Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Optimization and Psychometric Validation of the Chinese Barcelona Musical Reward Questionnaire (C-BMRQ)

Xintong Jiang
Institute of Psychology, Chinese Academy of Sciences
Shan Hua
Institute of Psychology, Chinese Academy of Sciences
Qin He
Institute of Psychology, Chinese Academy of Sciences
Yi Du
duyi@psych.ac.cn

Institute of Psychology, Chinese Academy of Sciences

Research Article

Keywords: Music reward sensitivity, Music anhedonia, Barcelona Music Reward Questionnaire, C-BMRQ, Cross-cultural adaptation, Measurement invariance, Psychometric validation

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11511069/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Music reward sensitivity—individual differences in the pleasure and motivation associated with music—is a measurable psychological construct. The Barcelona Music Reward Questionnaire (BMRQ) is a widely used five-factor instrument, but the existing Chinese adaptation relied on a geographically restricted student sample, omitted four items to achieve acceptable model fit, and did not examine test–retest reliability or measurement invariance. This study aimed to optimize and psychometrically validate the Chinese BMRQ (C-BMRQ) while preserving its original 20-item, five-factor structure.

Methods

Study 1 compared four Chinese versions—two literal translations, one culturally adapted translation developed here, and the previously published 20-item and 16-item versions—in two Beijing samples (N = 387 and 391). Study 2 validated the optimized version in a nationwide pooled sample of 1,650 Chinese adults, including a three-month retest subsample (N = 472). Analyses included confirmatory factor analysis (CFA), internal consistency, longitudinal measurement invariance, configural and metric invariance across gender and music-training groups, and convergent validity with the Goldsmiths Musical Sophistication Index (Gold-MSI) and the Physical Anhedonia Scale.

Results

The culturally adapted version showed better model fit than literal translations and allowed retention of all 20 items without deletion. In the pooled sample, the five-factor model reached acceptable fit (CFI = 0.921, TLI = 0.906, RMSEA = 0.066). Internal consistency was strong, and three-month test–retest correlations ranged from 0.64 to 0.75. Longitudinal measurement invariance was supported at the configural and metric levels. Configural and metric invariance were also supported across gender and music-training groups. Convergent validity was evidenced by theoretically consistent associations with Gold-MSI and the Physical Anhedonia Scale. The 10th percentile of the pooled distribution corresponded to a total C-BMRQ score of 69, providing a sample-derived research reference for low music reward sensitivity; this value requires independent clinical validation before diagnostic use.

Conclusions

The optimized C-BMRQ is a psychometrically robust and culturally appropriate instrument that retains the full 20-item, five-factor structure of the original BMRQ. It is suitable for research on individual differences in music reward and music anhedonia in Chinese-speaking populations, and offers a replicable template for culturally adapting instruments without sacrificing construct coverage.

Introduction

Music elicits profound physiological and emotional responses in listeners [1, 2], yielding pleasure and aesthetic experiences that are collectively recognized as “musical reward” [3, 4]. However, not everyone perceives music as rewarding. While some individuals are deeply immersed in extremely beautiful music, they often experience “chills”, manifested as scalp tingling and goosebumps [5] and subsequently activate the brain's reward system and release dopamine [6]. Others remain indifferent—they retain the ability to enjoy physical rewards such as food and money [7], but exhibit markedly reduced musical reward responsiveness, which is known as “music anhedonia” [8, 9].

This specific individual difference necessitates specialized measurement tools. Conventional general reward scales (e.g., monetary or food reward measures) are ill-suited for quantifying musical reward sensitivity, as they fail to capture domain-specific dimensions of musical experience. To address this, the Barcelona Music Reward Questionnaire (BMRQ) was developed as a critical tool for assessing this construct.

The Barcelona Music Reward Questionnaire (BMRQ) comprises five dimensions (4 items each)—Music Seeking, Sensorimotor, Emotion Regulation, Emotional Evocation, and Social Reward—and 20 items scored on a 5-point Likert scale. It has demonstrated robust reliability and validity across diverse populations globally, including North America [10], South America [11], European countries [12–14], and East Asia [15, 16].

Accumulating evidence suggests that musical reward sensitivity is shaped by a range of demographic and individual cognitive factors, with notable cross-cultural variability [10, 17]. Among these moderators, gender and musical training are the most robust and consistently replicated predictors of individual differences in musical reward experience across cultural contexts. Specifically, higher musical reward sensitivity was found in females [10, 18, 19] and musicians [10, 18, 20]. Consequently, evaluating measurement invariance across these key demographic subgroups is a psychometric prerequisite to ensure that the C-BMRQ yields valid and unbiased comparisons in future research.

Given profound cross-cultural divergence in musical aesthetic templates and perceptual frameworks between Eastern and Western populations, simple literal translation is insufficient for scale adaptation; rigorous relocalization is therefore crucial when adapting the BMRQ for Chinese contexts. Wang et al. (2023) undertook an initial Chinese adaptation of the BMRQ, laying a foundational groundwork for research on musical reward within Chinese-speaking contexts. Nevertheless, this initial adaptation carries certain practical and psychometric limitations that motivate further refinement: in terms of reliability, its highly homogeneous sample of students from only four universities in Changsha undermines generalizability, and the absence of a retesting procedure leaves test-retest reliability unexamined; regarding validity, translations with limited cultural contextual consideration yielded low factor loadings, prompting item deletions that left some dimensions with only two items, thereby compromising construct coverage; while the lack of norm development further restricts the scale’s applicability and score interpretation.

In the present investigation, we aimed to address these critical limitations through a rigorous re-localization process. In Study 1, we compared four versions: two developed here (a 20-item literal translation and a 20-item culturally adapted translation), and two versions from Wang et al. [19], including the 20-item original translation and their 16-item revised version, with the goal of selecting the optimal Chinese version. We hypothesized that, while retaining the full 20-item structure of the original scale, the culturally adapted version would yield improved model-fit statistics, higher factor loadings, and higher internal consistency relative to the literal translation.

Study 2 further validated the optimal candidate version using a nationwide pooled sample (N = 1,650) by evaluating its test-retest reliability and convergent validity against two theoretically relevant constructs: general reward sensitivity measured by the Chinese Physical Anhedonia Scale (CPAS), and music-related sophistication measured by the Goldsmiths Musical Sophistication Index (Gold-MSI). Following comprehensive psychometric evaluation, this candidate scale was finalized as the formal Chinese-version BMRQ (C-BMRQ). We additionally evaluated measurement invariance, explored demographic group differences, and generated normative score distributions to support future research on music-reward sensitivity among Chinese adults.

Method

Study 1: Optimization of the Chinese versions of Barcelona Music Reward Questionnaire (BMRQ)

The aim of the Study 1 was to re-validate the Barcelona Musical Reward Questionnaire (BMRQ) and identify the optimal Chinese version suitable for measuring musical reward experiences. With that purpose, we compared four Chinese versions: our original literally-translated (20 items, linguistically faithful to English) and adjusted-translated (20 items, culturally adapted) versions, and Wang et al. [19]'s original (20 items) and revised (16 items, four items deleted) versions.

Participants

Two rounds of data collection were conducted online through Wenjuanxing (https://www.wjx.cn) in the Beijing area. Data from the first round served for the preliminary validation in Study 1, whereas data from the second round were reserved for the pooled dataset reported in Study 2. In the first-round collection, 453 participants aged 18–35 years finished the literally-translated version of BMRQ, and in the second-round collection, 430 participants aged 18–35 years finished the adjusted-translated version of BMRQ (C-BMRQ). We excluded participants who failed the attention check (66 in the first round and 39 in the second round). The final sample for Study 1 (N = 387) came from the first-round collection (156 males, 231 females, 23.42 ± 3.54 years old). The second-round valid sample (N = 391; 123 males, 268 females, 22.46 ± 2.90 years old) is incorporated into Study 2's pooled dataset (see Study 2 Participants section). The prespecified five-factor, 20-item CFA model was fitted separately in each sample. Both samples

exceeded the COSMIN recommendation of at least seven participants per item and N ≥ 100 for factor analysis, corresponding to a minimum N of 140 for the present 20-item scale [21, 22]. This criterion was treated as a practical sample-adequacy benchmark rather than evidence of model-specific statistical power. In the first-round collection, 74 participants (26 males, 48 females) reported having professional music training before (mean duration: 6–9 years) and 30 participants (9 males, 21 females) reported having professional music training (mean duration: 3–5 years) in the second round. There was no overlap between the two rounds. All participants were provided with digital informed consent and paid 5 RMB for their participation. The research was approved by the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences (ID: H25086).

Measures

Barcelona Music Reward Questionnaire (BMRQ). The English version of BMRQ [10] was a 20-item self-report measure structured into five theoretically derived facets—Musical Seeking, Emotion Evocation, Mood Regulation, Sensory-Motor, and Social Reward—each comprising four items designed to assess distinct dimensions of musical reward experiences. Items 2 and Item 5 are reverse coded to mitigate acquiescence bias, requiring re-coding during data analysis. Participants respond to items on a 5-point Likert scale (1 = *Fully disagree* to 5 = *Fully agree*), yielding total and subscale scores for each facet. Higher total scores indicate stronger overall musical reward sensitivity, while subscale scores reflect specific dimensions (e.g., a high “Emotion Evocation” score suggests greater emotional responsiveness to music).

Translation and cross-cultural adaptation

Guided by the International Test Commission (ITC) Guidelines for Translating and Adapting Tests [23], we adopted a forward-backward translation procedure.

Forward translation and reconciliation.

Two independent translators—an English major and a music psychology Ph.D. candidate—conducted separate forward translations from English to Chinese. Both translators were bilingual and possessed deep knowledge of Chinese culture and held either test translation experience or domain expertise in the scale’s subject matter, ensuring accurate technical term translation. An expert then reviewed and reconciled the two Chinese versions.

Back-translation and review.

A backward translation was subsequently performed by an independent translator who was blind to the original English scale. Discrepancies were resolved through comparison between the original and back-translated English versions. Finally, a bilingual psychology translator provided feedback on equivalence and finalized the initial "literally-translated version" through comprehensive revisions.

Cultural adaptation and refinement.

The literally-translated version was first administered in a preliminary online survey. As it exhibited insufficient construct validity, the research team systematically revised the items to ensure cultural alignment. For instance, "16. At a concert I feel connected to the performers and the audience." was initially translated as "□□□□□, □□□□□□□□□□□□□□□□□□□□□□□□"; however, concerts are a highly specialized and infrequently experienced form within Chinese music performances, which limits the generalizability of statements related to them. Therefore, "*concert*" was replaced with the more inclusive term "*music performance*", transforming the English item into "□□□□□□□, □□□□□□□□□□□□□□□□□□□□□□□□" ("At a music performance, I feel a connection with the performers and other audience members"). Similarly, "10. Music often makes me dance." (initially translated as "□□□□□□□□□"). Given that Chinese respondents tend to express their musical engagement in a more subtle and reserved manner rather than *overt dancing*, this item was revised to "□□□□□□□□□□□, □□□□□□□□□□□□□" ("When exposed to strongly rhythmical music, I occasionally move or sway in time with the music.").

Furthermore, certain phrases from the original English questionnaire, such as "7. I inform myself about music I like." (initially translated as "□□□□□□□□□□□□□□□□□"), proved ambiguous and hard to comprehend within the Chinese cultural context. Thus, it was revised to "I will actively seek out the music I like to listen to" ("□□□□□□□□□□□□□□□□□") to better fit the music-seeking dimension. Moreover, overly intense frequency adverbs were toned down. For instance, "17. I spend quite a bit of money on music and related items." ("□□□□□□□□□□ □□□□□□□□□") was changed to "I am willing to spend money on music-related products and activities." ("□□□□□□ □□□□□□□□□□□□□□□"). The comparison between the literally-translated and adjusted-translated versions is shown in **Table S1** in the **Supplemental materials**.

Statistical analysis

All analyses were conducted in R (version 4.3.1) using the *lavaan* package [24] for structural equation modeling. Consistent with the structure reported in Mas-Herrero et al. [10], Confirmatory Factor Analyses (CFAs) were carried out on each sample separately using maximum-likelihood estimation with 2000-iteration bootstrap standard errors. Model fit was evaluated using both absolute and relative fit indices: χ^2 test, Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Tucker–Lewis Index (TLI). All standardized factor loadings and latent covariances were estimated and reported.

Study 2: Comprehensive Validation and Empirical Application of the C-BMRQ

Building on Study 1, which identified the optimal candidate Chinese version of the BMRQ (C-BMRQ), Study 2 focused on validating aspects left unaddressed. Although Study 1 established the C-BMRQ's basic psychometric properties, it did not evaluate its test-retest stability, subgroup equivalence, or convergent validity. In Study 2, we employed three datasets for further psychometric evaluation: one subsample drawn from Study 1 for retesting, and two datasets from separate independent behavioral studies that administered this candidate scale. This allowed us to: (1) evaluate three-month test-retest

reliability; (2) examine score differences across gender and music-training groups; (3) establish normative score distributions for Chinese-adult participants; and (4) explore convergent validity via with theoretically relevant self-report instruments: correlations with the PAS and Gold-MSI.

Participants

Dataset 1 (N = 1,019): To further evaluate the validity and reliability of the optimal Chinese version of BMRQ (C-BMRQ), we recruited a new sample of 1,087 participants (aged 18–35 years) across China to complete the online survey. After excluding 68 participants who failed the attention check, the final sample consisted of 1,019 individuals (278 males, 741 females, 22.31 ± 2.80 years old). Of the final sample, 215 participants (68 males, 147 females) reported prior professional music training (mean duration: 3–5 years).

The five-factor CFA model specified in Study 1 was refitted to Dataset 1 (N = 1,019). This sample exceeded the COSMIN benchmark and fell within the range reported in published factor-analytic validations of other BMRQ versions (N = 209—1,550), supporting its adequacy for structural-validity analysis without implying model-specific statistical power.

Additionally, 472 participants (89 males, 383 females) from the final sample completed the retest of the C-BMRQ after 3 months, and 92 (12 males, 80 females) of them reported prior professional music training (mean duration: 3–5 years). Given the relatively large sample and well-identified measurement structure, the sample size was considered adequate for longitudinal CFA and measurement invariance testing (Wolf et al., 2013; Kline, 2023).

Datasets 2 and 3 (from other behavioral studies, N = 136 for Dataset 2 and N = 110 for Dataset 3): To enhance the generalizability of our findings, we also incorporated data from two unpublished independent behavioral studies that applied the C-BMRQ. In Dataset 2, after excluding 10 participants who failed the attention check, 136 participants (81 males, 55 females; 22.18 ± 2.41 years old) were included, and their music training experience was not assessed. In Dataset 3, after excluding 3 participants who failed the attention check, 110 music responders (defined as individuals who experience “chill” responses to music; 48 males, 62 females; 23.53 ± 2.93 years old) were included, and of these 84 participants (36 males, 48 females) had professional music training experience (mean duration: 2 years). These two subsets, drawn from separate unpublished behavioral studies, were used solely for convergent-validity correlation analyses. Given the size of these extracted subsets, correlational results should be interpreted with appropriate caution.

Pooled data (N = 1,650, Study 1-round 2 and Study 2)

Subsequently, all samples that completed the C-BMRQ were pooled, including the second-round sample from Study 1 and the three datasets from the Study 2. Duplicate responses (5 from Dataset 2 and 1 from Dataset 3) were excluded by retaining only each participant’s first submission, yielding a final combined sample of 1,650 participants (527 males, 1,123 females; 22.41 ± 2.82 years old). This pooling strategy

aimed to increase statistical power and account for potential variability across different subgroups. Participant distributions across Study 1 and Study 2 datasets are presented in **Table S2**.

All participants in Study 2 provided informed consent, and received compensation of 5 RMB for online participation or 50 RMB per hour for offline participation. All studies in Study 2 were approved by the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences (ID: H25086).

Measures

Chinese-version Barcelona Music Reward Questionnaire (C-BMRQ). The C-BMRQ was identified as the optimal version among four versions in Study 1. It was adaptively translated to fit the Chinese context, demonstrates construct consistency with the original BMRQ, and exhibits excellent reliability and validity. The C-BMRQ consists of five dimensions and 20 items in total, each dimension with 4 items, and 5-point Likert rating scale (see the **Measures** section of **Method** in Study 1 and the adjusted-translated version in Table 1 for more details).

Table 1
Comparison of fit measures of 4 Chinese versions of BMRQ in Study 1

		Sample size	χ^2 (*df*)	CFI	TLI	RMSEA (90% CI)
Present Study	Literally translated	387	638.74 (160)	0.834	0.802	0.088 (0.081–0.095)
	Adjusted translated	391	452.19 (160)	0.899	0.880	0.068 (0.061–0.076)
Wang et al., 2023	Original	1120	1565.96 (160)	0.813	0.778	0.089 (0.085–0.093)
	Revised	1120	493.05 (94)	0.936	0.918	0.062 (0.056–0.067)

Note. Literally translated = literal word-for-word Chinese translation of the original BMRQ items; Adjusted translated = culturally fine-tuned Chinese translation from the present study; Original = the initial Chinese translation reported by Wang et al. (2023); Revised = the item-deleted 16-item revised version from Wang et al. (2023). CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root-mean-square error of approximation; df = degrees of freedom.

Chinese versions of the Revised Physical Anhedonia Scale (CPAS). We also employed the CPAS as a calibration scale for the reward-related pleasure sensitivity. [25]. CPAS was translated from the Physical Anhedonia Scale (PAS; [26]), which assesses self-reported deficits in experiencing pleasure from physical stimuli. Comprising 61 dichotomous (True/False) items, the scale evaluates an individual's inability to derive enjoyment from prototypical physical experiences, including gustatory (e.g., food),

tactile, sexual, or environmental domains (e.g., relaxing settings). Total scores range from 0 to 61, with higher scores indicating greater severity of physical anhedonia. The CPAS has demonstrated good internal consistency (Cronbach's α = .86) in a non-clinical Chinese population in prior research [25].

The Simplified Chinese version of the Goldsmiths Musical Sophistication Index (Gold-MSI). The Simplified Chinese version of the Gold-MSI [27] was employed as a calibration measure for musical sophistication. Adapted from the original English version, the Gold-MSI assesses broad individual differences in musical expertise across diverse populations. It consists of 39 items: 31 statements rated on a 7-point Likert scale and 7 items with ordered response categories. The 38 scored items (excluding one filler item) are organized into five subscales—Active Engagement, Perceptual Abilities, Musical Training, Singing Abilities, and Emotion—and a general factor (General Musical Sophistication, GMS). The GMS score ranges from 18 to 126, with higher scores indicating greater musical sophistication, a broader range of musical behaviors, and more effective strategies for achieving goals in musical contexts. All subscales of the Simplified Chinese Gold-MSI have demonstrated good internal consistency in Mainland China (Cronbach's α = .80 ~. 91).

Procedure

Dataset 1. Consistent with Study 1, participants were recruited online using Wenjuanxing (https://www.wjx.cn), and a Wechat group was established for follow-up retesting three months later. The online questionnaire assessed demographic variables, music training experience, the C-BMRQ, and the CPAS.

Dataset 2. Derived from an online behavioral experiment, this dataset obtained demographic information, music training experience and two subscales (General Musical Sophistication and Music Training) of the Gold-MSI via Gorilla experiment builder(https://gorilla.sc).

Dataset 3. Derived from an offline behavioral experiment, this dataset obtained demographic information, music training experience, and full-scale data for the C-BMRQ and Gold-MSI online via an online questionnaire on Tencent Questionnaire (https://wj.qq.com).

Statistical analysis

Reliability and Construct Validity. Reliability and construct validity were evaluated using Dataset 1 (N = 1019). Consistent with Study 1, we conducted all the validation analyses of the C-BMRQ in R (version 4.3.1) using the *lavaan* package [24], with the same model fit indices. To examine the test–retest stability of the construct, longitudinal measurement invariance (LMI) was evaluated using a series of nested single-group CFA models, testing configural, metric, scalar, and strict invariance [28]. Model comparisons employed χ^2 difference tests, complemented by changes in the comparative fit index ($\Delta CFI \leq 0.01$) [29] and the root mean square error of approximation ($\Delta RMSEA \leq 0.015$) [30]. The most constrained model without significant deterioration in model fit was selected to confirm the stability of the measurement

structure over time. Test–retest reliability was assessed using Spearman's rank-order correlation, and significance levels were adjusted using the Bonferroni correction for multiple comparisons.

Construct equivalence. We used pooled data excluding Dataset 2 (due to missing music training data; N = 1,519), to test the construct equivalence between groups defined by gender and music training using multigroup confirmatory factor analysis (MGCFA) [28]. Configural, metric, scalar, and strict invariance models were fitted sequentially, with model comparisons based on χ^2 difference tests, $\Delta CFI \leq 0.01$, and $\Delta RMSEA \leq 0.015$ to determine the highest level of invariance supported. Because the gender and training groups were unbalanced (male/female = 448/1071, trained/untrained = 329/1190), all estimates are presented as 95% confidence intervals derived from 1,000 bootstrap samples. For each grouping factor, we controlled for the other factor and used equal-sized subsamples based on the smaller group size to reduce potential bias from unequal group sizes in the invariance tests.

Musical Reward Sensitivity and Group Differences. The overall level of musical reward sensitivity as measured by the C-BMRQ in the Chinese sample was examined using the pooled data (N = 1,650). Following Mas-Herrero et al. [7], low music reward sensitivity was operationalized for descriptive research purposes as a C-BMRQ total score below the 10th percentile of the pooled sample. This sample-derived threshold was not treated as a clinically validated diagnostic cutoff for music anhedonia. To test the potential influencing factor of musical reward sensitivity, we employed a two-stage resampling approach to compare gender (male vs. female) and music training (with vs. without): for each of 1,000-iteration bootstrap subsamples with replacement, matching the size of the smaller group, we performed 1,000 permutation tests on the t-statistic for mean-difference. This yielded an observed t-value and its percentile position within the permuted t-distribution for each subsample. Statistical significance was determined by examining whether the 95% confidence interval of this percentile excluded $\alpha = 0.05$. Because music training data were not collected in Dataset 2, this dataset was excluded from analyses comparing music training groups.

Convergent Validity. Spearman's rank-order correlations were carried out to examine the associations between the C-BMRQ and CPAS, Gold-MSI in separately in Datasets 2 and 3. Significance levels were adjusted using the Bonferroni correction for multiple comparisons.

Results

Comparison of the Chinese versions (Study 1)

Consistent with the original study by Wang et al. [19], the literally-translated version of the BMRQ in the present research demonstrated inadequate model fit (Table 1; CFI = 0.834, TLI = 0.802, RMSEA = 0.088, 90% CI [0.081, 0.095]). In particular, the reverse-scored items (2, 5, 10, 17) showed factor loadings below 0.4 in both datasets (Table 2), replicating the pattern reported by Wang et al. [19].

Table 2
Comparison of the factor loadings of 3 Chinese versions of BMRQ in Study 1

	Item	Present Study		Wang et al., 2023	
		Literally translated	Adjusted translated	Original (Not reported)	Revised
Social Reward	1	0.52	0.62		0.63
	6	0.69	0.71		0.66
	13	0.51	0.63		0.47
	16	0.72	0.64		0.59
Music Seeking	7	0.62	0.69		0.59
	2*	-0.32	**-0.43**		deleted
	11	0.57	0.66		0.50
	17	0.40	**0.62**		deleted
Emotion Evocation	3	0.60	0.66		0.66
	8	0.76	0.68		0.69
	12	0.60	0.59		0.49
	18	0.37	**0.55**		0.51
Mood Regulation	4	0.76	0.67		0.70
	9	0.76	0.68		0.74
	14	0.64	0.80		0.66
	19	0.79	0.79		0.74
Sensory-Motor	10	0.36	**0.70**		deleted
	5*	-0.26	**-0.52**		deleted
	15	0.75	0.69		0.79
	20	0.81	0.67		0.73

Note. Values in red indicate items with low factor loadings, and values in bold denote improved items originally with low factor loadings. *Reverse Item.

Following cultural and linguistic adjustment (see **Table S1** in the **Supplemental materials** for version comparisons in the current study), all previously low-loading items displayed improved psychometric properties after revision (Table 1; CFI = 0.899, TLI = 0.880, RMSEA = 0.068, 90% CI [0.061, 0.076]).

Psychometric evaluation of the C-BMRQ (Study 2)

Factor structure and model fit.

With a larger sample (N = 1019), the model fit of the original five-factor structure improved and reached an acceptable level ($\chi^2(160) = 866.69$, $p < .001$, CFI = 0.921, TLI = 0.906, RMSEA (90% CI) = 0.066 (0.062–0.070)). The absolute value of factor loadings ranged from 0.55 to 0.80 and all were statistically significant (*p*s < .001, Fig. 1).

Internal consistency and test-retest reliability

Table S4 shows fit indices for a series of nested longitudinal CFA models examined using the first test and retest datasets from the sample in Study 1 (N = 472). Longitudinal measurement invariance analyses supported configural invariance (CFI = 0.921, TLI = 0.909, RMSEA = 0.051, 90% CI [0.048, 0.055]) and metric invariance ($\Delta\chi^2(\Delta df) = 15.00$ (15), $p = .451$). The scalar model showed acceptable absolute fit (CFI = 0.920, TLI = 0.912, RMSEA = 0.050, 90% CI [0.047, 0.054]), but fit deteriorated significantly relative to the metric model ($\Delta\chi^2(\Delta df) = 35.26$ (20), $p = .019$), suggesting partial scalar non-invariance across test and retest. Crucially, partial scalar non-invariance should not be taken as evidence of poor reliability or structural instability. Given the supported configural and metric invariance, the C-BMRQ showed stable factor structure and factor loadings across test and retest, although some item intercepts varied over time. This may reflect shifts in baseline response levels or response thresholds, consistent with the view that music reward sensitivity can vary with time, context, and experience rather than being completely fixed [10, 31]. Taken together, these results indicated that the factor structure and factor loadings of C-BMRQ were equivalent between the first test and retest (see **Figure S1** for details).

Complementing the longitudinal invariance results, the C-BMRQ total score and all five subscale scores showed strong positive test–retest correlations (ρs = 0.64–0.75, *p*s < .001, Bonferroni-corrected) (Fig. 2), further supporting the temporal stability and reliability of the scale.

Measurement Invariance Across Gender and Music Training Groups

After 1,000-iteration bootstrap, the configural invariance models across genders showed acceptable to good fit (**Table S5**), with CFI ranging from 0.906 to 0.927, TLI ranging from 0.889 to 0.913, and RMSEA ranging from 0.062 to 0.071. Across 1,000 bootstrap iterations, 675 metric-versus-configural comparisons were non-significant, whereas only 76 of these also showed non-significant scalar-versus-metric comparisons; none supported strict invariance. These findings support configural invariance across genders and provide preliminary evidence for metric invariance of the scale.

Across 1,000 bootstrap iterations, the music-training-based configural invariance models also showed acceptable to good fit (**Table S5**), with CFI ranging from 0.895 to 0.928, TLI ranging from 0.875 to 0.915,

and RMSEA ranging from 0.059 to 0.071. Among these iterations, 989 metric-versus-configural comparisons were non-significant, whereas all 989 showed significant deterioration in the scalar-versus-metric and strict-versus-scalar comparisons. These findings support configural invariance across music-training groups and provide strong resampling-based evidence for metric invariance of the scale. Note that metric invariance supports equivalence of factor loadings and permits comparisons of relationships among latent factors, but not latent factor means, which generally require scalar invariance. Because scalar invariance was not established, subsequent observed-score group comparisons were treated as exploratory and descriptive rather than as unbiased estimates of latent mean differences. Bootstrap resampling quantified sampling uncertainty in these contrasts but did not correct for potential bias arising from measurement non-invariance.

Convergent validity with Gold-MSI and PAS.

As shown in Table 3, the C-BMRQ total score and subscales were negatively correlated with PAS ($|\rho s| > 0.33$, Bonferroni-corrected), replicating findings from Mas-Herrero et al. (2013). After distinguishing music-related from other PAS items, C-BMRQ correlated negatively with both subscales. Among the C-BMRQ subscales, Sensory Motor showed the strongest association with music-related PAS items, whereas Social Reward showed the strongest association with non-music PAS items.

Table 3
Spearman Correlations of the Total and Subscale Scores of the C-BMRQ with the PAS and Gold-MSI in Study 2

		Total	SR	MS	EE	MR	SM
PAS (Dataset 1: n = 1019)	Total	-0.47***	-0.46***	-0.37***	-0.33***	-0.33***	-0.42***
	Music	-0.44***	-0.37***	-0.34***	-0.32***	-0.30***	-0.44***
	Other	-0.44***	-0.44***	-0.34***	-0.30***	-0.31***	-0.38***
Gold-MSI (Dataset 2: n = 136)	MT	0.26	0.29*	0.17	0.07	0.24	0.06
	GMS	0.44***	0.45***	0.33**	0.13	0.29*	0.20
Gold-MSI (Dataset 3: n = 110)	AE	0.45***	0.44***	0.54***	0.28	0.25	0.27
	PA	0.41***	0.39***	0.38**	0.36**	0.29	0.30
	MT	0.22	0.28	0.21	0.18	0.13	0.17
	SA	0.31*	0.38**	0.26	0.24	0.24	0.22
	E	0.56***	0.44***	0.45***	0.57***	0.51***	0.40***
	GMS	0.38**	0.42***	0.36**	0.28	0.25	0.28

Note. The optimal Chinese version of the Barcelona Music Reward Questionnaire; PAS = Physical Anhedonia Scale; Gold-MSI = Goldsmiths Musical Sophistication Index. SR = Social Reward; MS = Music Seeking; EE = Emotional Evocation; MR = Mood Regulation; SM = Sensory Motor. MT = Musical Training; GMS = General Musical Sophistication; AE = Active Engagement; PA = Perceptual Abilities; SA = Self-Attribution of Musical Ability; E = Emotions. Correlations were computed separately within each dataset. *$p_{Bonferroni} < .05$, **$p_{Bonferroni} < .01$, ***$p_{Bonferroni} < .001$.

Gold-MSI's Active Engagement was positively correlated with C-BMRQ with ρs ranging from 0.25 to 0.54. Significant correlations were limited to the C-BMRQ total score, Social Reward, and Music Seeking (*p*s < .001, Bonferroni-corrected), with Music Seeking displaying the highest correlation (ρ = 0.54). Perceptual Abilities was also positively correlated with the C-BMRQ with ρs ranging from 0.29 to 0.41; significant associations for the total score, Social Reward, Music Seeking, and Emotion Evocation (*p*s < .01, Bonferroni-corrected). Singing Abilities showed significant positive correlations only with the C-BMRQ total score and Social Reward (*p*s < .05, Bonferroni-corrected). The Emotions subscale showed significant positive correlations with the C-BMRQ total score and all subscales (ρs = 0.40–0.57, *p*s < .001, Bonferroni-corrected), with the strongest association observed for Emotion Evocation.

No consistent association was found between Music Training and the C-BMRQ. Although Dataset 2 showed a weak positive correlation between Music Training and Social Reward (ρ = 0.29, *p* = .021, Bonferroni-corrected), a similar effect in Dataset 3 did not survive correction (ρ = 0.28, *p* = .113, Bonferroni-corrected). General Musical Sophistication was significantly correlated with the C-BMRQ total

score, Social Reward, and Music Seeking in both datasets, with ρs ranging from 0.33 to 0.45 in Dataset 2 and from 0.36 to 0.42 in Dataset 3 (*p*s < .01, Bonferroni-corrected). Social Reward showed the strongest association with General Musical Sophistication in both datasets (Dataset 2: ρ = 0.45, Dataset 3: ρ = 0.42).

Group Differences in Musical Reward Sensitivity of Chinese population

Following Mas-Herrero et al. [7], the 10th percentile of the pooled C-BMRQ total-score distribution was used as a sample-derived reference threshold for low music reward sensitivity; item-level descriptive statistics for the pooled dataset are provided in **Table S3**. This cutoff was 69 in the present Chinese sample (Fig. 3), slightly higher than the cutoff of 65 reported by [7], indicating a potentially higher empirical threshold for identifying low music reward sensitivity in the present sample. Group differences by gender and music training status were examined in subsequent analyses.

Gender group comparisons revealed no consistent differences in the C-BMRQ total score or most subscales, except for Emotion Evocation and Sensory Motor (Fig. 4). For all other indices, the 95% bootstrap intervals of the *t* statistics included zero, indicating unstable and non-significant differences (**Figure S2**). Emotion Evocation yielded a 95% bootstrap interval [-3.85, -0.43] (males < females); across 1,000 bootstrap iterations there were 590, 308, and 106 tests reaching $p < .05$, $p < .01$, and $p < .001$, respectively. Sensory Motor yielded a 95% bootstrap interval [-4.65, -1.21] (males < females); across 1,000 bootstrap iterations there were 855, 633, and 355 bootstrap tests reaching $p < .05$, $p < .01$, and $p < .001$, respectively. The observed gender differences in Emotion Evocation and Sensory Motor are broadly consistent with previous findings using the Chinese version of the BMRQ [19] and international adaptations (e.g., [10]; [13]), suggesting that gender-related differences in these dimensions may be reproducible within this measurement framework.

Music-trained individuals scored significantly higher than non-musicians on the C-BMRQ total score and all subscales (Fig. 5), with all 95% bootstrap intervals of the *t* statistics above zero (**Figure S3**). The total score yielded a 95% bootstrap interval [4.24, 7.97], with 998 of 1,000 bootstrap tests reaching $p < .001$. Across subscales, Social Reward showed the largest difference, 95% bootstrap interval [5.06, 8.82], with all 1,000 tests reaching $p < .001$. Followed by Music Seeking, [3.46, 7.25], with 986 tests reaching $p < .001$. Emotion Evocation showed a 95% bootstrap interval of [2.61, 6.53], with 972 and 892 tests reaching $p < .01$ and $p < .001$, respectively. Mood Regulation showed a 95% bootstrap interval of [2.40, 6.10], with 960 and 858 tests reaching $p < .01$ and $p < .001$, respectively. Sensory Motor showed a 95% bootstrap interval of [1.79, 5.55], with 965, 876, and 858 tests reaching $p < .05$, $p < .01$, and $p < 0.001$, respectively. These results indicate a stable association between music training and higher music reward sensitivity across all C-BMRQ dimensions.

Discussion

The present research completed the development and comprehensive validation of the C-BMRQ. While Wang et al. [19] provided an initial adaptation, our study addresses key methodological gaps: (1) their item-deletion procedure altered the original theoretical structure, and tests of measurement invariance and population-specific reference values were absent; (2) their sample was geographically restricted. To address these, we employed a geographically diverse nationwide sample and a culture-adaptive strategy that preserved the original structure. The core methodological breakthrough of the present study lies in the systematic comparison of four questionnaire versions. Results indicated that the culturally adapted version yielded significantly better model fit than both literal translations and the item-deletion version. Both literal translations exhibited consistently poor model fit, confirming the inherent limitation of pure literal translation in capturing cultural connotations; the clear advantage of the culturally adapted version demonstrates that "linguistic revision integrated with cultural adaptation" is a more rigorous strategy than item deletion [32, 33].

A primary methodological contribution of our study is achieving construct validity and stability without deleting any items. Adhering to the principle that theoretical integrity should take precedence over statistical fit [37], we fully preserved the original five-factor structure. Psychometric research indicates that subscales with fewer than 3 items may face compromised internal consistency and incomplete construct capture [38]. Wang et al.'s [19] item-deletion strategy reduced the Music-seeking and Sensory-Motor dimensions from 4 items to only 2. However, music-seeking and sensory-motor responding are central to the "wanting" component of reward processing [39, 40], and willingness to pay for music is a robust behavioral indicator of reward value [41–43]. Truncating these eliminated the motivational component of music reward. Furthermore, eliminating reverse-coded items undermines the scale's ability to mitigate response biases (e.g., acquiescence bias) and accurately capture individual differences. Therefore, strict literal translation without contextual modification leads to misinterpretation, whereas our refined wording better reflects how Chinese individuals engage with music, providing empirical evidence that cultural adaptation is superior to item deletion.

Beyond the optimized adaptation strategy, the C-BMRQ demonstrated psychometric robustness and cross-group applicability. The scale's test-retest reliability met psychometric standards [34], and it achieved measurement invariance across gender and musical training groups. More importantly, longitudinal and group-difference results confirm that music reward sensitivity possesses a stable structural foundation while exhibiting malleability across time and experience. Specifically, the partial scalar non-invariance observed across test and retest plausibly reflects the inherent malleability of this dynamic affective–motivational trait [10, 31], while the stable underlying factor structure underscores the instrument's integrity. Regarding cross-group differences, the observed gender disparities (females > males in Emotion Evocation and Sensory-Motor) and higher scores among music-trained individuals align with prior findings [10, 13, 19, 31], further supporting the construct validity of the C-BMRQ.

Critically, we verified concurrent and discriminant validity using two theoretically complementary criterion measures. The Gold-MSI assessed objective musical skills, with the C-BMRQ showing significant positive correlations, confirming that higher music reward sensitivity aligns with greater

musical competence. The CPAS measured generalized loss of pleasure capacity, and the C-BMRQ exhibited significant negative correlations, supporting that reduced music reward sensitivity co-occurs with broader anhedonia traits [8, 9, 35, 36]. Together, these dual perspectives confirm the cross-cultural generalizability of the music reward construct and the C-BMRQ's specificity for assessing music anhedonia.

Additionally, from a practical standpoint, the 10th percentile of the pooled C-BMRQ distribution corresponded to a total score of 69. This sample-derived value was higher than the cutoff of 65 reported in previous Western-sample research [10]. This difference highlights potential cross-cultural variation in score distributions and the need to evaluate normative reference values within specific populations [11, 12, 14, 44]. This issue may be particularly relevant given recent evidence of atypical music-reward processing in Chinese clinical and neurodevelopmental samples [45]. The value of 69 may serve as a preliminary research reference for low music reward sensitivity in comparable Chinese adult samples. However, the present study did not include a clinical reference standard or evaluate diagnostic accuracy. Independent studies using clinical interviews, behavioral and physiological measures of music-evoked pleasure, assessments of non-musical reward processing, and appropriate diagnostic-accuracy analyses are required before the threshold can be used for clinical screening or decision-making.

Strengths and limitations

A key strength of this study is the rigorous multi-version comparison, which provides empirical evidence for the superiority of cultural adaptation over item deletion while fully preserving the original 20-item construct coverage. Additionally, the large nationwide sample (N = 1,650*N* = 1,650) and comprehensive psychometric evaluations—particularly the longitudinal and cross-group measurement invariance tests—substantially enhance the robustness and generalizability of the findings.

However, several limitations should be acknowledged. First, the lack of rigorous cross-cultural comparisons with Western samples and the exclusion of specific populations (e.g., low-education, elderly, and clinical groups) limit the generalizability of the conclusions. Second, despite using an independent behavioral sample for criterion validation, the core scale evaluation relied primarily on self-report data. Future research should incorporate objective behavioral tasks (e.g., music preference choice paradigms) or physiological indicators (e.g., skin conductance responses) to enable multi-method cross-validation. Nevertheless, the C-BMRQ remains a reliable, valid, and culturally appropriate tool for assessing music reward sensitivity in the Chinese population.

Conclusions

The optimized C-BMRQ is a psychometrically robust and culturally appropriate instrument that fully retains the original 20-item, five-factor structure. It provides a reliable and valid tool for assessing individual differences in music reward sensitivity and music anhedonia in Chinese-speaking populations. Furthermore, by demonstrating that cultural adaptation can yield better psychometric properties than

item deletion, this study offers a replicable methodological template for the cross-cultural adaptation of instruments without sacrificing construct coverage.

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences (ID: H25086). All participants provided digital informed consent prior to participation.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Number: 32561160135) to Yi Du and the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (Grant Number: E4JY292266) to Yi Du.

Author Contribution

Conceptualization: Y. Du and X.T. Jiang; Funding acquisition: Y. Du; Formal Analysis: S. Hua, X.T. Jiang and Q. He; Methodology: X.T. Jiang, S. Hua, Q. He and Y. Du; Investigation: X.T. Jiang, S. Hua and Q. He; Visualization: X.T. Jiang, S. Hua and Y. Du; Supervision: Y. Du; Writing (original draft): X.T. Jiang and S. Hua; Writing (review and editing): X.T. Jiang, S. Hua, and Y. Du. All authors read and approved the final manuscript.

Acknowledgement

The authors would like to thank Hou Chen for collecting the online questionnaire.

Data Availability

The datasets and materials generated and analyzed during the current study are openly available in the Science Data Bank (ScienceDB) repository at https://doi.org/10.57760/sciencedb.psych.00482.

References

1. Blood AJ, Zatorre RJ. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proc Natl Acad Sci U S A. 2001;98(20):11818-23.
2. Salimpoor VN, Benovoy M, Longo G, Cooperstock JR, Zatorre RJ. The rewarding aspects of music listening are related to degree of emotional arousal. PLoS One. 2009;4(10):e7487.
3. Brattico E, Pearce M. The neuroaesthetics of music. Psychol Aesthet Creat Arts. 2013;7(1):48–61.
4. Zatorre RJ, Salimpoor VN. From perception to pleasure: music and its neural substrates. Proc Natl Acad Sci U S A. 2013;110 Suppl 2:10430-7.
5. Panksepp J. The emotional source of "chills" induced by music. Music Percept. 1995;13:171–207.
6. Salimpoor VN, Benovoy M, Larcher K, Dagher A, Zatorre RJ. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nat Neurosci. 2011;14(2):257 – 62.
7. Mas-Herrero E, Zatorre RJ, Rodriguez-Fornells A, Marco-Pallares J. Dissociation between musical and monetary reward responses in specific musical anhedonia. Curr Biol. 2014;24(6):699–704.
8. Martinez-Molina N, Mas-Herrero E, Rodriguez-Fornells A, Zatorre RJ, Marco-Pallares J. Neural correlates of specific musical anhedonia. Proc Natl Acad Sci U S A. 2016;113(46):E7337-45.
9. Martinez-Molina N, Mas-Herrero E, Rodriguez-Fornells A, Zatorre RJ, Marco-Pallares J. White matter microstructure reflects individual differences in music reward sensitivity. J Neurosci. 2019;39(25):5018-27.
10. Mas-Herrero E, Marco-Pallares J, Lorenzo-Seva U, Zatorre RJ, Rodriguez-Fornells A. Individual differences in music reward experiences. Music Percept. 2013;31(2):118 – 38.
11. Pereira CS, Bortoloti R. Validation of the Brazilian Portuguese Version of the Barcelona Music Reward Questionnaire (BMRQ). Music Sci. 2025;29(4):660-9.
12. Lippolis M, Sorensen SD, Petersen B, Vuust P, Brattico E. How culture and musical engagement shape musical reward sensitivity in Danish teens: a validation study of the Danish Barcelona Musical Reward Questionnaire with 4641 adolescents. Scand J Psychol. 2025;66(2):190–209.
13. Mannino A, Di Maso M, Viganò M, Corbo M, Sozzi M. Italian validation of the Barcelona Music Reward Questionnaire. Music Percept. 2024;41(4):307 – 19.
14. Saliba J, Lorenzo-Seva U, Marco-Pallares J, Tillmann B, Zeitouni A, Lehmann A. French validation of the Barcelona Music Reward Questionnaire. PeerJ. 2016;4:e1760.
15. Honda S, Mas-Herrero E, Isoda M, Muraki M, Lorenzo-Seva U, Kitayama Y, et al. The Japanese version of the Barcelona music reward questionnaire (J-BMRQ) confirms the cross-cultural generalizability of the "five-factor" model. PLoS One. 2026;21(2):e0340700.
16. Wang J, Xu M, Jin Z, Xia L, Lian Q, Huyang S, et al. The Chinese version of the Barcelona Music Reward Questionnaire (BMRQ): Associations with personality traits and gender. Music Sci.

2021;27(1):218 – 32.
17. Belfi AM, Bai E, Stroud A, Twohy R, Beadle JN. Investigating the role of involuntary retrieval in music-evoked autobiographical memories. Conscious Cogn. 2022;100:103305.
18. Fuentes-Sánchez N, Espino-Payá A, Prantner S, Sabatinelli D, Pastor MC, Junghöfer M. On joy and sorrow: Neuroimaging meta-analyses of music-induced emotion. Imaging Neurosci. 2025;3.
19. Wang J, Xu M, Jin Z, Xia L, Lian Q, Huyang S, et al. The Chinese version of the Barcelona Music Reward Questionnaire (BMRQ): Associations with personality traits and gender. Music Sci. 2023;27(1):218 – 32.
20. Matthews TE, Lumaca M, Witek MAG, Penhune VB, Vuust P. Music reward sensitivity is associated with greater information transfer capacity within dorsal and motor white matter networks in musicians. Brain Struct Funct. 2024;229(9):2299 – 313.
21. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539 – 49.
22. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21(4):651-7.
23. International Test Commission. International Guidelines on Test Adaptation. https://www.intestcom.org. Accessed 26 Oct 2026.
24. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. J Stat Softw. 2012;48(2):1–36.
25. Chan RCK, Wang Y, Yan C, Zhao Q, McGrath J, Hsi X, et al. A study of trait anhedonia in non-clinical Chinese samples: evidence from the Chapman Scales for Physical and Social Anhedonia. PLoS One. 2012;7(4):e34275.
26. Chapman LJ, Chapman JP, Raulin ML. Scales for physical and social anhedonia. J Abnorm Psychol. 1976;85(4):374 – 82.
27. Müllensiefen D, Gingras B, Musil J, Stewart L. The musicality of non-musicians: an index for assessing musical sophistication in the general population. PLoS One. 2014;9(2):e89642.
28. Brown TA. Confirmatory factor analysis for applied research. 2nd ed. New York, NY: The Guilford Press; 2015.
29. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling. 2002;9(2):233 – 55.
30. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct Equ Modeling. 2007;14(3):464–504.
31. Mas-Herrero E, Zatorre RJ, Marco-Pallarés J. Understanding individual differences to specific rewards through music. Trends Cogn Sci. 2026;30(2):175 – 86.
32. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public

Health. 2018;6:149.
33. Maneesriwongul W, Dixon JK. Instrument translation process: a methods review. J Adv Nurs. 2004;48(2):175 – 86.
34. ten Berge JMF. J. C. Nunnally and I. H. Bernstein. Psychometric Theory. New York: McGraw-Hill, 1994, xxiv + 752 pages, £51.95. Psychometrika. 1995;60(2):313-5.
35. Kathios N, Patel AD, Loui P. Musical anhedonia, timbre, and the rewards of music listening. Cognition. 2024;243:105672.
36. Belfi AM, Loui P. Musical anhedonia and rewards of music listening: current advances and a proposed model. Ann N Y Acad Sci. 2020;1464(1):99–114.
37. DeVellis RF. Scale development: Theory and applications. Thousand Oaks, CA: Sage Publications, Inc; 1991.
38. Worthington RL, Whittaker TA. Scale development research: a content analysis and recommendations for best practices. Couns Psychol. 2006;34(6):806 – 38.
39. Berridge KC, Kringelbach ML. Affective neuroscience of pleasure: reward in humans and animals. Psychopharmacology (Berl). 2008;199(3):457 – 80.
40. Berridge KC, Kringelbach ML. Pleasure systems in the brain. Neuron. 2015;86(3):646 – 64.
41. Mas-Herrero E, Dagher A, Farres-Franch M, Zatorre RJ. Unraveling the temporal dynamics of reward signals in music-induced pleasure with TMS. J Neurosci. 2021;41(17):3889-99.
42. Mas-Herrero E, Dagher A, Zatorre RJ. Modulating musical reward sensitivity up and down with transcranial magnetic stimulation. Nat Hum Behav. 2018;2(1):27–32.
43. Salimpoor VN, van den Bosch I, Kovacevic N, McIntosh AR, Dagher A, Zatorre RJ. Interactions between the nucleus accumbens and auditory cortices predict music reward value. Science. 2013;340(6129):216-9.
44. Carraturo G, Ferreri L, Cardona G, Lorenzo-Seva U, Rodriguez-Fornells A, Brattico E. The Italian version of the extended Barcelona Music Reward Questionnaire (eBMRQ): a validation study and association with age, gender, and musicianship. PeerJ. 2025;13:e18403.
45. Jin Z, Huyang S, Jiang L, Lian Q, Wu D. Music reward sensitivity associated with frontostriatal function and structure in congenital amusia. Ann N Y Acad Sci. 2025;1554(1):176 – 86.

Figures

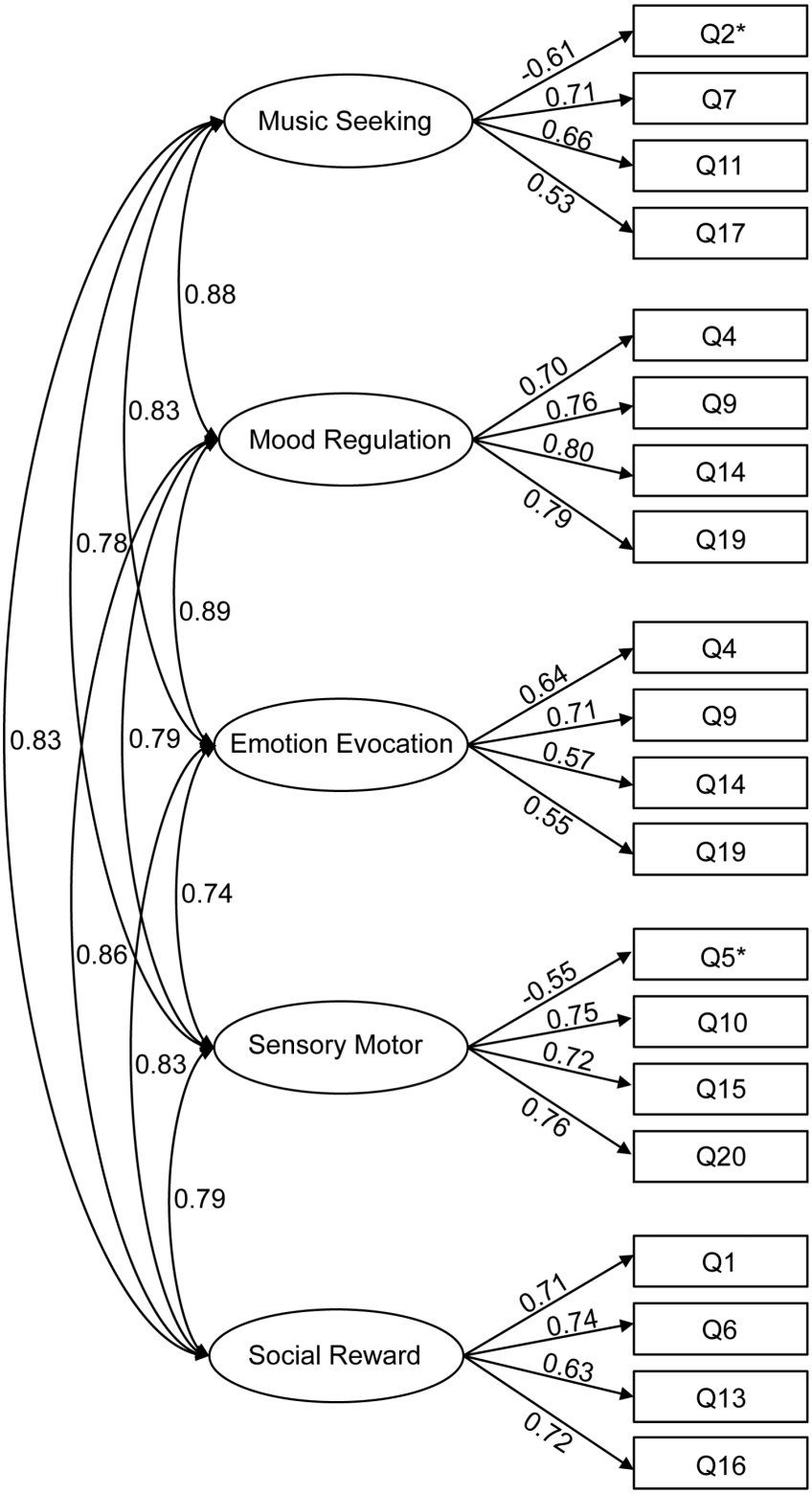


Figure 1

Factor loadings of C-BMRQ in CFA (Study 2 Dataset 1; N = 1019). *Reverse Item.

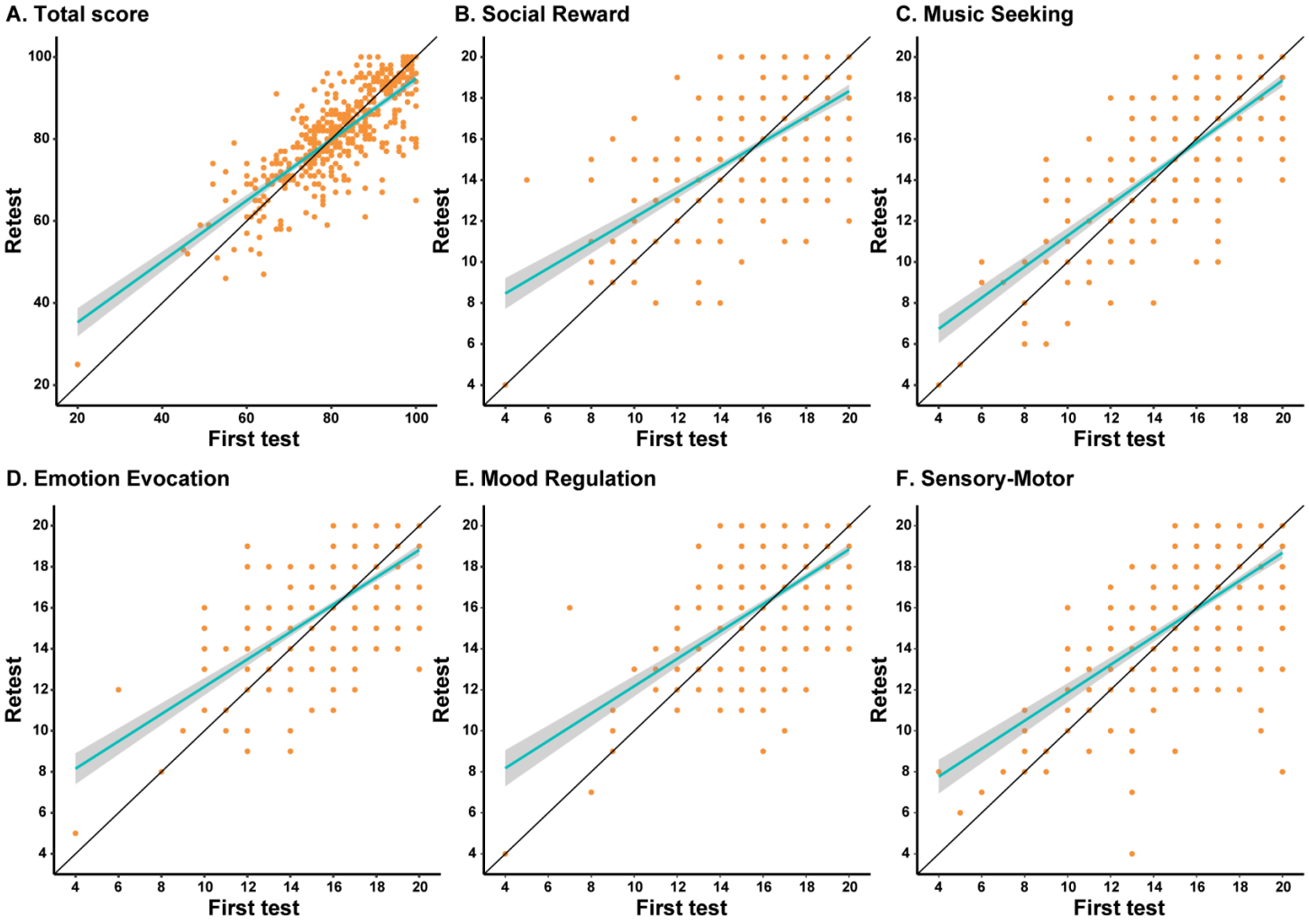


Figure 2

Test-retest reliability of the C-BMRQ. Scatter plots depicting first-test versus three-month retest scores for (A) Total score of C-BMRQ, and subscales (B) Social Reward, (C) Music Seeking, (D) Emotion Evocation, (E) Mood Regulation, (F) Sensory Motor. Spearman correlations with Bonferroni correction. ***$p_{Bonferroni}$ < .001. The cyan line represents the regression line; the shaded gray band indicates the 95% confidence interval around the regression estimate. The black solid line denotes the line of perfect agreement between first-test and retest scores.

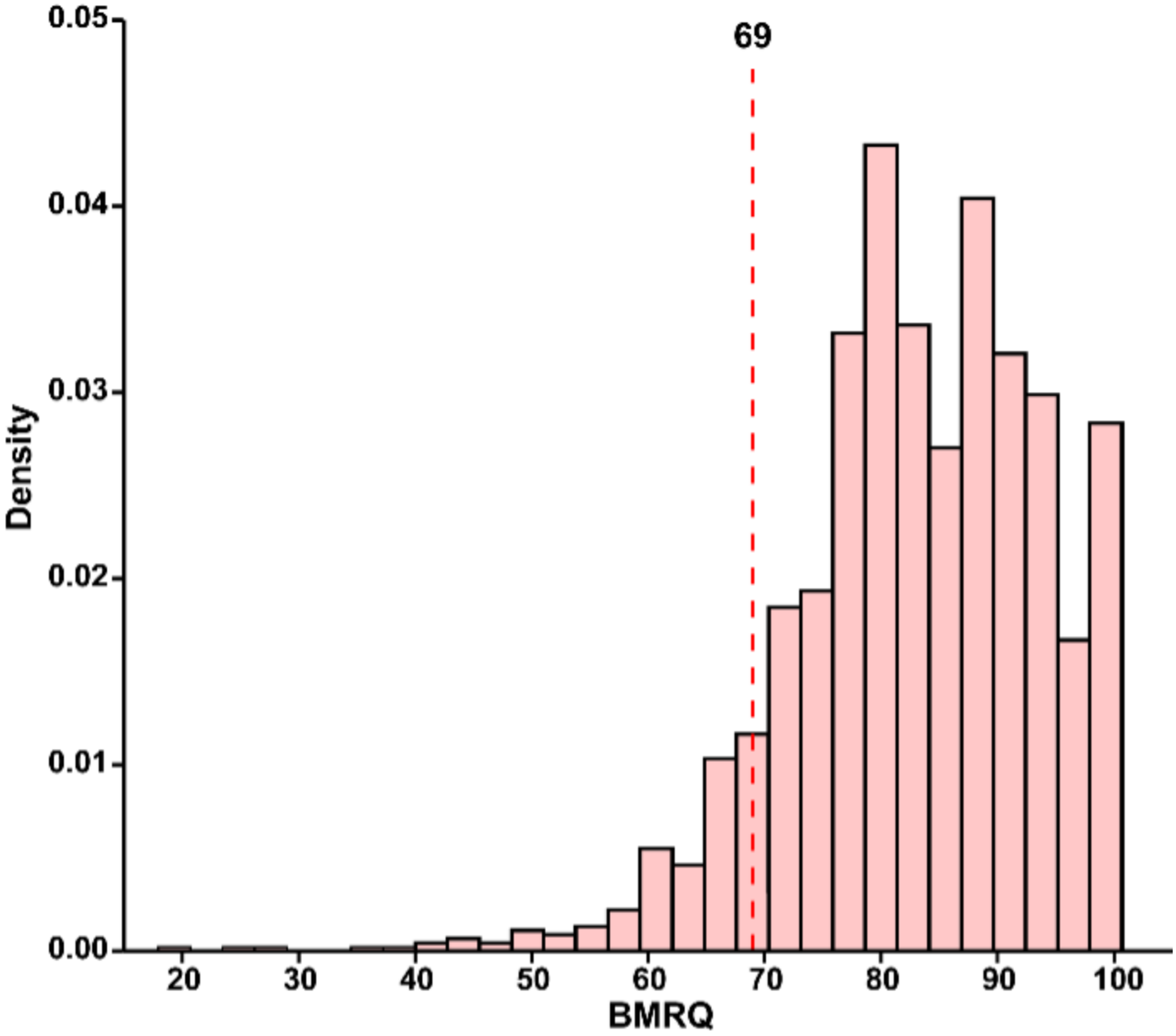


Figure 3

Global music-reward-sensitivity: total-score distribution of the C-BMRQ in Chinese adult participants. The red dashed vertical line marks the 10th percentile threshold for total C-BMRQ scores (value = 69), which was used as a sample-derived reference threshold for low music reward sensitivity. The y-axis denotes density; the x-axis represents the total C-BMRQ scores.

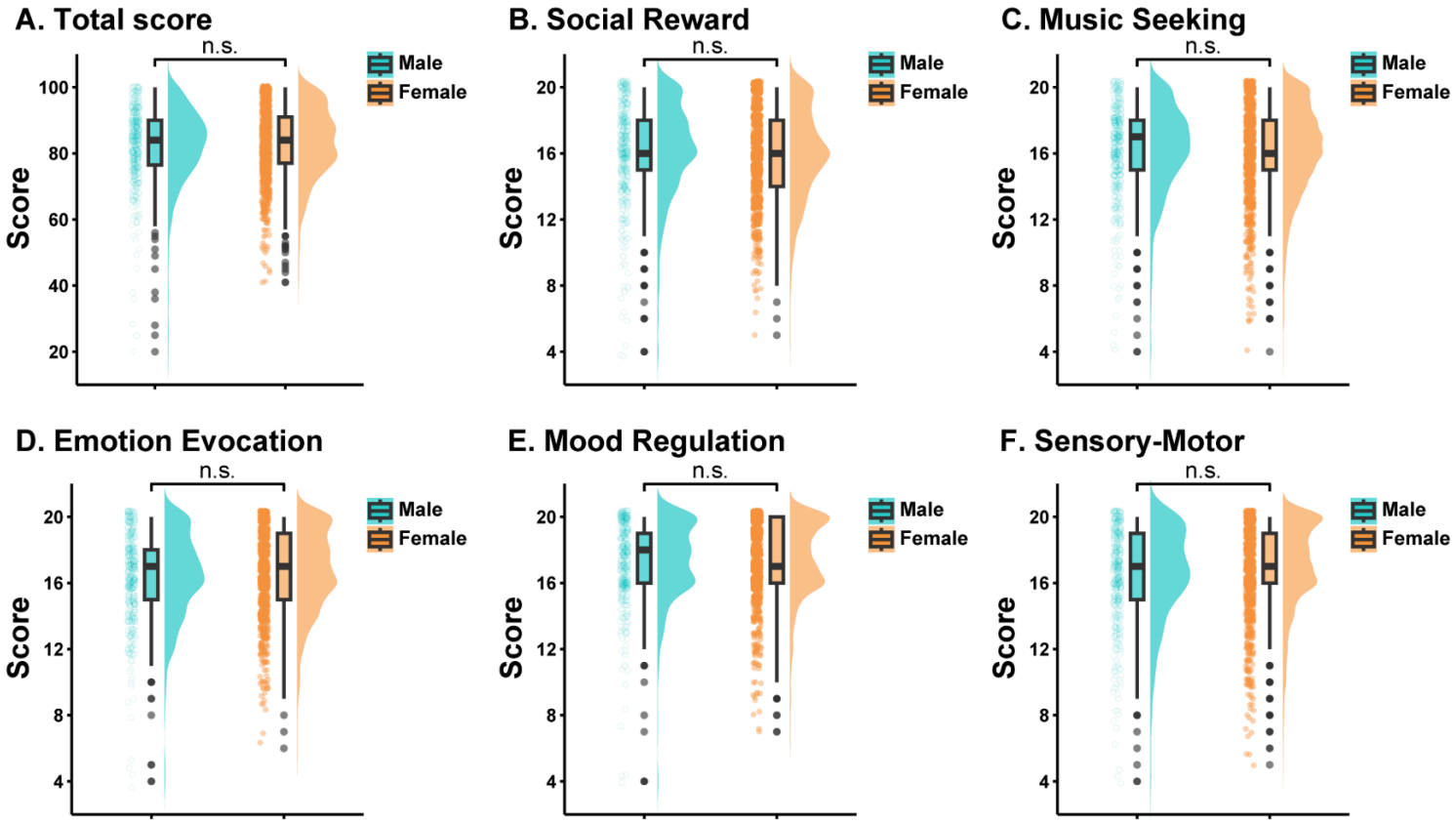


Figure 4

Gender comparisons on the C-BMRQ total and subscale scores in the Chinese sample. Violin plots comparing male and female participants on the C-BMRQ (A) total score and subscales (B) Social Reward, (C)Music Seeking, (D) Emotion Evocation, (E) Mood Regulation, (F) Sensory Motor. n.s. = non-significant. 95% bootstrap confidence intervals (1,000 iterations) for observed t-values are displayed.

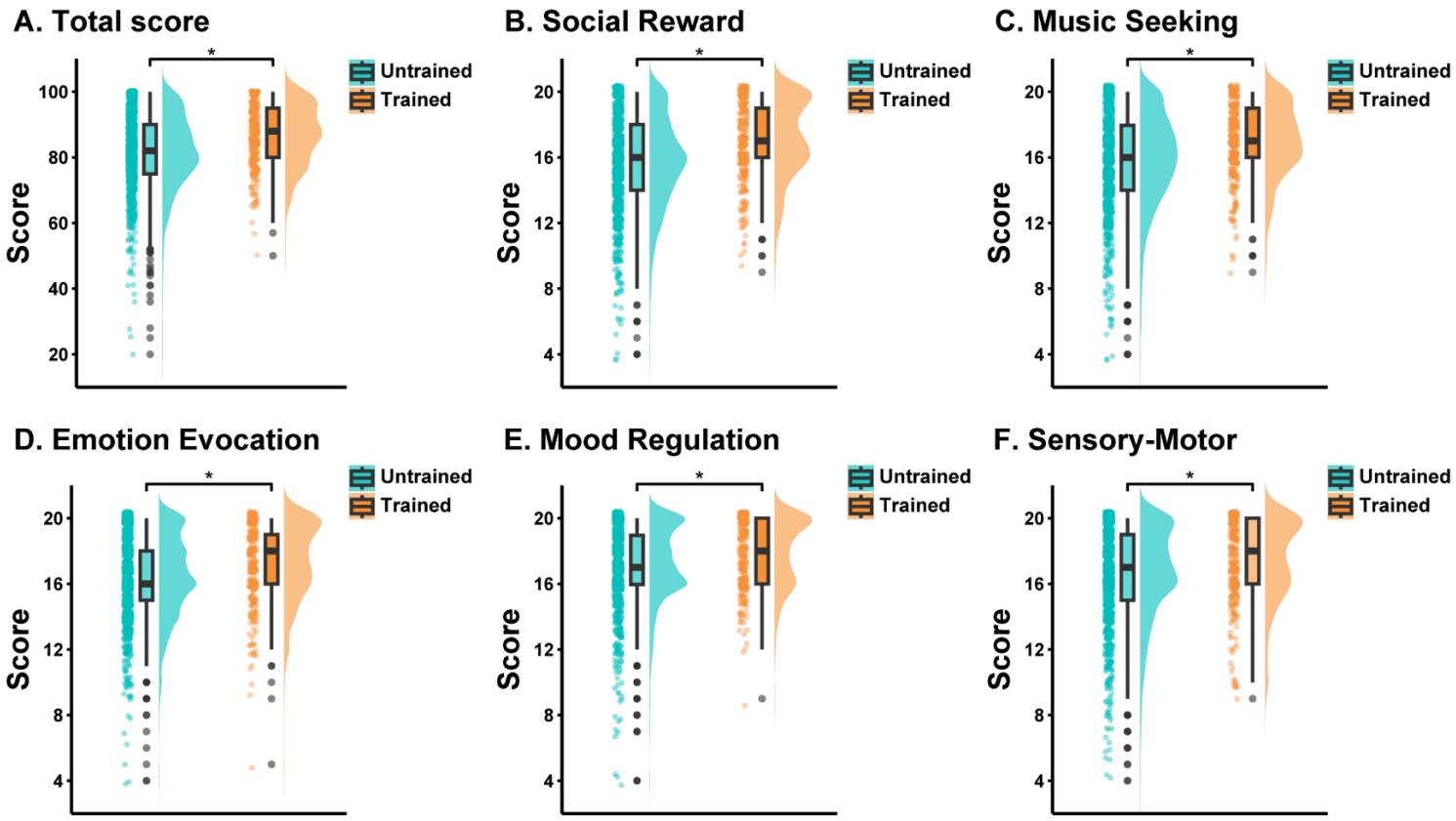


Figure 5

Comparisons of C-BMRQ total and subscale scores between musically trained and untrained participants in the Chinese sample. Violin plots comparing C-BMRQ scores for musically untrained versus trained participants on the C-BMRQ (A) total score, and subscales (B) Social Reward, (C)Music Seeking, (D) Emotion Evocation, (E) Mood Regulation, (F) Sensory Motor. * indicates a statistically significant group difference derived from the 1000-iteration bootstrap permutation test.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementalmaterialsfinalversion.docx