Pakistan Journal of Psychological Research, 2021, Vol. 36, No. 4, 645-663

https://doi.org/10.33824/PJPR.2021.36.4.36

Adaptation and Validation of Short Test of Music Preference Scale for Students With Music and Non-Music Majors

Muhammad Faran and Farah Malik

University of the Punjab

Music is a universal phenomenon however, despite its unified properties, the taste and preference of music may still vary as a function of ethnicity and culture. So, the present study aimed to adapt and validate the short test of music preference scale for music and non-music Pakistani students. In Phase I, the cultural adaption of the scale was carried out while the content validity index (Lawshe, 1975) was also established. However, in phase II, the Short test of Music Scale (STOMP) was validated, yielding confirmatory factor analysis. For the empirical evaluation, a sample of 561 students, including both 286 music and 275 non-music students of undergraduate level with the age range of 18-26 years were recruited. The psychometric evolution of STOMP turned into excellent validity and reliability estimates for first-order constructs. Moreover, strict measurement invariance was established for STOMP across music and non-music students. The validation of this scale would be a little effort to pave the way for music psychology to make research available to measure the construct indigenously.

Keywords. Adaption, validation, music preferences scale, music and non-music, students

Every person has a different preference for genres of music. Everyone has their own likings and disliking in music. Through this liking and disliking, music preferences are developed, which are defined by Rentfrow and Gosling (2003) as the definition of music preferences is fairly self-explanatory, it refers extent to which a person prefers or likes a particular kind of music over another. Because music preference is usually seen as a long-term effective evaluation, it is usually described synonymously with musical taste, which has

Muhammad Faran and Farah Malik, Institute of Applied Psychology, University of the Punjab, Lahore, Pakistan.

Correspondence concerning this article should be addressed to Muhammad Faran, Institute of Applied Psychology, University of the Punjab, Lahore, Pakistan. Email: mfaranakbar@gmail.com

traditionally been defined as a person's overall attitude towards a collective musical phenomenon (Julin & Sloboda, 2011).

There has been a voluminous existing literature exploring various aspects of musical preference and taste, only two model of music preference (Hargreaves, North, & Tarrant, 2006; LeBlanc 1982) explicitly bounds the phenomenon together. LeBlanc's (1982) interactive theory of music preference is the approach which is useful in formally identifying a large number of types of variables that fall into different broad categories. For example, the listener's cultural environment includes the variables of media, peer group, family, educators and authority figures, and incidental conditioning. While the listener characteristics such as attention and mental processing through to a preference decision at a given moment, which then influences subsequent behaviour. LeBlanc (1982) argued that the feature of music accounted for a great disparity in children's uttered musical preferences, while culture and age are also accounted.

Moreover, Hargreaves et al. (2006) proposed a model which is based on four very simpler reciprocal feedback constructs of musical responses which comprised of the interaction among the four major variables that is, music, listener, situation and context to elicit a desired response. Musical features includes a reference system (genre, preferences), collative variables (complexity and familiarity) and prototypically. Secondly, the characteristics of the listeners in term of individual difference which include, gender, age, personality; while the knowledge of the music, preference and style, and identity are also are accounted. Thirdly, the reaction of the listener in response to music is also influenced by the physiological, cognitive and affective factors. Lastly, the situation and context which refer to socio-cultural context, day to day situation or circumstances and accompany of others.

Both model shared a tripartite division between music, listener and the context as well the interaction between the large internal and external factors. However the culture should not be accounted as a factor but rather considered as a mediator through which all real-life experiences interpreted (Cole, 1998; Lamont, 2006). Similarly, Cattell and Anderson (1953) argued that music preference is an unconscious process and reflects the inner personality types. Thus, after this piece of information has been revealed, the researchers have started taking more interest into finding the linkage between personality and music preferences. Likewise, other researchers, Little and Zuckerman (1986) conducted a study and they found that there is a positive relationship between music preference specially rock, heavy metal and punk music and sensation seeking. However, with personality traits like

extraversion and psychoticism; rap and dance music is preferred (McCown, Keiser, Mulhearn, & Williamson, 1997). Rentfrow and Gosling (2003) grouped 14 genres into four factors which were labeled as reflective and complex, intense and rebellious, upbeat and conventional and energetic and rhythmic music preference. They argued that reflective and complex music preferences comprised of blues, jazz, classical, and ethnic music genres. This particular factor accounted by refined skills and knowledge that is rarely accessible from different resources such as, through the education system or through the excessive use mass media. From this point of view, these preferred choices construct presume certain qualities (Rentfrow & Gosling, 2003); whereas intense and rebellious music includes punk, trance, funk, alternative and hard rock/heavy metal. The types included in this construct have low apparency in the media. People who prefer this type of music also signify some distance from mainstream preferences. Participants attaining high scores in this construct are conscious about the fact that their choice is inharmonious from other wide range of preferences (Rentfrow & Gosling, 2003). While upbeat and conventional music preferences includes rock, rock 'n' roll, Greek rock, country, religious, and art-popular is associated to stereotypes of a deviating musical culture. In the Greek culture, rock music is a socially critical type that is related to cultural lifestyles (Rentfrow & Gosling, 2003). Moreover, energetic and rhythmic included pop, soul, rap, hip-hop, Greek pop, Greek rap, Laika pop etc. it mainly had pop music in common, and it is easy to learn, typically focused on gender relations and emotional needs. The mainstream music industry uses rap as it helps create sensations, sentimental lyrics, and different non-verbal elements (Gardikiotis & Baltzis, 2012). Rentfrow, Goldberg, and Lovitin (2011) also proposed the five-factor structure of music preference which provides a better fit for the data. The factors included mellow, unpretentious, sophistication, intense and contemporary. However, Gosling (2003) with his colleagues, keeps on adding new genres to develop a more comprehensive measure to gauge music preferences, for example, STOMP-R, which is comprised of 23 genres of music. However, the genres included of extended so far basically belong to the western continent. There is a dire need to adapt the STOMP to a more valid and reliable scale to measure music preferences in Asia, particularly in Pakistan.

Rationale of the Study

Since, research within this field may seem quite far-reaching, numerous technical limitations exist in the current body of knowledge.

Only a few studies have focused on the non-Western cultures (Chamorro-Premuzic et al., 2009; Rana & North, 2007; Upadhyay, Shukla, & Chakraborty, 2016). Generally, it is considered that music is a universal phenomenon. However, the structure of musical preferences is quite culturally bound and as a function of ethnicity (Eerola, Himberg, Toiviainen, & Louhivuori, 2006; Gans, 1974; Gregory & Varney, 1996; Saarikallio, 2008). Researches which have been conducted indigenously so far focused on measuring only western genres as the preferences of music. However, the subcontinent culture, especially Pakistan, is predominantly classical music listeners and performers due to cultural roots. At the same time, the associated genres in classical music, including semi-classical and *qawali*, were never considered while gauging the music preferences (Mehboob, 2010; Naz, 2008). Hence, the present study adapts the music preference scale by including the indigenous genres and validates it to measure the structure of music preferences more accurately and reliably for the Pakistani population.

Method

The current study is comprised of two phases; in phase I adaptation of the scale was carried out while in phase II validation of the STOMP was also taken into account to determine the psychometric properties including the measurement invariance.

Phase I: Adaptation of STOMP

New music genres were added to the existing structure of musical preferences to adapt to the test, which is listening in Pakistan. The iterative review process by the research team and study's expert panel was used until consensus was reached about the structure of music preferences. The expert panel of key stakeholders included two music teachers, four professional music artists, and two music composers. This ended up in a final scale comprised of 16 musical styles. Moreover, the content of the finalized measure (by the expert panel meeting) was validated by another panel of experts, adapting the criteria established by (Lawshe, 1975). In response options, along with the original scale, the option 'essential,' 'useful, but not essential,' or 'not necessary' was included. The content validity index was .99 for the six experts (including two psychologists, two music teachers, two music artists, and two music composers) (Lawshe, 1975), which indicated that all items were relevant to the content regarding the Pakistani cultural context.

Translation procedure. The short test of music preferences was available in English. This translation aimed to generate a translation of the English version of the test into the Urdu language that is conceptualized to the original and can easily be understood by participants. The steps followed in translating a short test of music are as follows:

Forward translation. This process aimed to obtain a translation of the measure in a target language (Urdu) that was conceptually equivalent to the original scale and easily understandable for the respondents. Two forward translations were obtained from two bilinguals (English & Urdu) who were native speakers of the target language and fluent in the source language. The purpose was to obtain a consensus target language version. The consensus was developed in a meeting with the researchers between the two forward translations, and that translation was given preference, which completed the meaning behind the items in English and on which there was mutual consent. The translators made every effort to stay as close to the literal meaning behind the item as they could. In that way, a final version of the Urdu translation was completed.

Backward translation. The purpose of backward translation was to obtain a translation into English (source language) of the target language version (Urdu). For this purpose, two backward translations were done. The Urdu translated version of the short test of the music preference was given to two translators who were proficient in both English and Urdu and had not been part of the forward translation procedure. Then researchers developed a consensus to make a final version of both English translations for comparing it with the original one. After getting a final version of backward translation, it was then compared with the original version short test of music preferences (STOMP).

Proofreading. Proofreading aimed to ensure that no typing spelling or grammatical mistakes remain in the target language version. The translators made every effort to stay as close to the literal meaning behind the item as they could. In that way, a final version of the Urdu translation was completed. Certain discrepancies regarding the translation were noted, and the research team advised changes in the proofreading process.

Measure

Short Test of Musical Preferences (STOMP) developed by (Rentfrow & Gosling, 2003) included 14-basic genres and measuring

four distinctive preferences of music that is, reflective and complex music included classical and blues genres. While intense and rebellious music comprised of alternative and rock music genres. Upbeat and Conventional consisted of country and religious music, and energetic and rhythmic accounted rap/hip-hop and soul/funk genres of music. All genres on the STOMP scale was prefaced with, "*I like*"; whereas 7-point Likert type response scale was used where 1 = *not at all* and 7 = *very much*. The Cronbach's alpha were found to be .78, .85, .79, and .76 for reflective and complex, intense and rebellious, upbeat and conventional, and energetic and rhythmic; respectively. However, after the adaptation process two more genres that is, semi-classical and *qawali* were added in the current measure. The high scores on sub-factors of the measure represent high preference for that certain factor.

Try Out

The purpose of try out was to check the understanding and comprehensibility of the adapted version of the scale. For this purpose sample consisted of 30 adults with the age ranged between 18 to 26 years. These individuals were both men and women. Gender was equally distributed. Convenient sampling strategy was used. Participants were approached personally either at their departments and institutes. They were briefed about the purpose of the study. Participants were instructed to underline any ambiguity if they found during the response. It was observed that most of the participants were not familiar with certain genres. However, for their understanding and ease, famous artist and songs of each particular genres were mentioned. The participants did not report any major problem then. They understood the conceptual meaning of each item and no ambiguities were found. The Cronbach's alphas of the adapted version were ranging from .78 to .81 of all factors.

Phase II: Determining the Psychometric Properties of the Scale

To determine the psychometric properties of the scale, confirmatory factor analysis was conducted on 16 items with 7 point Likert scale to validate the factor structure of the scale. Structural equation modeling (SEM) through AMOS (Analysis of Moment Structure) version 24.0 was employed to validate the factor structure of the short test of music preference scale. STOMP scale consisted of four sub-factors, labeled as reflective and complex, intense and rebellious, upbeat and conventional, and energetic and rhythmic.

Sample

It is considered that the greater the size of sample is better to validate a measure. So, the sample size was drawn according to the minimum criteria of 10:1 (10 cases per parameter) (Haier et al., 2010). However, for the empirical evaluation 561 students including both 286 music and 275 non-music students of graduate or post graduate level. These 561 students were 281 men and 280 women with age range of 18-26 years (M = 21.70, SD = 2.54). The sample was purposely drawn from different institutes of Punjab.

Inclusion criteria (students with music major). Participants who were formal students of music enrolled in the regular degree program. Participants who enrolled in the music program for at least six months were included. Only those participants were included who used to engage in music for at least two hours a day and only those participants were taken who used to engage with music at least six days a week.

Inclusion and exclusion criteria (students with non-music major). Only those participants were included who had a habit of listening to music. Only those participants were taken part that use to listen to music, at least half an hour a day. Only those participants were taken part that use to listen to music, at least four days a week.

Exclusion criteria. Participants who were listening to music lesser than six months were excluded. Participants with any physical impairment were excluded.

Procedure

Permission to use the scale to validate and translate into the Urdu language was taken from their authors. The letter was authenticated the researcher's identities and the topic under investigation. Music and non-music students were selected through purposive sampling. After explaining the nature and purpose of the research, individual consent was taken from the study participants. All the queries raised while responding to the scale by the participants were answered by the researcher. Then the demographic information sheet and Short Test of Musical Preferences (STOMP) were administered to participants. Confidentiality and anonymity of the data were ensured.

Results

Table 1

Fit Indices of Confirmatory Factor Analysis for the Short Test of Music Preferences Scale

Model	χ^2	*df*	χ^2/df	GFI	CFI	NNFI	RMSEA	SRMR
Initial Mode	813.88	196	4.15	.83	.88	.87	.07	.08
Model Fit	513.70	192	2.68	.91	94	.93	.05	.06

Note. N = 600, All changes in chi square values are computed relative to model, χ^2 >.05, GFI = Goodness of Fit Index, CFI = Comparative Fit Index, NNFI = Non-Normed Fit Index; RMSEA = Root Mean Square Error of Approximation, SRMR = Standardized Root Mean Square.

Table 1 indicted the fit indices of the model for short test of music preferences scale. The absolute fit of STOMP were, χ^2 (192) = 513.70, $p < .05$. The initial investigation of model fit indices included, absolute and relative fits provided an indication of poor model fit. Hair et al. (2010) argued that the chi-square test is sensitive to the size of the sample, number of parameters that are accounted in a model and non-normality of the distribution.

So, the investigators recommend the several relative fit indices to determine the model fit Hence, the indices of relative fit of the model including CFI, NFI, GFI, RMSEA, SRMR were consulted. Hu and Bentler (1999) recommended the criteria of relatives indices as χ^2/df fall under the range zero and three; while, RMSEA and SRMR indices would be .08 or lesser and CFI, NNFI and GFI would be .90 or higher. Since, the RMSEA and SRMR for the initial model were ended up with the values of .07 and .08; whereas the GFI, CFI, and NNFI values for STOMP were .83, .88, and .87; respectively which indicted the poor model fit according to the standard criteria of relative model fit indices.

Hence, procedure of model modification was initiated. The modification of the model was determined in a single step in order to attain the model fit as per the criteria. A covariance was drawn as suggested by the modification indices. Modification indices suggested to draw a covariance across the error terms of the indicators of the latent factors of the STOMP. The items from each respective factors were similar in term of content and context (Kenny, 2011). Moreover, Tomás and Oliver (1999) have argued that, in a survey based research covariance between the error terms of indicators of the latent factors can be legitimately drawn.

Arbuckle (2012) suggested the criteria of adding any parameter in model to improve the model fit. So, only those covariances were

added in a model which changes the chi square with 4.0 or above. After drawing the covariance across the error terms, the indices of absolute and relative fit (GFI, CFI, NNFI, RMSEA, and SRMR) were compared once again. Finally, the values of RMSEA and SRMR after modification in the model were found to be .05 and .06; respectively. In addition, the GFI, CFI and NNFI indices were .91 .94, and .93; respectively. So, the model was best fit to estimate the further investigation of the measurement model.

Music Students Non-Music Students

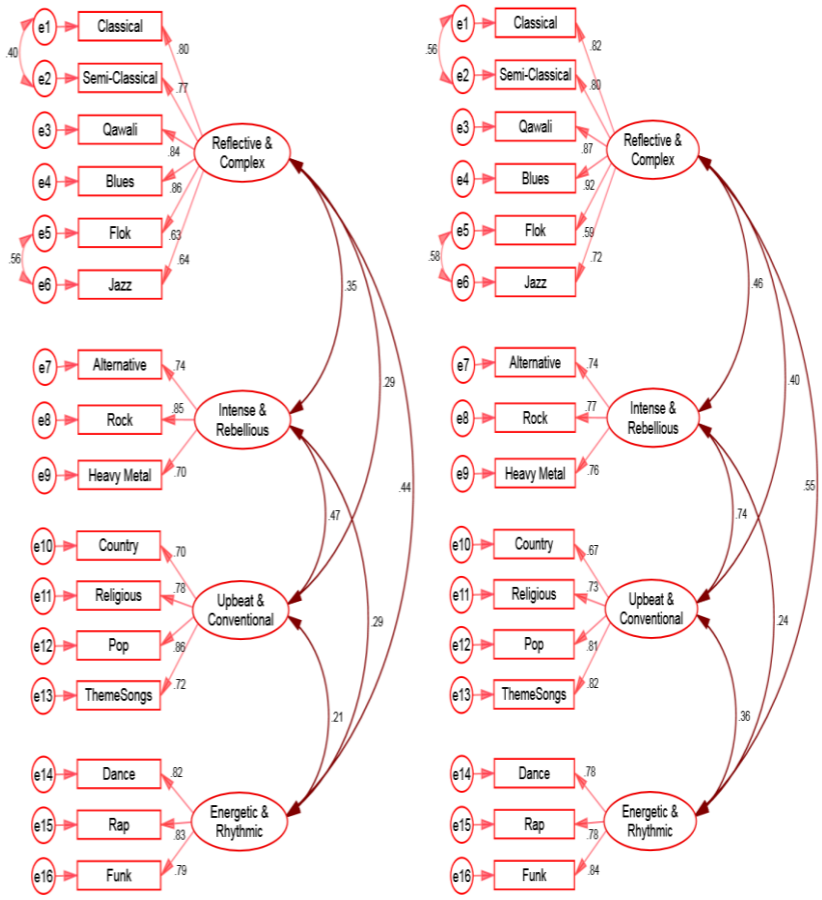

Figure 1. First Order Confirmatory Factor Analysis of STOMP.

The first basic question of this study was to investigate the students' social competency status in general. Accordingly, the result is summarized and presented in Table 1.

Table 2

First order CFA for Short Test of Music Preferences Scale

Factors	Music (n = 286)				Non-Music (n = 275)			
	ω	AVE	MSV	λ	ω	AVE	MSV	λ
Reflective and Complex	.89	.63	.30		.88	.58	.19	
Classical				.82				.80
Semi-Classical				.80				.77
Qawali				.87				.84
Blues				.92				.86
Folk				.59				.63
Jazz				.72				.64
Intense and Religious	.79	.60	.55		.80	.59	.23	
Alternative				.74				.74
Rock				.77				.85
Heavy Metal				.76				.70
Upbeat and Conventional	.85	.58	.55		.85	.59	.26	
Country				.67				.70
Religious				.73				.78
Pop				.81				.86
Theme Songs				.82				.72
Energetic and Rhythmic	.84	.64	.16		.86	.66	.17	
Dance				.78				.82
Rap				.78				.83
Funk				.84				.79

Note. ω = Mcdonald's Reliability; AVE = Average Variance Extracted; MSV = Maximum Shared Variance; λ = Standardized Factor Loading

Psychometric properties of the short test of music preferences scale were determined through CFA to establish the reliability, convergent validity, and discriminant validity of the factors. Table 2 shows McDonald's omega reliability and AVE values were above their cutoff values of .70 and .50; respectively (Henseler, Hubona, & Ray, 2016).

Factor loadings of the items for each respective constructs were examined to determine the convergent validity. The percentage of AVE of each factors that is reflective and complex, intense and rebellious, upbeat and conventional and energetic and rhythmic for music and non-music students were in acceptable range; whereas the McDonald's omega reliability was ranging from .79 to 89 for music students and .80 to 88 for non-music students.

Table 3

Descriptive Statistics and Fornell-Larcker Criterion for the Factors of Music Preference Scale

Factors	*K*	Music (n = 286) *M (SD)*	Non-Music (n = 275) *M (SD)*	Intense and Rebellious	Reflective and Complex	Upbeat and Conventional	Energetic and Rhythmic	
Intense & Rebellious	3	15.27 (4.3	14.38 (4.51)	**0.75** 0.77	–			
Reflective & Complex	6	27.53 (9.6	23.96 (9.59)	**0.46** 0.35	**0.80** 0.76	–		
Upbeat & Conventional	4	20.68 (5.5	18.89 (5.79)	**0.74** 0.47	**0.41** 0.29	**0.76** 0.77	–	
Energetic & Rhythmic	3	15.81 (4.0	13.36 (4.96)	**0.24** 0.29	**0.55** 0.44	**0.36** 0.20	**0.80** 0.82	–

Note. k = Number of Items; Bold Figures = Results of Students From Music Group; Unbold Figures = Results of Students From Non-Music Group.

However, two different criteria were followed to establish the discriminant validity (Henseler, Hubona, & Ray, 2016; Voorhees et al., 2016). First, the square root of average variance extracted AVE values for each scale was greater than the construct's respective correlation with all other factors (Fornell & Larcker, 1981) (see Table 3). Secondly, the average variance of a factor should be greater than the variance which it shared with all other factors, means average variance (AVE) extracted should be greater than maximum shared variance (MSV) (Haire et al., 2010). So, the evidence revealed that MSV of all respective factors were less than the AVE.

Measurement Invariance for the Short Test of Music Preferences Scale

Measurement invariance test was also applied to assess the generalizability of the measure across two different populations i.e., music and non-music. The purpose of measurement invariance was to assess that whether an instrument measure is interpreted conceptually or contextually in a similar manner by participants across different groups (Byrne & van de Vijver, 2010). Byrne, Stewart, Kennard, and Lee, (2007) argued that whenever mean scores of the latent variable are intended to compare across two or more different groups, it is always important determine at least strong or partial measurement invariance of the measurement device. Nevertheless, if the homogeneity or invariance of an assessment measure is failed to establish, the validity of the instrument and conceptualization extracted from the data is may be anomalous (Byrne, 2008), while the results based on differences across groups cannot be valid.

Measurement invariance of a measure is the degree to which parameters encompassing the measurement are homogeneous across different groups (Byrne, 2008) and it is determined at three stages, at the first stage weak (invariance of factor loadings also called matric invariance including configural invariance) is established. While at the second stage, strong invariance (factor loadings of the items and intercepts invariance i.e., matric and scalar invariance) is determined while at last stage, strict invariance (invariance of factor loadings, mean intercepts, factor covariance and error variance invariance) is established. Invariance of the measures can also be categorized in full and partial measurement invariance where full Invariant model included (configural, matric, scalar, factor variance-covariance and error variance invariance); whereas partially invariant model accounted for configural, matric and scalar invariance across groups (Hair, 2010). Evaluation of the measurement invariance involves a

series of sequential set of steps of nested models that typically begun with the establishment of a well-fitting baseline model (unconstrained configural model) (see Table 4).

Table 4

Testing for the Strict Measurement Invariance of STOMP Scale

Model	χ^2	*df*	$\Delta\chi^2$	Δdf	CFI	ΔCFI	RMSEA
Unconstrained configural	400.48	96	-		.95	-	.07
Configural invariance	513.70	192	113.22	96	.94	.009	.05
Matric invariance	537.21	208	23.51	16	.93	.005	.05
Scalar invariance	560.58	224	23.37	16	.92	.008	.05
Factor covariance invariance	574.02	232	13.44	8	.92	.005	.06
Error variance invariance	597.16	248	23.14	16	.91	.008	.06

Note. CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square; $\Delta\chi^2$ = Chi Square Change; *df* = Degree of Freedom Change; CFI = Comparative Fit Index.

In investigation of measurement invariance of a measure, equality of the parameters of the constrained models across different groups are compared. The procedure of measurement invariance comprises evaluation of the fit series of successively constrained models against a preceding equally constrained nested model. The constraints on nested model are successively added and then models are progressively analyzed. For the evaluation of the comparison of the nested models which are constrained, researchers suggested to conduct the likelihood ratio test (chi square test of difference). This chi square difference value (delta chi square) is distributed as chi square, with degrees of freedom equal to the difference in degrees of freedom (delta degree of freedom). If the chi-square difference test is statistically nonsignificant (p >.05), in the contrast of two nested models, it suggests that the two models are invariant across different groups (Hair et al. 2010).

However, the chi square test is sensitive to the sample size, number of parameters to be estimated and non-normality of the distribution (Hair et al., 2010). Cheung and Rensvold (2002) suggested a robust criterion for the evaluation of invariance investigation, the change in cumulative fit index (delta CFI), to determine whether the models compared are invariant or not. If the change in CFI is 0.01 or less, it is considered that all equal constrained specified for the nested models are acceptable; whereas, when there is a change in CFI greater than 0.01 across two nested models, the most restrictive model does not invariant.

In first stage unconstrained model was compared with the well fitted multi-group (constrained model) which depicted that the structure of factor was invariant across the music and non-music students (test of invariance of the configural model where $\Delta\chi^2$ = 113.22 with Δdf = 96 at p >.05 and ΔCFI was .009. So, it is concluded that the number of latent factors and the structure of factor loadings of the items of short test of music preferences scale were similar across the different groups i.e., music and non-music students. Consequently, the findings were validating the configural invariance of the measurement model and allowed the investigation for further stringent invariant models that is matric, scalar, factor variance, and error variance invariances.

After establishment the configural invariance, the most important test of invariance, that is, matric invariance (equal factor loadings) was carried out. As it can be observed when the item loadings of the latent factors of short test of music preference scale are equally constrained across both groups that is music and non-music. The differences in the ΔCFI between the configural model and the constrained model (matric) did not exceeded 0.01. Moreover, the $\Delta\chi^2$ = 23.51 with Δdf =16 at p >.05 was also indicating the establishment of matric invariance across groups.

At third stage invariance of mean and intercept (scalar invariance) were investigated. Equality of mean and intercept (scalar invariance) is considered as the most stringent test for measurement invariance, in which means and intercepts for the latent factors are compared to be equal across groups. The findings indicated that scalar invariance was established, as ΔCFI of both constrained models i.e., (matric and scalar) confirmed a considerable improving with .008, and $\Delta\chi^2$ = 23.37 with Δdf = 16 at p >.05 was also indicating the homogeneity of means and intercept across music and non-music students.

Factor-covariance across both groups was also tested, which indicated that both constrained models i.e., scalar and factor-covariance were invariant across both groups (music and non-music) as ΔCFI was .005 and the $\Delta\chi^2$ = 13.44 with Δdf = 8 at p >.05, which indicated that the correlations between the latent factors were similar across groups. Moreover, the evidence of error variance invariance was also indicated the equality of error variance across both groups. Whereas the variance of constrained models that is, factor covariance and error variance were found to be invariant as ΔCFI was .008 and the $\Delta\chi^2$ = 23.14 with Δdf =16 at p >.05. Hence, the results showed the strict or full measurement invariance of musical identity scale.

Discussion

The current study aimed to adapt and validate the factor structure of short test of music preferences scale. Short test of music preferences scale was originally developed by (Rentfrow & Gosling, 2003). Short test of music preferences scale is comprised of four distinctive factors First factor is labeled as reflective and complex which included four genres i.e., classical, semi-classical, *qawali*, jazz, blues and folk. Second factor labeled as intense and rebellious included three genres that is, alternative, rock, heavy metal. Third factor labeled as upbeat and conventional, included country, religious, pop and theme songs; while the last factor labeled as energetic and rhythmic and comprised of three genres of music including dance, rap, and funk.

STOMP Scale was validated across music and non-music students. Prior to data collection the scale was adapted to elaborate the genres by adding the definition of genres, famous songs composed in particular genre and famous artist to enhance the content validity of the measure so that participants can accurately reposed on the items. Further, the psychometric evolution of the short test of music preferences scale was carried out which turned into excellent reliability and validity estimates. McDonald's reliability (omega coefficient) was used to determine the internal consistency of musical identity scale. Hayes and Coutts (2020) argued that methodologists have warned that Cronbach's alpha (α) is not an optimal coefficient of internal reliability; while, the use of McDonald's omega (ω) as a measure of reliability is more optimal for confirmatory factor analysis. In addition, the evidence of validity, that is average variance extracted for convergent validity and maximum shared variance for discriminant validity, also fall within the acceptable ranges (Hair et al., 2010). Moreover, the Fornell and Larcker (1981) criteria was also taken in account while determining the discriminant validity of the scale, which also ended up with the excellent estimates. Further, measurement invariance with strict invariance, including (configural, factor loadings, mean intercepts, factor covariance, and error variance invariances) was also established (Hair et al., 2010). The findings of measurement invariance of strict invariance were also supported. Hence, it was concluded that the short test of the music preferences scale is invariant at all aspects of measurement invariance across music and non-music students.

The sub-continent adults are engaged in more music genres than the westerns. Though rock, jazz, blues, classical, rap, religious, pop, country, folk genres seemed to be common in adults of both cultures

(Upadhyay, Shukla, & Chakraborty, 2016), various other unique genres, including semi-classical (*ghazal*) and *qawali* were also identified by respondents. The adapted new genres were in line with the indigenous study conducted by (Faran, Akram, Tahir, & Malik, 2021; Rana, Ajmal, & North, 2011; Rana & North, 2007). The identified new genres were loaded under the structure of reflective and complex because of two reasons. Firstly it has strong psychometric support in term of validity and reliability, secondly, both genres are considered as the sub genres of the classical music and classified under the umbrella of classical music genre. In Pakistani sample, only the nomenclature of factor solutions (music genres) was bit unique, however, the factor structure was quite consistent with the existing literature (Rentfrow & Gosling, 2003).

Limitations and Recommendations

So far, the short test of music scale has been adapted and validated only in the Urdu language version for the Pakistani population. The self-report method was used to measure the music preferences. This limitation is compounded by the fact that participants completed the surveys in one sitting, and the results of this study are therefore not verified across a longitudinal study. The STOMP has also been verified only for music and non-music students only. Further studies including musicians and professionals from different backgrounds as well as vocalists or composers are highly recommended. Finally, the small sample size prevented establishing norms for the scale to make it psychometrically stronger. However, further studies will be conducted to establish the norms with the use of the STOMP to define low, moderate, and high scores for better diagnostic possibilities.

Conclusion and Implications

For the scale of short test of music preferences, multi-group confirmatory factor analysis validated the factor structure with excellent reliability and validity estimates across two groups that is, music and non-music students. Additionally, measurement invariance test also revealed that the short test of music preferences scale was invariant at all levels of strict invariance that is, (configural, matric, scalar, factor covariance, and error variance). The study may help the researchers working in the musical behavior of young adults that support their mental health and overall quality of life by understanding their music preferences.

References

Arbuckle, J. L. (2012). *IBM, SPSS, AMOS. User's Guide*. New York: IBM Corporation.

Byrne, B. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema, 20*(4), 872-882.

Byrne, B. M., & van de Vijver, F. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing, 10*(2), 107-132.

Byrne, B. M., Stewart, S. M., Kennard, B. D., & Lee, P. (2007). The Beck Depression Inventory-II: Testing for measurement equivalence and factor mean differences across Hong Kong and American adolescents. *International Journal of Testing, 7*(3), 293-309.

Cattell, R. B., & Anderson, J. C. (1953). *The IPAT Music Preference Test of Personality*. Champaign, IL: Institute for Personality and Ability Testing.

Chamorro-Premuzic, T., Swami, V., Furnham, A., & Maakip, I. (2009). The Big Five personality traits and uses of music in everyday life: A replication in Malaysia using structural equation modeling. *Journal of Individual Differences, 30*(1), 20-27.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(1), 233-255.

Cole, M. (1998). *Cultural psychology: A once and future discipline*. Massachusetts: Harvard University Press.

Eerola, T., Himberg, T., Toiviainen, P., & Louhivuori, J. (2006). Perceived complexity of Western and African folk melodies by Western and African listeners. *Psychology of Music, 34*(1), 341-375.

Faran, M., Saba, A., Tahir, N., & Malik, F. (2021). Music engagement and flourish: Mediating role of emotion regulation. *Journal of Behavioral Science, 31*(2), 25-45.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Gans, H. J. (1974). *Popular culture and high culture: An analysis and evaluation of taste*. New York: Basic Books.

Gardikiotis, A., & Baltzis, A. (2012). Rock music for myself and justice to the world!: Musical identity, values, and music preferences. *Psychology of Music*, *40*(2), 143-163.

Gregory, A. H., & Varney, N. (1996). Cross-cultural comparisons in the affective response to music. *Psychology of Music, 24*(1), 47-52.

Hair, J. D., Anderson, R. E., Tatham, R. L., & Black, W. C. (2010). *Multivariate data analysis* (7th ed.). New Jersey: Prentice Hall.

Hargreaves, D. J., North, A. C., & Tarrant, M. (2006). Musical preference and taste in childhood and adolescence. In G. E. McPherson (Ed.), *The child as musician: A handbook of musical development* (pp. 135-154). New York: Oxford University Press.

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability: But…. *Communication Methods and Measures, 14*(1), 1-24.

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management and Data Systems, 116*(1), 2-20.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Juslin, P. N., & Sloboda, J. (2011). *Handbook of music and emotion: Theory, research, applications*. Oxford, UK: Oxford University Press.

Kenny, D. A. (2011). *Correlated errors*: *Re-specification of latent variable model*. Retrieved from http://davidakenny.net/cm/respec.html

Lamont, A. (2006). *Musical communication*. Oxford, UK: Oxford University.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563-575.

LeBlanc, A. (l982). An interactive theory of music preference. *Journal of Music Therapy, 119*(1), 28-45.

Little, P., & Zuckerman, M. (1986). Sensation seeking and music preferences. *Personality and Individual Differences, 7*(1), 575-577.

McCown, W., Keiser, R., Mulhearn, S., & Williamson, D. (1997). The role of personality and gender in preferences for exaggerated bass in music. *Personality and Individual Differences, 23*(1), 543-547.

Mehboob, S. (2010). *Effect of music on happiness,* (Unpublished BS Honors. Thesis), Department of Psychology, Government College University. Lahore, Pakistan.

Naz, S. (2008). *The relationship between music preference dimensions and personality traits among university students* (Unpublished M.Phil. thesis), National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

Rana, S. A., Ajmal, M. A., & North, A. C. (2011). Importance of music for Pakistani youth. *Journal of Social and Clinical Psychology*, 1(*9*), 27-35.

Rana, S. A., & North, A. C. (2007). The role of music in everyday life among Pakistanis. *Music Perception, 25*(1), 59-73.

Rentfrow & Gosling (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84*(6), 1236-1256.

Rentfrow, P. J., Goldberg, L. R., & Levitin, D. J. (2011). The structure of musical preferences: A five factor model. *Journal of Personality and Social Psychology, 100*(6), 1139.

Saarikallio, S. (2008). *Cross-cultural investigation of adolescents' use of music for mood regulation*. In Proceedings of the 10th International Conference on Music Perception and Cognition, Sapporo, Japan.

Tomás, J. M., & Oliver, A. (1999). Rosenberg's Self-Esteem Scale: Two factors or method effects. *Structural Equation Modeling, 6*, 84-98.

Upadhyay, D., Shukla, R., & Chakraborty, A. (2016). Factor structure of music preference scale and its relation to personality. *Journal of Indian Academy of Applied Psychology*, *43*(1), 104-113.

Voorhees, C. M., Michael K. B., Roger C., & Edward Ramirez. (2016). Discriminant validity testing in marketing: An analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science, 44*(1), 119-134.

Received 4 December 2020

Revision received 28 October 2021