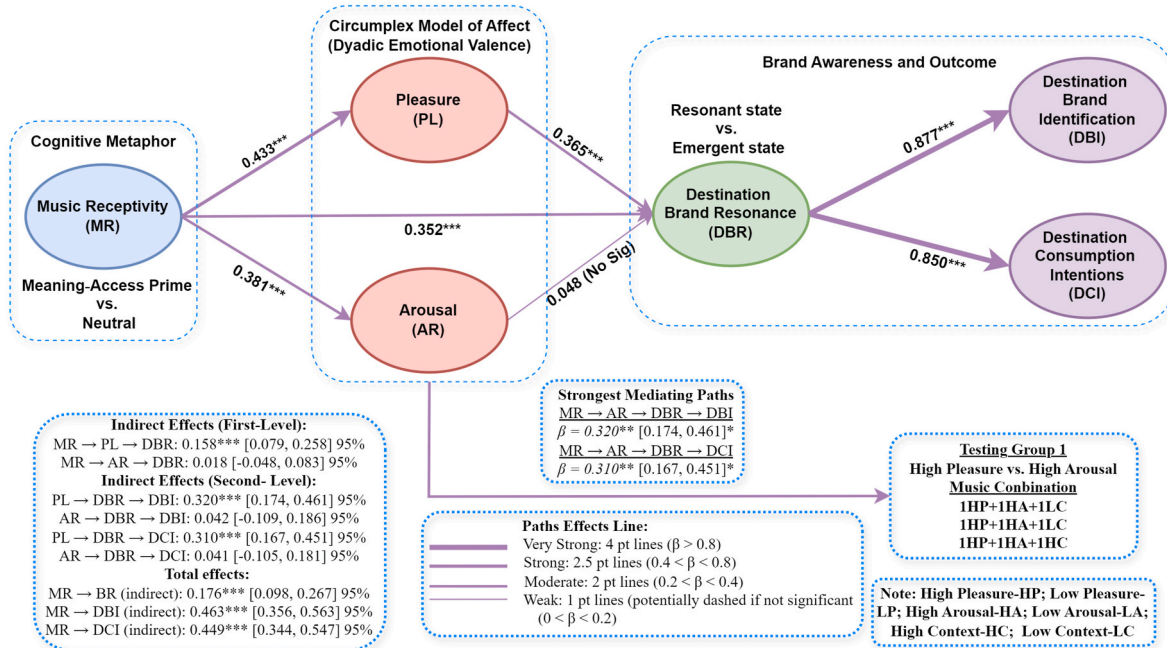


Fig. 11. Path result - Group 3.

carried the majority of the effect (66.7 % and 67.9 %, respectively). Without the direct path, the models in these cells produced inconsistent inferences (e.g., pleasure appearing dominant despite weak pleasant tone), underscoring the need to retain the direct route.

Downstream links were stable. DBR → DBI paths ranged $\beta = 0.813-0.877$, and DBR → DCI ranged $\beta = 0.759-0.850$, all with credibility intervals excluding zero. Variance explained in DBR increased when the direct path was included, with the largest gain in Group 3 ($R^2: 0.316 \rightarrow 0.412, \Delta R^2 = 0.096$). Group 2 showed the highest DBR R^2 (0.457) but the smallest improvement ($\Delta R^2 = 0.017$), consistent with stronger mediation under pleasant, lower-activation content. The inferences above are based on the continuous DBR scores. A simple

median split, reported as a supplementary illustration, shows a pattern that is in line with the continuous results: participants in the higher-DBR segment ("Resonant") reported higher identification ($M_{Res} = 5.41$ vs. $M_{Emer} = 4.62, t(618) = 8.94, p < .001$) and higher consumption intention ($M_{Res} = 5.18$ vs. $M_{Emer} = 4.36, t(618) = 8.51, p < .001$). This split is descriptive and does not replace the continuous-model estimates.

The evidence shows that receptivity exerts both indirect and direct influences on DBR, and that the balance between the two depends on the effect profile of the stimulus. Models that omit the direct MR → DBR path are misspecified, especially under low-pleasure conditions, where pleasure and arousal do not account for most of the movement in resonance. The complete model should therefore be retained for theory

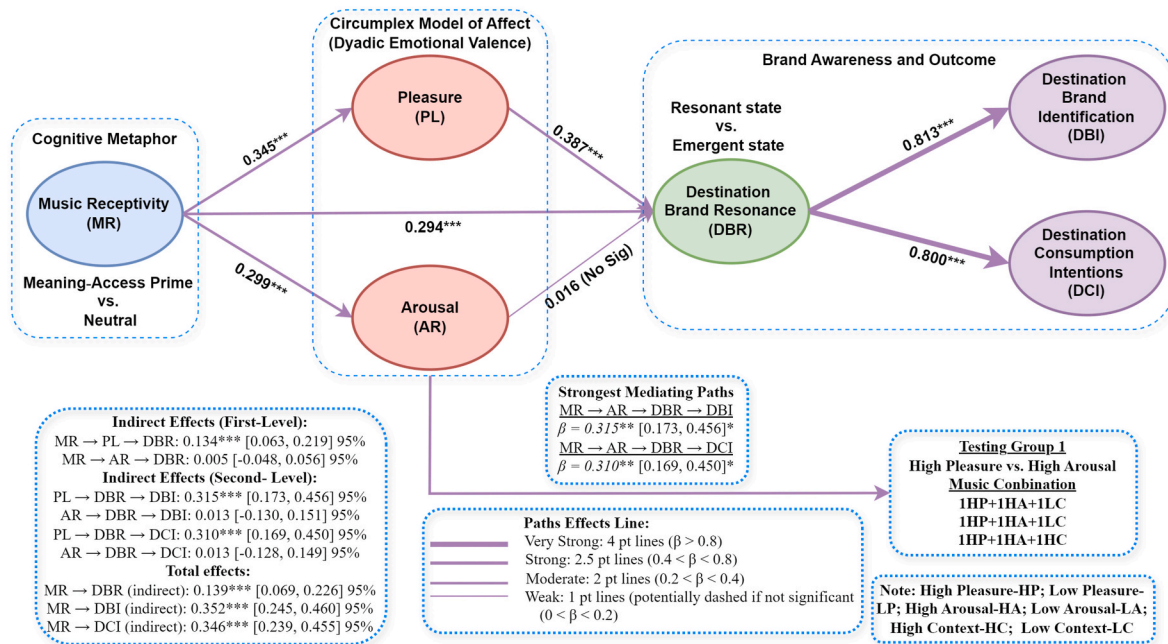


Fig. 12. Path result - Group 4.

and for application.

4.5.3. Discussion

Study 3 is designed to evaluate the complete mechanism in one experiment. The key questions are whether (a) the affective route from receptivity to DBR via pleasure and arousal remains operative when DBI and DCI are added, (b) a direct receptivity → DBR path improves the model, consistent with CMA’s claim that meaning access can consolidate into a bond independent of momentary affect, and (c) the pathway is stable across affect cells and cultural-embedding levels. Interpreting DBR both continuously and as Resonant versus Emergent allows a clear reading of downstream consequences: if DBR functions as the proximal bridge, the Resonant state should display higher DBI and DCI than the Emergent state, while the structural model attributes those differences to the earlier appraisal steps.

5. General discussion

Across three experiments, the findings reveal a consistent pattern: pleasure is the primary emotional driver of destination brand resonance, arousal plays a secondary and conditional role, and a meaning-access state, induced by the receptivity prime, enhances both emotional pathways while also contributing directly to resonance.

Study 1 established the baseline effect. Pleasure showed a stable positive link to destination brand resonance in every effect cell, including settings with low activation. Arousal helped only when pleasure was high. Interaction terms were not supported. Response-surface analyses showed symmetry benefits at moderate or low intensities, but once both dimensions were high, overall intensity sustained resonance, and adding symmetry yielded little extra gain. High-context edits delivered a small advantage on the appraisal check, consistent with easier access to culturally familiar cues.

Study 2 showed that the same clips produce a stronger effect when viewers are placed in a meaning-access state. MAP versus Neutral raised pleasure and arousal in all four cells. The lift was largest when both dimensions were already high; in effect, sparse settings MAP shifted arousal more than pleasure, indicating that receptivity can supply activation when the stimulus is quiet. These results confirm that receptivity acts upstream of the dyadic emotions rather than merely

tracking them.

Study 3 tested the full model, including identification and consumption intention. Bayesian models favored the specification that includes a direct path from receptivity to resonance. The gain in fit was largest in the low-pleasure cells, where pleasure and arousal alone did not account for most of the change in resonance. Mediation shares varied with the effect profile: arousal carried most of the effect when both dimensions were high; pleasure and arousal both contributed under high pleasure/low arousal; in low-pleasure settings, the direct path dominated with a modest pleasure share. Resonance then showed strong, stable paths to identification and consumption intention. A descriptive split of resonance into Resonant versus Emergent states produced clear mean gaps on both downstream outcomes, matching the structural paths.

These results outline a clear mechanism. When culturally embedded sound-image pairings are processed in a meaning-access state, they enhance emotional response and, in low-pleasure contexts, activate a cognitive route to brand resonance. That resonance, in turn, reliably predicts downstream behavioral intentions.

5.1. Theoretical contributions

This research presents a concise account of how culturally embedded music-image pairings shape destination responses by integrating CMT with the dyadic effect axes of pleasure and arousal. Across three experiments, pleasure emerges as the dominant effect route to destination brand resonance, while arousal contributes only under a favorable tone. Interaction terms are not supported, and response-surface analyses reveal additive effects with two regimes: at moderate or low intensity, the alignment of pleasure and arousal improves outcomes; once both dimensions are high, overall intensity sustains resonance, and added symmetry offers little extra gain. Prior work shows that audio parameters such as tempo and vocal acoustics systematically shift affect in travel ads (Barnes et al., 2023), that short-form destination videos influence travel intention through framing and affect intensity (Gan et al., 2023), that music effects in tourism settings depend on presence, tempo, congruence, and liking (Trompeta et al., 2022), that multimodal metaphor frames emotional appraisal (O’Dowd, 2024), and that culture shapes how consumers parse cues along the journey (Shavitt and Barnes,

2020); these streams have typically offered valuable but distinct perspectives on emotion, metaphor, and culture; often emphasizing valence-based effects or assuming multiplicative relationships, and only rarely addressing a receptivity state that enables both indirect (affect-based) and direct pathways to bonding in creator-made short videos. The present studies extend this work by refining CMA for short, nonverbal destination communications, clarifying when affective activation enhances outcomes, and demonstrating a necessary direct receptivity-to-resonance path with context-dependent mediation.

On this basis, the present findings contribute a culturally grounded extension: high-context orientation functions as a compensatory decoding mechanism when hedonic tone is weak. Under low-pleasure conditions, viewers exposed to culturally dense symbolic cues still achieved meaningful levels of resonance, especially when accompanied by moderate or high arousal. This pattern suggests that, when culturally primed, receptivity activates a direct semantic integration route independent of affect. High-context audiences, accustomed to metaphor-rich communication, can extract brand meaning even in sparse affective environments, provided symbolic clarity is preserved. This interpretation positions cultural context not as background noise but as a structural amplifier of meaning, offering a second pathway to brand bonding that complements emotion-based routes.

MAP induction reliably increases pleasure and arousal in response to identical stimuli, confirming that receptivity lies upstream of the dyadic emotions. Bayesian structural models also require a direct receptivity → resonance path, with the largest gains in fit under low-pleasure content. This pattern suggests that when a pleasant tone is weak, access to culturally patterned mappings can still consolidate into a bond, even without a strong hedonic lift. In mediation terms, receptivity operates as a competitive partial mediator: indirect affect routes remain, but a direct path is needed to capture meaning consolidation that is not reducible to momentary affect.

CMT is extended by demonstrating that metaphorical structure in music can guide destination brand meaning through both affective appraisal and a direct meaning-access route, whose weight shifts with the effect context. The evidence supports two coordinated functions. First, metaphorically encoded cues in sound and image move the dyadic emotions, with pleasure as the reliable conduit. Second, receptivity heightens the accessibility of those cues and can form resonance directly when a pleasant tone is limited. This dual-route view clarifies why models that force all influence through emotion misattribute variance, especially in low-pleasure settings.

High-context edits show a slight advantage in perceived metaphorical richness and imagery, and the MAP induction enhances access to symbolic cues across contexts. The results indicate partial orthogonality between cultural structure and affect: the two routes can reinforce each other when tone and activation are high. However, cultural structure can also compensate when a pleasant tone is weak or activation dominates. This accounts for the observed compensation in the low-pleasure cells and explains why direct receptivity effects are strongest.

Resonance ultimately drives destination brand identification and consumption intention, positioning short-form, creator-driven video content within a broader framework of destination brand equity. The split of resonance into Resonant and Emergent states adds a useful state view to the theory. The state contrast reveals a threshold-like shift in outcomes: once appraisal and meaning access are sufficient to move viewers into the Resonant state, identification and intention rise sharply relative to the Emergent state.

5.2. Practical implications

Marketers can treat pleasure, arousal, and cultural embeddedness as separate levers that must be co-designed. Pleasure carries destination brand resonance most reliably, so every edit should first secure a pleasant tone that fits the place. Arousal should then be adjusted or restrained to match the tone and task. When the tone is tranquil but

positive, raising activation through tempo, rhythmic density, and dynamic contrast can improve results. When activation is high but the tone is dark, adding pleasant musical cues, major-mode passages, consonant harmony, brighter timbres, and upward melodic motion restores balance. When both tone and activation are low, neither cue moves resonance enough to matter; the scene and the soundtrack need to be rebuilt rather than tuned.

The results show that pleasure, arousal, and cultural cues can be adjusted separately, allowing a destination marketer to design different versions of the same short video for various goals. A calm, high-end or wellness-type scene should combine a pleasant musical tone with slow visual pacing and avoid high-energy percussion, because high arousal without high pleasure weakens resonance in this context. A festival or city-energy scene should do the opposite: fast cuts and high-tempo music are acceptable because high pleasure and high arousal move together. For historical or high-context scenes, adding recognizable cultural sounds or visual motifs from the target audience helps move arousal into interpretation rather than tension. Before releasing a clip, a simple pre-flight checklist can be used: confirm that the soundtrack is pleasant enough for the brand, ensure that the musical energy matches the visual pace and goal, and verify that culture-specific cues are clear for the audience that will view it first. On platforms, this can be tested using a two-cell A/B design (same visuals, low-arousal music vs. high-arousal music) and compared based on watch-through rates, likes, shares, and click-through rates to destination pages. This turns the theoretical finding “alignment of pleasure and arousal improves resonance” into a concrete content-testing routine.

Cultural embeddedness should be added on top of the tone–energy plan, not used instead of it. High-context cues, traditional instruments, pentatonic figures, emblematic visuals, and ritual or heritage motifs make symbolic meaning easier to access and are most helpful when a pleasant tone is weak or activation dominates. Low-context cues, clear and explicit imagery, and straightforward musical phrasing work well when the tone is pleasant and calm. This design logic explains why Group 3 benefited from adding pleasant cues to high activation and why Group 2 benefited from adding activation to a pleasant, calm state. The aim here is to match the effect profile to what the audience can decode quickly, not to force one recipe on every market.

In high-context markets, implicit cultural symbols and musical metaphors are not decorative; they function as primary carriers of meaning. Marketers should approach these cues as strategic tools rather than aesthetic flourishes. When pleasure is low but activation is high, as seen in the G3 response-surface group, campaigns benefit from emphasizing familiar instruments, pentatonic melodic figures, and heritage-laden visuals. These high-context signals help restore emotional balance and prevent resonance collapse. In contrast, low-context markets require more explicit verbal anchors to ensure that symbolic elements are understood as intended. This distinction highlights the importance of culturally sensitive creative logic: the symbolic density must align with the audience’s decoding fluency. Future research should replicate these effects in low-context samples to more precisely define the boundary between implicit and explicit messaging strategies.

Receptivity can be influenced before or during exposure. Short prompts, captions, or on-screen copy that invite viewers to “listen for the city’s rhythm,” “hear the river lines in the strings,” or “notice how drum patterns follow the crowd’s pace” act like a meaning-access prime. They orient attention to symbolic relations in the clip and raise both pleasure and arousal under the same stimulus. This is a low-cost tool for platforms built on creator content, where marketers curate rather than produce every asset. When space allows, thumbnails and first frames should surface the key cultural cue so that viewers enter the clip already tuned to its meaning.

Destination brand resonance can be monitored as a continuous score and a state. The continuous score is sensitive to creative tweaks and useful for modeling. The state split, Resonant versus Emergent, helps

with campaign control. A useful target is to move a larger share of the audience across the threshold into the resonant state, because identification and spending plans rise sharply once that state is reached. In practice, this means testing whether edits that raise a pleasant tone or add meaning-access prompts increase the share of Resonant viewers, and then allocating spend to the combinations that deliver the largest shift.

Implementation should follow a simple test–learn loop. For each destination, build paired edits that differ on one lever at a time: pleasant tone, activation, or cultural embeddedness. Pretest pleasure and arousal to confirm the intended shift, confirm that the appraisal items indicate symbolic meaning, and then track DBR and downstream metrics. Where MAP-style prompts enhance affect or DBR, retain them; where they do not, reallocate effort to tone and activation. Short videos allow rapid iteration, and the present results provide a rule set to guide those iterations toward stronger resonance, identification, and planned spending.

6. Limitations and future research

The experiments relied on random assignment and controlled clips, so the effects of the affective environment in Study 1 and of MAP versus Neutral in Studies 2–3 on pleasure, arousal, and destination brand resonance can be read as causal within the lab-style setting. A main limitation is the time scale: all measures were taken right after exposure. This is standard for stimulus–response work, but it means the study reports immediate states and intentions rather than later platform behavior. Future work should add follow-up waves or link survey IDs to platform logs to test whether the lift from MAP and the direct receptivity path stay, grow, or fade over repeated views.

The appraisal check in this study relied on participants' immediate reports that the clip carried symbolic or culturally meaningful cues. Future studies should therefore incorporate an external layer of coding or machine-based descriptors (such as tempo, mode, timbral brightness, motif frequency, and density of destination cues) to triangulate the manipulation without altering the viewer's task, and implement comprehensive acoustic and visual coding protocols. The analysis of musical parameters should include tempo measurement in beats per minute, modal classification of tonal systems, spectral analysis for timbral brightness and harmonic complexity, and rhythmic density through onset detection. Systematic coding of visual elements requires measuring cultural emblem frequency, color palette distributions, shot duration analysis for editing pace, and the proportion of traditional versus contemporary markers. Such protocols would enable more rigorous manipulation verification, facilitate meta-analytic synthesis, and permit investigation of whether specific feature combinations differentially activate metaphorical processing. A follow-up study can test whether it reveals an inverted-U pattern in arousal, or adding short music-emotion sets can check whether particular emotions contribute beyond the dyad. Mixture or threshold models can then be used to formalize the observed Resonant versus Emergent states and to estimate how much MAP increases the probability of moving into the Resonant state.

The model centers on the dyadic effect axes—pleasure and arousal—with CMT providing the content logic. This choice keeps the mechanism compact and allows for tracing how meaning-access prompts and audio–visual edits move viewers along two main emotional dimensions; however, it does not test whether finer labels, such as awe, nostalgia, or tenderness, add incremental predictive value. A calm, high-end, or wellness scene should pair a pleasant musical tone with slow visual pacing and avoid high-energy percussion, because high arousal without high pleasure weakens resonance in that context. A festival or city-energy scene can do the opposite: fast cuts and high-tempo music are appropriate because high pleasure and high arousal move together there. For historical or other high-context scenes, adding recognizable cultural sounds or visual motifs for the target audience helps channel arousal into interpretation rather than tension. Before

release, the clip can be checked on three key points: the soundtrack is pleasant enough for the brand, the musical energy aligns with the visual pace and goal, and culture-specific cues are clear to the intended audience. On platforms, this can be verified with a two-cell A/B test that keeps visuals constant and varies musical arousal (low vs. high), then compares watch-through, likes, shares, and click-through rates to destination pages.

Cultural embeddedness in the present studies was manipulated through high-versus low-context edits rooted in Chinese motifs, and the small but consistent advantage for the high-context version on the appraisal check suggests that Chinese viewers accessed meaning more easily in that condition. This pattern is consistent with China being a high-context culture, where audiences are accustomed to deriving meaning from nonverbal cues, such as style, symbols, and sound. However, it is not yet clear whether the same level of strength of the direct cognitive path to DBR would hold in low-context markets such as the United States, Germany, or the Nordic countries. The next step is to repeat the design with two matched samples, one from a high-context culture (e.g., China) and one from a low-context culture (e.g., the U. S.), to show both groups culturally adjusted clips and to estimate the model in a multigroup form. The working expectation is that the affective paths through pleasure and arousal will remain relatively stable across cultures because they are tied to general responses to music and editing, while the path that requires decoding of cultural signs will be stronger in the high-context sample. Portugal failed pilot (high-context) further indicates that socioeconomic status should be recorded, because it may dampen affect and DBR even when cues are culturally fit. Such a program of cross-cultural replications would clarify whether destination marketers need separate audio–visual versions for different markets or whether a single global version can be used with only caption or framing changes, such as musical metaphors, that will enhance or transform familiar destination perceptions.

Cultural embeddedness was manipulated with high-versus low-context edits rooted in Chinese motifs. The small but consistent advantage for high-context edits on the appraisal check suggests easier access to meaning, yet broader tests are needed. Cross-cultural experiments should vary traditions, instruments, rhythmic archetypes, and visuals, while measuring cultural familiarity at the respondent level. The Portugal pilot points to the need to record socioeconomic status and time pressure, which may dampen affect and DBR even when cues are culturally fit.

CRedit authorship contribution statement

Joston Gary: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yang Gu:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation. **Hannah Wang:** Writing – review & editing, Resources, Project administration, Investigation, Conceptualization. **Xixing Zhou:** Writing – review & editing, Resources, Project administration, Investigation, Formal analysis, Conceptualization. **Yan Feng:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Conceptualization. **António Carrizo Moreira:** Writing – review & editing, Supervision, Project administration, Formal analysis, Data curation, Conceptualization.

Informed consent statement

Informed consent was obtained from all subjects involved in the study. Writing consent was obtained from all individuals involved in this study.

Other statement

Due to funding constraints, sample sizes were calculated using G*Power software to achieve the minimum effect size detection threshold with acceptable Type II error rates.

Ethics statement

This study adhered to the ethical principles set forth in the Declaration of Helsinki. Prior to responding to these questionnaires, participants were provided with written informed consent materials detailing the study's objectives and procedures. Consent was obtained via the online survey platform used for data collection. Participants were clearly informed that their involvement was completely voluntary and that they could withdraw at any point without facing any negative consequences. Furthermore, robust confidentiality protocols were implemented to safeguard their responses.

Funding

This work was financially supported by FCT - Portuguese Foundation for Science and Technology (Grant No. UI/BD/154682/2022) for this study. We also appreciate the funding support from Linköping University — Library.

Declaration of competing interest

The authors have declared no conflict of interest.

Acknowledgments

We sincerely thank the editor and anonymous reviewers for their thoughtful feedback and valuable suggestions, which significantly improved the quality of this paper. We wish to acknowledge Bruno Huang (黄凌奕, brunohuang2012@gmail.com) and Ricardo Miguel Yang (杨安速, ricardo.yang00@gmail.com), bachelor's degree in Informatics and Computing Engineering at Faculty of Engineering, University of Porto in Portugal, as well as Filipe Chen (陈秒, up202104945@up.pt), a master's degree student in Management at FEP – School of Economics and Management, University of Porto in Portugal. We are deeply grateful for their contributions to the design of the experimental stimuli, particularly in the development of locally relevant Portuguese music and video materials. Their feedback in the early stages of the project offered highly practical insights that aligned closely with the tone and preferences of local audiences. They also played a key role in evaluating the cultural fit between Portuguese language context and the music–video pairings used in the study. Although this component of the research was eventually discontinued due to funding limitations and challenges in obtaining a sufficient Portuguese sample, we wish to express our highest respect and appreciation for their dedication and contributions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jretconser.2025.104637>.

Data availability

Please refer to the research data only for supplementary material (Supplementary_A to E).

References

- Asano, R., Lo, V., Brown, S., 2022. The neural basis of tonal processing in music: an ALE meta-analysis. *Music Sci.* 5. <https://doi.org/10.1177/20592043221109958>.
- Ayrton, R., 2020. The case for creative, visual and multimodal methods in operationalising concepts in research design: an examination of storyboarding trust

- stories. *Sociol. Rev.* 68 (6), 1229–1249. <https://doi.org/10.1177/0038026120903918>.
- Baiocchi, A., Dantas, D., Pessôa, L., Colbert, F., 2024. Cultural branding of cities: the role of live music in building a city's brand. *Cult. Trends* 34, 230–248. <https://doi.org/10.1080/09548963.2024.2345836>.
- Bakri, J., 2024. The power of metaphor in the representation of mental images in the language of tourism print advertising: Lakoff and Johnson's model of conceptual metaphor. *CaLLs (Journal of Culture, Arts, Literature, and Linguistics)*. <https://doi.org/10.30872/calls.v10i1.15468>.
- Barnes, S., Boehnke, K., Kumar, V., Yu, Z., 2023. Smooth talking and fast music: understanding the importance of voice and music in travel and tourism ads via acoustic analytics. *J. Trav. Res.* 62 (8), 1657–1676. <https://doi.org/10.1177/00472875231185882>.
- Blanco-Encomienda, F.J., Rosillo-Díaz, E., 2025. Exploring brand attitude in e-commerce using the S-O-R model: the role of information quantity, information quality and source credibility. *J. Prod. Brand Manag.* <https://doi.org/10.1108/JPBMM-04-2024-5095>. Advance online publication.
- Cao, D., Meadows, M., Ma, X., 2024. Thinking fast and slow: a revised SOR model for an empirical examination of impulse buying at a luxury fashion outlet. *Eur. J. Market.* 58 (1), 342–368. <https://doi.org/10.1108/EJM-01-2022-0046>.
- Chen, Y., Liu, T., 2024. Metaphorical mapping and cultural significance in Chinese death-related idiomatic expressions. *Metaphor Symbol* 39, 149–168. <https://doi.org/10.1080/10926488.2024.2303660>.
- Coronel, M., Irimías, A., 2022. "A very special song from Queen to you!" the role of music in destination promotional videos. *Consumer Behav. Tour. Hospital.* <https://doi.org/10.1108/cbth-11-2021-0265>.
- Demjén, Z., Semino, E., 2020. Metaphor, metonymy and framing in discourse. <https://doi.org/10.1017/9781108348195.011>.
- Ervas, F., Rossi, M., Ojha, A., Indurkha, B., 2021. The double framing effect of emotive metaphors in argumentation. *Front. Psychol.* 12. <https://doi.org/10.3389/fpsyg.2021.628460>.
- Fan, Y., Wong, I., Lin, Z., 2023. How folk music induces destination image: a synthesis between sensory marketing and cognitive balance theory. *Tour. Manag. Perspect.* <https://doi.org/10.1016/j.tmp.2023.101123>.
- Forceville, C., 2008. Metaphor in pictures and multimodal representations. In: Gibbs Jr., R.W. (Ed.), *The Cambridge Handbook of Metaphor and Thought*. Cambridge University Press, Cambridge, pp. 462–482 chapter.
- Gan, J., Shi, S., Filieri, R., 2023. Short video marketing and travel intentions: the interplay between visual perspective, visual content and narration appeal. *Tour. Manag.* 99, 104795. <https://doi.org/10.1016/j.tourman.2023.104795>.
- Golubkova, E., 2021. Mixed multimodal metaphors in advertising in English. <https://doi.org/10.15405/epsbs.2021.12.76>.
- Hall, E.T., 1976. *Beyond Culture*. Anchor Press, Garden City, NY.
- Haoyue, L., Cho, H., 2024. Factors influencing intention to engage in human-chatbot interaction: examining user perceptions and context culture orientation. *Univ. Access Inf. Soc.* 24, 607–620. <https://doi.org/10.1007/s10209-023-01087-7>.
- Hornikx, J., le Pair, R., 2017. The influence of High-/Low-Context culture on perceived Ad complexity and liking. *J. Global Market.* 30 (4), 228–237. <https://doi.org/10.1080/08911762.2017.1296985>.
- Jiang, Mengyin, Garrod, Brian, 2025. Stirring the emotions: modelling the role of emotions in developing satisfaction and loyalty in tea tourism experiences. *Curr. Issues Tourism.* <https://doi.org/10.1080/13683500.2025.2478285>.
- Khairani, K., Aruan, D.T.H., Gayatri, G., Hati, S.R.H., 2025. Urban tourism experiences: linking destination attributes to tourist behaviour through stimulus-organism-response theory. *Tour. Recreat. Res.* 1–18. <https://doi.org/10.1080/02508281.2025.2521274>.
- Kong, J., Lou, C., 2023. Do cultural orientations moderate the effect of online review features on review helpfulness? A case study of online movie reviews. *J. Retailing Consum. Serv.* <https://doi.org/10.1016/j.jretconser.2023.103374>.
- Lakoff, G., Johnson, M., 1980. The metaphorical structure of the human conceptual system. *Cogn. Sci.* 4 (2), 195–208. <https://doi.org/10.1207/s15516709cog0402.4>.
- Lan, C., Yin, Z., 2020. A cognitive approach to the conceptual metaphors in Shi Jing (the Book of Poetry). *Appl. Linguist. Rev.* 14, 173–195. <https://doi.org/10.1515/applirev-2020-0076>.
- Liu, J., 2025. Implicit communication and cultural symbols in advertising: an applied analysis from the perspective of media semiotics. *Advances in Social Behavior Research.* <https://doi.org/10.54254/2753-7102/2025.22339>.
- Liu, Y., Li, D., 2022. Multimodal metaphor (re)framing: a critical analysis of the promotional image of China's Hubei Province in the post-pandemic era on new media. *Soc. Semiot.* 34, 269–288. <https://doi.org/10.1080/10350330.2022.2094233>.
- Moreno-Lobato, A., Di-Clemente, E., Hernández-Mogollón, J., Campón-Cerro, A., 2023. How emotions sound. A literature review of music as an emotional tool in tourism marketing. *Tour. Manag. Perspect.* <https://doi.org/10.1016/j.tmp.2023.101154>.
- Noble, J., 2022. The career of metaphor hypothesis and vocalicity in contemporary music. *Interdiscip. Sci. Rev.* 47, 259–278. <https://doi.org/10.1080/03080188.2022.2035108>.
- Osemwegie, O., 2025. Cross-cultural narrative branding in global entertainment: strategic communication for market penetration and retention. *Int. J. Appl. Res. Soc. Sci.* <https://doi.org/10.51594/ijarss.v7i3.1819>.
- O'Dowd, N.A., 2024. The framing effects of metaphor and metonymy in multimodal campaigns about plastic pollution: a solutions-oriented and emotion-oriented study. *Metaphor Symbol* 40 (1), 76–97. <https://doi.org/10.1080/10926488.2024.2429988>.
- Peh, M.L., 2025. Digital storytelling: crafting story content with moving images across multiple platforms. In: Kan, R.Y.P., Khoo, C.S.G. (Eds.), *Signature Pedagogies for*

- Professions in Arts and Design. Springer, Singapore. <https://doi.org/10.1007/978-981-96-2616-8-8>.
- Rather, R., Najjar, A., Jaziri, D., 2020. Destination branding in tourism: insights from social identification, attachment and experience theories. *Anatolia* 31, 229–243. <https://doi.org/10.1080/13032917.2020.1747223>.
- Requejo, M., 2016. Music in multimodal narratives: the role of the soundtrack in digital stories. In: Mildorf, J., Kinzel, T. (Eds.), *Audionarratology: Interfaces of Sound and Narrative*. De Gruyter, Berlin, Boston, pp. 29–46. <https://doi.org/10.1515/9783110472752-003>.
- Richardson, R.M., Smith, S.W., 2007a. The influence of high/low-context culture and power distance on choice of communication media: students' media choice to communicate with professors in Japan and America. *Int. J. Intercult. Relat.* 31 (4), 479–501. <https://doi.org/10.1016/j.ijintrel.2007.01.002>.
- Richardson, R.M., Smith, S.W., 2007b. The influence of high/low-context culture and power distance on choice of communication media: students' media choice to communicate with professors in Japan and America. *Int. J. Intercult. Relat.* 31 (4), 479–501. <https://doi.org/10.1016/j.ijintrel.2007.01.002>.
- Shavitt, S., Barnes, A.J., 2020. Culture and the consumer journey. *J. Retailing* 96 (1), 40–54. <https://doi.org/10.1016/j.jretai.2019.11.009>.
- Tian, D., Abdullah, M., 2024. Chinese-english translation strategies for metaphors based on domestication and foreignization. *Arab World English J. Transl. Liter. Stud.* <https://doi.org/10.24093/awejlts/vol8no4.4>.
- Ting, L., Ahn, J., 2023. Luxury brand cultural symbolism: signaling effect on customer well-being. *J. Strat. Market.* 32, 590–605. <https://doi.org/10.1080/0965254X.2023.2275134>.
- Trompeta, M.-A., Karantinou, K., Koritos, C., Bijmolt, T.H.A., 2022. A meta-analysis of the effects of music in tourism and hospitality settings. *J. Bus. Res.* 138, 130–145. <https://doi.org/10.1016/j.jbusres.2021.08.067>.
- Warner-Soderholm, G., 2013. Beyond a literature review of Hall's context dimension: scale development, validation & empirical findings within a Norwegian study. *Int. J. Bus. Manag.* 8 (10), 27–40. <https://doi.org/10.5539/ijbm.v8n10p27>.
- Wu, C., Yama, H., Zakaria, N., 2023. How do Japanese and Chinese view each other? Understanding the meaning of low-context culture in intercultural communication. *Glob. Netw.* <https://doi.org/10.1111/glob.12440>.
- Xu, Z., 2024. The evolution of cultural symbols in graphic design. *Front. Art Res.* <https://doi.org/10.25236/far.2024.060415>.
- Yang, X., Hou, J., Arth, Z., 2021a. Communicating in a proper way: how people from high-/low-context culture choose their media for communication. *Int. Commun. Gaz.* 83, 238–259. <https://doi.org/10.1177/1748048520902617>.
- Yang, X., Hou, J., Arth, Z.W., 2021b. Communicating in a proper way: how people from high-/low-context culture choose their media for communication. *Int. Commun. Gaz.* 83 (3), 238–259. <https://doi.org/10.1177/1748048520902617>.
- Yang, X., Hou, J., Arth, Z.W., 2021c. Communicating in a proper way: how people from high-/low-context culture choose their media for communication. *Int. Commun. Gaz.* 83 (3), 238–259. <https://doi.org/10.1177/1748048520902617>.
- Yen, C., Teng, H., Chang, S., 2020. Destination brand identity and emerging market tourists' perceptions. *Asia Pac. J. Tourism Res.* 25, 1311–1328. <https://doi.org/10.1080/10941665.2020.1853578>.
- Yu, N., 2023. Metaphors from perception and culture. *Cognitive Linguis. Stud.* <https://doi.org/10.1075/cogls.00106.yu>.
- Zhang, X., 2024. Cognitive perspective of metaphors in Chinese Hua'er folksongs. *Front. Commun.* <https://doi.org/10.3389/fcomm.2023.1274622>.
- Zhang, C., Lin, Z., Jin, S., 2022. What else besides war: deliberate metaphors framing COVID-19 in Chinese online newspaper editorials. *Metaphor Symbol* 37, 114–126. <https://doi.org/10.1080/10926488.2021.1948333>.
- Zhu, N., Xu, H., Zhang, X., Lu, Y., 2025. Research on the factors influencing tourist loyalty to outdoor music festivals: an application of stimulus-organism-response paradigm. *Front. Psychol.* 16, 1553211. <https://doi.org/10.3389/fpsyg.2025.1553211>.
- Zhuang, M., Zhang, H., Li, P., Shen, C., Xiao, X., Zhang, J., 2023. Connecting tourists to musical destinations: the role of musical geographical imagination and aesthetic responses in music tourism. *Tour. Manag.* <https://doi.org/10.1016/j.tourman.2023.104768>.
- Zoghaib, A., Luffarelli, J., Feiereisen, S., 2023. Branding with music: how can music contour and tonality enhance perceived brand innovativeness and brand evaluations? *Psychol. Market.* <https://doi.org/10.1002/mar.21875>.
- Juston Gary** is a Ph.D. Candidate in Marketing at University of Aveiro (Portugal, EU) and Linköping University (Sweden, EU). He holds an M.Phil. in Business Administration (Marketing) from the City University of Macau, Macau S.A.R; B.Eng. in Civil Engineering from Chongqing Institute of Engineering, China; B.Sc. in Psychology from JSS Academy of Higher Education and Research, India; B.Sc in Computer Science from University of The People, USA. His research applies quantitative methods in marketing, psychology, and AI, focusing on SEM, machine learning, and cross-cultural studies, particularly on music's influence on consumer behavior. He has published in Q1 Journals. He has also trained in Business Analytics, Data Science, and Digital Marketing, holding data analysis roles with The Hong Kong Polytechnic University and City University of Macau. He is ABS Journal Reviewer.
- Yang Gu** is a PhD candidate in Business Research Unit at Iscte - Instituto Universitário de Lisboa and Maastricht University. He holds a MRes of Research in Management and Innovation from SKEMA Business School and Université Côte d'Azur, a Bachelor degree of Business Administration from City University of Macau. His research interests include organization theory, institutional theory, corporate governance and organization behavior.
- Yudan Wang (Hannah Wang)** was born in Sichuan, China. She began her piano studies at the age of 5 under the influence of her aunt. After moving to Beijing to study at the China Conservatory of Music, Yudan completed her postgraduate piano studies at the University of Aveiro in Portugal from 2019 to 2021 and continued her doctoral studies at University of Aveiro, from 2025 to present, pursuing PhD program at University of Évora, while she also attended the Conservatorio Superior de Música de Liceu in Spain from 2021 in order to further her studies to help her doctoral project. Since 2015, Yudan has been the winner of many international piano competitions and she regularly performs in Europe and Asia, as well as performing chamber music with a variety of artists. In 2018, she won the first prize in China at the 20th Santa-cecilia International Piano Competition and the first prize in China at the Nagoya Piano Competition in Japan. In April 2019, she gave a recital in Beijing and then went to Portugal to participate in the Viseu International Piano Competition, where she won the Excellence Award. In May, she won the Bronze Prize in China at the XXI Santa-cecilia International Piano Competition (the Gold and Silver Prize was not available that year) and became the only Chinese pianist in the adult category selected for the Asian competition. October 2020, she invited to perform as a distinguished pianist at the Sardeal International Piano Festival in the seminary. In July 2021, she graduated with the highest score ever, with a score of 19 (20). In the same month, Yudan published her article (Aural and visual imagery in the multimedia music performance) in *Avanca International* magazine. In the same month she received an invitation from the Chinese Director of the Confucius Institute in Portugal, Han Ying, to collaborate with the Confucius Institute for a year in the promotion of Chinese festivals. In August Yudan participated in the obidos Festival in Portugal and performed as an outstanding pianist. In September, she was accepted to study traditional Spanish music with Dainel Ligorio, a disciple of the Spanish national pianist Larrocha and a Catalan pianist, at the Liceu Conservatory of Music in Barcelona, Spain. In November she was invited to perform at the opening of the Ceramics Biennale by Carlos Morais, the Portuguese Director of the Portuguese Confucius Institute.
- Xixing Zhou** is currently pursuing her PhD in Cultural Studies at the School of Social Sciences, University of Minho, Portugal. She holds a Bachelor's degree in Literature from Henan University of Science and Technology, China, and a Master's degree in International Chinese Education from Soochow University, China. Her research areas include Education, Second Language Teaching, Cultural Studies, Intercultural Communication, and Communication Science. Miss. Zhou's research focuses on enhancing cultural diversity and intercultural communication competence in the context of globalization.
- Yan Feng** is now a postdoctoral researcher at CIES-ISCTE (Lisbon University Institute) in POIESIS Project (Probing the impact of integrity and integration on societal trust in science). She has a Bachelor's degree in International Trade and Economy and a Master's degree in Industrial Economy in China, and then received her Ph.D. in Management (specialization in Human Resources and Organizational Behavior) from ISCTE-Lisbon University Institute in 2022. Her research interest include research integrity, cross-cultural studies, thinking styles and emotional regulation in working context, behavioral science.
- António Carrizo Moreira** received his Bachelor's degree in Electrical Engineering and a Master's degree in Management from the University of Porto, Portugal. He holds a Ph.D. from the University of Manchester, England. He received scholarships from two important institutions: Science and Technology Research National Board (JNICT) and Science and Technology Foundation (FCT). He is Professor at Universidade Aberta, Portugal.