Journal of Exercise Science & Fitness 22 (2024) 322–328

Contents lists available at ScienceDirect

Journal of Exercise Science & Fitness

journal homepage: www.elsevier.com/locate/jesf

Psychometric property and measurement invariance of perceived benefits and barriers of physical activity in Chinese adolescents

Jun Dai [a,*], Han Chen [b], Haichun Sun [c]

[a] Shanghai University of Sport, School of Physical Education, Shanghai, 200438, China
[b] Valdosta State University, Department of Teacher Education KSPE, Valdosta, GA, 31698, USA
[c] University of South Florida, College of Education, Tampa, FL, USA

ARTICLE INFO

Keywords:
Benefits and barriers of physical activity
Psychometric property
Measurement invariance
Predictive validity
Structural equation modeling

ABSTRACT

Background: The perceived benefits and barriers to physical activity play crucial roles in determining daily physical activity levels. However, previous studies have employed tools lacking adequate validation, leading to inconsistent conclusions about the impact of these two factors. Therefore, this national, population-based study was conducted to assess the psychometric properties, measurement invariance, and predictive validity of the Chinese versions of the perceived benefits (C-PBEPA) and barriers to physical activity (C-PBAPA) scales.

Methods: The final sample (N = 2942, 49.3 % for boys) was randomly split into two subsamples. The first subsample was used for exploratory factor analysis and the second subsample was used for confirmatory factor analysis. Measurement invariance across gender and age groups were examined. Structural equation models were developed to examine the predictive validity of the revised C-PBEPA and C-PBAPA on moderate to vigorous PA.

Results: The results showed that both scales were unidimensional, had excellent model fit (e.g., $X^2/df < 3$, CFI >0.9, RMSEA <0.06) and demonstrated convergent validity. Findings also revealed lack of scalar invariance for C-PBAPA between preadolescents and adolescents' groups (ΔCFI >0.01) and supported the predictive validity of both scales ($p < 0.001$).

Conclusion: The study demonstrated that the revised C-PBEPA and C-PBAPA are valid scales for measuring Chinese adolescents' perceived benefits and barriers to PA.

1. Introduction

The benefits of regular physical activity (PA) have been well documented. These benefits include reduced risk of developing numerous diseases including but not limiting to type 2 diabetes, hypertension, depression, cardiovascular diseases, osteoporosis, and cancers.[1] Despite these benefits, physical inactivity (lack of PA) has been alarming prevalent worldwide. For example, the citizens of China, which has the largest population in the world, were found to have decreased PA levels during a study conducted from 1991 to 2011.[2] The decline of PA typically begins during adolescent years.[3,4] Thus, it is especially important to design strategies targeting adolescents in order to increase their level of PA.

The Health Belief Model (HBM) was developed to explain and predict individuals' health behaviors.[5] According to the HBM, people's perceived benefits of and barriers to PA are the two most prominent determinants influencing their heath behaviors such as leisure time PA.[5,6] Perceived benefits are people's assessments of the value of engaging in health-enhancing behavior while perceived barriers are individuals' assessments of obstacles preventing them from engaging in health-related behaviors.[5] Indeed, numerous studies have found the relationship between perceived benefits and/or barriers and PA.[7–9] For instance, both exercise benefits and barriers were found to be related to young adults' PA, as measured by the 43-item exercise benefits/barriers scale.[8] There are, however, some contradictory findings, indicating no relationship between benefits and/or barriers to PA.[10,11] These equivocal findings were partly due to the inconsistent and/or unstandardized instruments utilized in the studies. For example, when employing other benefits (9-item) and barriers (5-item) scales that have not been sufficiently validated, it was discovered that the benefits and barriers of PA were not correlated with MVPA for children, adolescents, and adults.[10]

A study by Robbins, Wu, Sikorskii, and Morley[12] is among the few

* Corresponding author.
E-mail addresses: daijun19800703@126.com (J. Dai), hanchen@valdosta.edu (H. Chen), Sun@usf.edu (H. Sun).

https://doi.org/10.1016/j.jesf.2024.05.002
Received 14 December 2023; Received in revised form 25 May 2024; Accepted 29 May 2024
Available online 30 May 2024

examining the psychometric property of the perceived benefits (PBEPA) and barriers (PBAPA) to PA for adolescents. Exclusively using exploratory factor analysis (EFA), their findings revealed that both PBEPA and PBAPA contain two factors. Nevertheless, both scales retain items with relatively low standardized factor loadings as well as high cross loadings on different factor. Two items cross loading in the perceived barriers scale were even higher than their own factor loadings. In addition, a robust factor analysis of a specific scale should undergo assessments for both convergent and discriminant validity. Convergent validity examines the degree to which items are related to the associated construct. Discriminant validity assesses the extent to which one latent factor is discriminant from other latent variable(s). However, these tests were rarely conducted in previous studies on PBEPA and PBAPA. Thus, it is necessary to re-examine PBEPA and PBAPA in order to better understand their dimensionality using more strict methods. Furthermore, adolescents' beliefs towards PA may be influenced by being in different cultures. However, it is not clear if PBEPA and PBAPA contain the same constructs in a different cultural context. In contrast to Western cultures, which are individualism-oriented, Chinese culture is rooted in collectivism.[13] This cultural distinction may impact individuals' perspectives on psychological variables related to exercise. Moreover, Chinese students experience higher academic stress compared to their counterparts in Western countries,[14] which could also influence perceptions of the benefits and barriers of PA. Considering the consistent PA decline among the Chinese population and the crucial PA transition period during adolescence, it is important to better understand Chinese adolescents' perceived benefits and barriers to PA. The first step, therefore, is to examine the factor structure and validate the Chinese version of the PBEPA (C-PBEPA) and PBAPA (C-PBAPA).

Scholars are also interested in exploring gender and/or age differences in the perceived benefits and barriers to PA.[15–20] These studies were conducted with the assumption that the perceived benefits and barriers to PA scales used in their studies were invariant across gender and/or age. To the best of our knowledge, however, very few studies have examined the measurement invariance test of these scales between gender groups, and no invariance test was conducted between different age groups. One study found that the perceived barriers of the PA scale were invariant at the factor variance levels across race groups for adolescent females.[21] Another study found the revised exercise benefits and barriers scales were invariant between genders at the item intercepts level.[22] As indicated by Vandenberg and Lance,[23] measurement invariance of the instrument needs to be established across groups before any meaningful between-groups comparison can be made.

Accordingly, we conducted a national, population-based study recruiting Chinese adolescents as participants. The main purposes of the study were to examine 1) factor analysis of C-PBEPA and C-PBAPA, 2) measurement invariance of C-PBEPA and C-PBAPA across gender and age groups, and 3) the predictive validity of C-PBEPA and C-PBAPA.

2. Methods

2.1. Participants

The current study is part of a large, national population-based survey investigating multivariate factors influencing Chinese adolescents' leisure time PA. In order to obtain a nationwide representative of adolescents, a total of seven cities in different geographic areas of mainland China were selected as recruitment sites. At least five public schools (elementary, middle, or high schools) were then randomly selected and contacted from each one of the seven cities. There were 24 schools from seven cities that agreed to participate in the study. The study targeted participants who were enrolled in 5th-11th grades. For each participating school, at least one grade level was randomly selected by the researchers. The head teacher(s) of the associated grade(s) at each participating school then randomly selected at least one class to participate in accordance with their own schedule. Due to preparing for the entrance examination for high school or university, students enrolled in the 9th and 12th grades during data collection were excluded from participation. In total, 3252 students participated in the national survey. For the current study, after eliminating participants with missing data, the final sample included 2942 participants whose ages were between 10 and 19 years old. The final sample was then categorized into either pre-adolescent (10–14 years) or adolescent groups (15–19 years) based on World Health Organization criteria.[24] Table 1 includes the distribution of the final sample by gender, education levels, and age groups across each recruiting city. Permission to conduct the study was approved by the Ethics Review Board in China, the schools, the participants, and their parents.

2.2. Instruments

The PBEPA and PBAPA were translated from English into Chinese independently by two bilingual scholars. The translated scales were then back translated into English in order to check the translation quality and content meaning equivalence.[25] Discussions were conducted between scholars for consensus and resolved some minor discrepancies. No additional items were included, and the wording of the original scales remained unchanged. In order to reduce response bias, 5-point Likert scales with a neutral point were employed in the study.

PBEPA. The perceived benefits of PA was assessed using the 10-item PBEPA scale adopted from a previous study.[12] The 10 items were preceded by the stem sentence "A major reason for being physically active or exercising for me is to …" The respondents were then asked to rate each one of the 10 items on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores of each item indicate greater perceived benefits of PA. In Robins et al.'s study, the 10-item PBEPA contains two factors representing outcome orientation (six items) and personal satisfaction resulting from PA (four items).

PBAPA. The perceived barriers to PA was assessed with the nine-item PBAPA scale adopted from previous research.[12] Following the stem sentence, "Please show how true each statement is regarding certain barriers or problems that prevent or stop you from exercising, being active, or doing sports …" respondents were asked to rate each item anchored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores of each item represent greater perceived barriers to PA. Robbins and her colleagues found that the nine-item PBAPA includes two factors measuring negative personal emotions (four items) and personal sense of immobilization (five items).

Weekly moderate to vigorous PA (MVPA). The participants' weekly MVPA was measured using a Chinese version of the International Physical Activity Questionnaire-short form (IPAQ-S). For the purpose of the current study, the students were asked to self-recall two different types of activities (i.e., moderate and vigorous) they performed during the past seven days. The IPAQ-S includes descriptions as well as several examples for both moderate and vigorous PA. Based on the IPAQ-S protocol, the students recorded the number of days as well as the duration in minutes (activities lasting below 10 min per day were not counted) for when they participated in moderate and vigorous PA during the past week. The total weekly MVPA was obtained by adding the total minutes recorded for either moderate or vigorous PA from each day. The reliability and validity of the IPAQ-S has been established in a cross-national study.[26]

2.3. Data analysis

The final sample ($N = 2942$) was randomly split into two subsamples using random case selection procedure in SPSS. The first subsample ($n = 1469$) was used for EFA and the second subsample ($n = 1473$) was used for confirmatory factor analysis (CFA). The descriptions of the specific analytic strategies are described in the following paragraph.

EFA was conducted in SPSS version 22.0 on sample 1. It was expected that two-factor models would be extracted for both C-PBEPA and C-

Table 1
Distribution of participants by gender, education levels, and age groups across recruiting cities.

		Gender		Education Levels			Age groups	
	Schools	Boys	Girls	Upper elementary (5th-6th)	Middle (7th-8th)	High (10th-11th)	Pre-adolescents	Adolescents
	N	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Total	24	1450 (49.3 %)	1492 (50.7 %)	614 (20.9 %)	1293 (43.9 %)	1035 (35.2 %)	1733 (58.9 %)	1209 (41.1 %)
Shanghai (Largest city in eastern China)	4	251 (48.8 %)	263 (51.2 %)	100 (19.5 %)	195 (37.9 %)	219 (42.6 %)	277 (53.9 %)	237 (46.1 %)
Guangzhou (Largest city in south central China)	2	114 (47.7)	125 (52.3 %)	0	93 (38.9 %)	146 (61.1 %)	93 (38.9 %)	146 (61.1 %)
XiAn (Largest city in south central China)	4	243 (55.2 %)	197 (44.8)	39 (8.9 %)	195 (44.3 %)	206 (46.8 %)	230 (52.3 %)	210 (47.7 %)
Ürümqi (Largest city in western China)	4	167 (46.6 %)	191 (53.4 %)	101 (28.2 %)	161 (45 %)	96 (26.8 %)	220 (61.5 %)	138 (38.5 %)
Yuling (North Shaanxi Province)	4	235 (47.3 %)	262 (52.7 %)	127 (25.6 %)	324 (65.2 %)	46 (9.3 %)	396 (79.7 %)	101 (20.3 %)
Chuzhou (East Anhui Province)	3	210 (47.7 %)	230 (52.3 %)	98 (22.3 %)	179 (40.7 %)	163 (37 %)	265 (60.2 %)	175 (39.8 %)
Heihe (northeast region of Heilongjiang Province)	3	230 (50.7 %)	224 (49.3 %)	149 (32.8 %)	146 (32.2 %)	159 (35 %)	252 (55.5 %)	202 (44.5 %)

Note. N = number.

PBAPA scales and these two factors in each scale would be correlated with each other based on previous research.[12] We therefore used the Principle Axis Factoring extraction method with Promax (oblique) rotation on both the 10-item C-PBEPA and nine-item C-PBAPA separately. The number of factors extracted was determined by examining scree plot[27] and Kaiser' criteria where eigenvalues greater than 1 for all factors were retained.[28]

CFA was conducted using AMOS version 20.0 on sample 2. First, a two-factor solution identified from Robbins et al.'s study[12] was examined. Second, in case a different factor solution was retrieved from the EFA test, CFA was conducted again on sample 2 examining this different factor solution model and comparing it with Robbins et al.'s two-factor model. The following multiple fit indices were used to assess model fit: chi-square (χ^2)/degree of freedom (df), comparative fix index (CFI), goodness of fit index (GFI), normed fit index (NFI), Tucker-Lewis index (TLI), root-mean-square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Following the recommendations, the accepted cutoffs for χ^2/df should be below 3[29]; the values of CFI, GFI, NFI, and TLI should be greater than 0.90 [30]; and the thresholds for RMSEA and SRMR should be less than 0.06 and 0.08, respectively.[30] We also retained those items with significant standardized factor loadings equals to or above the 0.6 threshold as recommended by Hair et al.[31]

Convergent and discriminant validity of the constructs were examined using several tests. Convergent validity examines the degree to which items are related to the associated construct. Convergent validity of the constructs is established if the values of the composite reliability (CR) and the average variance extracted (AVE) exceeds the criterion of 0.70 and 0.50, respectively.[32] Discriminant validity assesses the extent to which one latent factor is discriminant from other latent variable(s). The discriminant validity test is established if the square root of the AVE extracted by a construct is higher than the correlations between the factors.[33]

After the best measurement models for both the C-PBEPA and C-PBAPA scales were obtained, we conducted measurement invariance tests for each scale. Following the recommendation of Vandenberg and Lance[23] to use progressively restrictive stages, the invariance tests across gender and age groups (i.e., pre-adolescents and adolescents) for each scale were conducted at configural, metric, and scalar levels. The configural test aims to establish a baseline model across groups while the metric and scalar invariance tests assess the equivalence of factor loadings and both factor loadings and indicator intercepts, respectively. The metric and scalar invariance is established if the CFI change is < 0.01.[34]

Finally, in order to test the predictive validity of both revised C-PBEPA and C-PBAPA scales, several structural equation models (SEMs) were developed that contained C-PBEPA, C-PBAPA, and weekly MVPA for groups of boys, girls, pre-adolescents, and adolescents. It is hypothesized that the perceived benefits of PA is positively related to participants' weekly MVPA while the perceived barriers to PA is negatively related to it for each group.

3. Results

3.1. EFA

For the PBEPA scale, the requirements for conducting EFA were met by the value of the KMO test (KMO = 0.913) and the outcome of Bartlett's Test of Sphericity, X^2 (45) = 7001.75, $p < 0.001$. The EFA test for PBEPA found one factor with an eigenvalue >1 and this single-factor explained 47.61 % of the variance. The assessment of the scree plot also supported the one-factor solution for PBEPA. For the PBAPA scale, the prerequisites for conducting an EFA test were also fulfilled by the value of the KMO test (KMO = 0.916) and the results of Bartlett's Test of Sphericity, X^2 (36) = 6333.483, $p < 0.001$. The EFA for PBAPA extracted one factor with an eigenvalue >1, which explained 49.46 % of the variance. The result of the scree plot also revealed a one-factor model for PBAPA.

3.2. CFA

Two-factor models. We conducted a three-step item deduction approach in order to improve model fit. In step 1, we removed any items with low factor loading (i.e., standardized factor loading <0.60). We continued this approach after each item was removed, deleting the item with the lowest factor loading each time. The CFA test was then respecified following each item removal. We repeated this step until all items' standardized factor loadings were above 0.60. In step 2, we checked the modification indices (MI) to identify any item(s) with high cross loading on the item from a different factor or on a different factor (MI > 25).[35] This step was also repeated, removing the item with the highest cross loading each time. The CFA test was again respecified following each item removal. We repeated this step until items had no or trivial cross loadings. In step 3, MI was assessed to identify pairs of items within each factor with high MI resulting from error covariances. Error covariances usually occur when items are similarly worded.[35] Generally, correlating

indicators' error terms is discouraged as this step requires strong theoretical justification.[35,36] We therefore decided to remove one item in the pair in order to improve model fit and reduce redundancy when error covariances existed.[37] In order to decide which item would be eliminated when a pair of items with the highest MI was identified, we calculated each item's total MI with all other items within the same factor and removed the one with the highest total MI. We repeatedly conducted this step and respecified the CFA test following each item removal. In the case of a factor with three indicators, item deduction did not proceed further in order to satisfy the minimum item requirement. Table 2 outlines the specific steps for item deduction as well as fit indices for each model respecification.

As can be seen from Table 2, the initial two-factor models for both scales generated from a previous study[12] did not demonstrate acceptable model fit (e.g., for both initial models X^2/df > 3, RMSEA >0.06). After items deduction, the values of fit indices for the final six-item model of C-PBEPA were still not satisfactory (e.g., X^2/df > 3, RMSEA >0.06). Similarly, for C-PBAPA, even after some problematic items were removed, the final six-item model still lacked model fit (the value of X^2/df and RMSEA almost stayed the same).

Convergent and discriminant validity of the revised two-factor models. The final two-factor models for both scales were examined for convergent and discriminant validity. For both revised scales, the values of CR and AVE for both factors exceeded the criterion, supporting the convergent validity for both factors of the C-PBEPA and C-PBAPA (see Table 3). However, results of the discriminant validity test found that the values of the square root of AVE (in bold) were lower than the values of factor correlation for both scales. This finding revealed that the revised two-factor C-PBEPA and C-PBAPA lack discriminant validity (see Table 3).

One-factor vs. two-factor models. The findings from both EFA and discriminant validity tests for both scales imply that both C-PBEPA and C-PBAPA contain only one factor. Therefore, we utilized a two-step item deduction approach (we only followed step 1 and step 3 as cross-loading check is unnecessary for the one-factor model) described previously to improve model fit. The procedure for items deduction was also included in Table 3. For C-PBEPA, the initial 10-item model did not demonstrate satisfactory model fit (X^2/df > 3, RMSEA >0.06). After we dropped several problematic items, the final six-item scale showed excellent model fit. Similarly, for C-PBAPA, the initial nine-item model had unacceptable model fit (X^2/df > 3, RMSEA >0.06). After some items deduction, the final four-item model demonstrated great model fit. Additionally, CR (0.89 and 0.84 for C-PBEPA and C-PBAPA, respectively) and AVE (0.57 and 0.56 for C-PBEPA and C-PBAPA, respectively) exceeded the cut-off criterion, supporting the convergent validity for both revised single-factor scales. Further, both revised scales demonstrate good internal consistency (Cronbach's alpha = 0.89 and 0.83 for the revised C-PBEPA and C-PBAPA, respectively). Model comparisons clearly show that the final one-factor models are superior than the two-factor models for both C-PBEPA and C-PBAPA.

Table 3
Convergent and discriminant validity of the revised two-factor models for C-PBEPA and C-PBAPA.

	CR	AVE	Personal satisfaction resulting from PA	Outcome orientation
Revised 8-item two-factor PBE-PA scale				
Personal satisfaction resulting from PA	0.82	0.60	**0.78**	
Outcome orientation	0.77	0.53	0.90 Negative personal emotions	**0.73** Personal sense of immobilization
Revised 7-item two-factor PBA-PA scale				
Negative personal emotions	0.75	0.51	**0.71**	
Personal sense of immobilization	0.77	0.53	0.96	**0.73**

Note. CR = composite reliability; AVE = average variance extracted; Square root of AVE in bold on diagonals; off diagonals are Pearson correlation of constructs.

3.3. Measurement invariance

The results for measurement invariance tests for the revised one-factor C-PBEPA are included in Table 4. The findings show that C-PBEPA is invariant between gender for sample 1 and total sample at both configural, full metric, and full scalar levels (ΔCFI <0.01). For sample 2,

Table 2
Item deduction procedure and the fit indices following each item removal for both two-factor and one-factor models of scales.

	X^2	df	X^2/df	GFI	CFI	RMSEA	NFI	TLI	SRMR
PBE-PA two factor solution									
Initial two factor_10 items	456.88***	34	13.44	0.94	0.95	0.09	0.94	0.93	0.04
delete 1 (factor loading = 0.48)	336.82***	26	12.96	0.95	0.96	0.09	0.95	0.94	0.03
delete 3 (factor loading = 0.57)	180.93***	19	9.52	0.97	0.98	0.08	0.97	0.96	0.03
cross loading MI (7,6) 39.366; remove 7	113.23***	13	8.71	0.98	0.98	0.07	0.98	0.97	0.03
crossloading MI (8,5) 20.27; remove 8 (Final model[a])	74.52***	8	9.32	0.98	0.98	0.08	0.98	0.97	0.02
PBA-PA two factor solution									
Initial two factor_9 items	541.87***	26	20.84	0.92	0.92	0.12	0.92	0.89	0.05
delete Bar1 (loading 0.53)	486.02***	19	25.58	0.92	0.93	0.13	0.92	0.89	0.05
cross loading MI (8,9) 107.471; remove 9	279.58***	13	21.51	0.95	0.95	0.12	0.94	0.91	0.04
cross loading MI (8,7) 44.063; remove 7 (Final model[b])	163.92***	8	20.49	0.96	0.96	0.12	0.95	0.92	0.04
PBE-PA one factor solution									
Initial one factor_10 items	583.46***	35	16.67	0.93	0.93	0.10	0.93	0.91	0.04
delete 1 (loading 0.46)	457.38***	27	16.94	0.94	0.94	0.10	0.94	0.92	0.04
delete 3 (loading 0.57)	310.23***	20	15.51	0.95	0.96	0.10	0.95	0.94	0.03
cross loading MI (4,5) 129.755; remove 5	89.65***	14	6.40	0.98	0.99	0.06	0.98	0.98	0.02
cross loading MI (6,9) 35.152; remove 6 (Final model[c])	25.52***	9	2.84	1.00	1.00	0.04	0.99	0.99	0.01
PBA-PA one factor solution									
Initial one factor_9 items	550.48***	27	20.39	0.92	0.92	0.12	0.92	0.90	0.05
delete 1 (loading 0.52)	486.63***	20	24.33	0.92	0.93	0.13	0.92	0.89	0.05
cross loading MI (4,5) 106.059; remove 4	278.74***	14	19.91	0.95	0.95	0.11	0.95	0.93	0.04
cross loading MI (2,3) 90.365; remove 2	138.42***	9	15.38	0.97	0.97	0.10	0.97	0.95	0.03
cross loading MI (8,9) 41.428; remove 9	44.37***	5	8.88	0.99	0.99	0.07	0.99	0.97	0.02
cross loading MI (5,7) 18.828; remove 5 (Final model[d])	5.02	2	2.51	1.00	1.00	0.03	1.00	1.00	0.01

Note. [a] outcome orientation (2,9,10); personal satisfaction (4,5,6); [b] NegPerEmotion (4,5,8); Immobilization (2,3,6); [c] item (2,4,7,8,9,10); [d] item (3,6,7,8); ***p < 0.001.

Table 4
Measurement invariance between gender and age groups for C-PBEPA.

	X^2	df	CFI	RMSEA	ΔCFI
Sample 1 for EFA					
Gender (2 levels)					
Configural invariance	137.66	18	0.97	0.07	
Full Metric invariance	147.49	23	0.97	0.06	0.001
Full Scalar invariance	185.27	29	0.96	0.06	0.008
Age (2 levels)					
Configural invariance	153.46	18	0.97	0.07	
Full Metric invariance	165.64	23	0.96	0.07	0.002
Full Scalar invariance	200.46	29	0.96	0.06	0.007
Sample 2 for CFA					
Gender (2 levels)					
Configural invariance	75.65	18	0.99	0.05	
Full Metric invariance	84.78	23	0.99	0.04	0.001
Full Scalar invariance	144.04	29	0.97	0.05	0.012
Age (2 levels)					
Configural invariance	44.11	18	0.99	0.03	
Full Metric invariance	47.49	23	1	0.03	0.001
Full Scalar invariance	67.58	29	0.99	0.03	0.004
Total sample					
Gender (2 levels)					
Configural invariance	158.13	18	0.98	0.05	
Full Metric invariance	174.39	23	0.98	0.05	0.002
Full Scalar invariance	260.22	29	0.97	0.05	0.009
Age (2 levels)					
Configural invariance	144.77	18	0.99	0.05	
Full Metric invariance	154.89	23	0.98	0.04	0.001
Full Scalar invariance	203.11	29	0.98	0.05	0.005

a gender invariance test at the scalar level was not established (ΔCFI >0.01). In addition, C-PBEPA is invariant between preadolescent and adolescent groups at both configural, full metric, and full scalar levels for sample 1, sample 2, and total sample (ΔCFI <0.01).

Table 5 lists outcomes for the invariance test of the revised one-factor C-PBAPA. Gender invariant tests were established at both configural, full metric, and full scalar levels for all three groups (ΔCFI <0.01). Full scalar invariance between age groups was not established for all three groups (ΔCFI >0.01).

Table 5
Measurement invariance between gender and age groups for C-PBAPA.

	X^2	df	CFI	RMSEA	ΔCFI
Sample 1 for EFA					
Gender (2 levels)					
Configural invariance	2.04	4	1	0	
Full Metric invariance	3.08	7	1	0	0
Full Scalar invariance	25.22	11	0.993	0.03	0.007
Age (2 levels)					
Configural invariance	6.99	4	0.999	0.02	
Full Metric invariance	7.36	7	1	0.01	0.001
Full Scalar invariance	63.26	11	0.975	0.06	0.025
Sample 2 for CFA					
Gender (2 levels)					
Configural invariance	5.06	4	1	0.01	
Full Metric invariance	9.21	7	0.999	0.02	0.001
Full Scalar invariance	30.36	11	0.991	0.04	0.008
Age (2 levels)					
Configural invariance	6.88	4	0.999	0.02	
Full Metric invariance	12.31	7	0.998	0.02	0.001
Full Scalar invariance	71.97	11	0.972	0.06	0.026
Total sample					
Gender (2 levels)					
Configural invariance	1.02	4	1	0	
Full Metric invariance	2.95	7	1	0	0
Full Scalar invariance	44.2	11	0.992	0.03	0.008
Age (2 levels)					
Configural invariance	5.55	4	1	0.01	
Full Metric invariance	7.82	7	1	0.01	0
Full Scalar invariance	118.76	11	0.975	0.06	0.025

3.4. Predictive validity of C-PBEPA and C-PBAPA

In order to examine the predictive validity of the revised scales, we developed four SEMs that include C-PBEPA, C-PBAPA, and weekly MVPA for different gender and age groups. Findings show that both models fit the data very well (X^2/df = 5.43 to 3.75; CFI, GFI, NFI, and TLI >0.95 for all models; RMSEA <0.05 for all models). As also can be seen in Fig. 1, the participants' perceived benefits of PA is positively related and perceived barriers to PA is negatively related to their weekly MVPA for boys, girls, preadolescents, and adolescents groups.

4. Discussion

This study examined the psychometric property of the Chinese version PBEPA and PBAPA scales. Findings support the unidimensionality of both C-PBEPA and C-PBAPA. The revised C-PBEPA and C-PBAPA showed excellent model fit, demonstrated convergent validity, and had good internal reliability. Findings also revealed lack of scalar invariance for C-PBAPA between preadolescents and adolescents' groups and supported the predictive validity of both scales.

Generated from a comprehensive literature review and several existing instruments from Garcia et al.'s study,[38] the original item pool was revised based on adolescent girls' feedback. This revision elicited 20 and 23 items for measuring perceived benefits and barriers to PA, respectively.[39] In order to reduce response burden and completion time, several items were removed from both scales to generate PBEPA and PBAPA. The removed items were non-major benefits and barriers identified by girls and did not contribute to increasing internal reliability.[12] Using the EFA test, Robbins and her colleagues further tested both scales' construct validity and found that both PBEPA and PBAPA contain two factors. However, the Robbins et al. study had at least two major flaws. First, several items that were retained in both scales had very low factor loadings (e.g., one item from PBEPA and three items from PBAPA had factor loadings <0.5). Second, several items that were retained in both scales had high cross loadings (e.g., >0.5). For example, two retained items from PBAPA had higher cross loadings than their own factor loadings. In the current study, results from EFA supported the unidimensionality of the Chinese version of PBEPA and PBAPA. We also conducted numerous CFA tests to example the model fit of the original two-factor solution for both scales. Despite removing several problematic items due to low factor loading or cross-loading, the revised two-factor model for both scales still exhibited X^2/df values much higher than the cut-off of 3, indicating an unacceptable model fit. Additionally, the revised two-factor solution did not pass the discriminant validity test. This finding suggests that the two factors identified in the previous study,[12] namely Outcome Orientation and Personal Satisfaction for PA Benefits, and Negative Personal Emotions and Personal Sense of

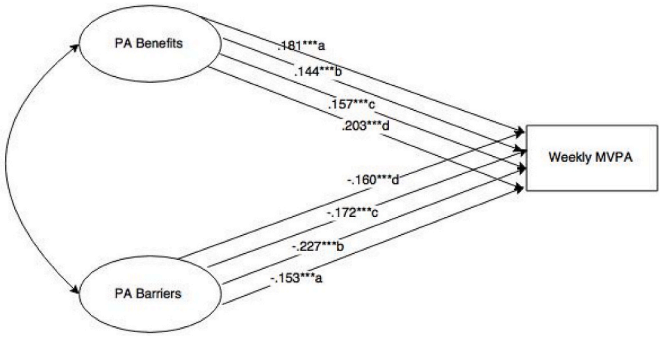

Fig. 1. Validity of the revised one-factor perceived benefits and barriers to PA scales in predicting weekly MVPA for boys, girls, preadolescents, and adolescents' groups
Note. a = boys; b = girls; c = preadolescents; d = adolescents; ***$p < 0.001$; all numbers are standardized regression weights.

Immobilization for PA Barriers, may belong to a single construct. Subsequent multiple CFA tests based on a single-factor solution also identified several problematic items due to low factor loading and cross-loading. After the removal of these items, the one-factor solution of C-PBEPA and C-PBAPA demonstrated good model fit. In the CFA test, it is crucial to establish robust and unidimensional constructs.[40] Consequently, correlations among item residuals should be avoided, and the elimination of problematic items is essential to enhance model fit.[41] The present study thus investigated MI, removing items with the highest total MI one at a time, and subsequently reconfiguring the CFA model after each item removal. Revised C-PBEPA and C-PBAPA demonstrated good mode fit, convergent validity, and good internal reliability. Additionally, Robbins et al. utilized a 4-point Likert scale, compelling respondents to choose opinions leaning toward either end of the spectrum in response to the questions. In contrast, the current study adopted a 5-point Likert scale that includes a neutral midpoint. Research indicates that, when compared to even-numbered scales, odd-numbered Likert scales with a neutral midpoint option can mitigate response bias.[42]

Measurement invariance examines the degree to which an instrument is measured similarly across groups.[23] Failure to establish measurement invariance of a scale would make it difficult to conclude whether the differences between groups come from variances in measurement or result from true latent differences. We conducted invariance tests at the scalar level (indicators' intercepts invariance) as scalar invariance is the prerequisite for scores comparisons between groups. Our study found that that C-PBEPA is invariant between age groups at configural, metric, and scalar levels for sample 1, sample 2, and total sample. These findings demonstrated that for both preadolescents and adolescents' groups, C-PBEPA was viewed as the same construct, and had the same factor loadings and indicators intercepts. Thus, meaningful comparisons in perceived benefits of PA can be made using C-PBEPA between these two age groups. In contrast, C-PBAPA is not invariant at the scalar level between two age groups for sample 1, sample 2, and total samples. This finding seems to preclude using C-PBAPA for comparing perceived barriers to PA between preadolescents and adolescents in China. However, C-PBAPA was invariant at the scalar level across gender for sample 1, sample 2, and total sample, supporting the use of this instrument for gender comparison in perceived barriers of PA. Finally, C-PBEPA was found invariant at the scalar level across gender only for sample 1 and total sample. Further study is needed investigating scalar invariance of C-PBEPA between gender before conclusions can be made regarding if scholars can use C-PBEPA for gender comparisons of perceived benefits of PA among Chinese adolescents.

SEM results indicated positive associations between perceived benefits of PA and weekly MVPA, while perceived barriers to PA showed negative associations for boys, girls, preadolescents, and adolescents. This finding is promising and is consistent with previous research.[7,43,44] It may suggest that C-PBEPA and C-PBAPA are valid tools in assessing Chinese adolescents' perceived benefits and barriers to PA. Interventions designed to enhance Chinese adolescents' PA should include strategies targeting the improvement of their perceptions of PA benefits. Strategies should also be designed to help adolescents overcome their perceived PA barriers.

This study is not without its limitation. First, PBEPA and PBAPA were generated mainly from Garcia and colleagues' study[38] and was modified based on feedback from adolescent girls. Future studies should consult opinions from both adolescent boys and girls and include more items in both benefits and barriers scales before examining their factorial analysis. Second, the Robbins et al. study[12] used a 4-point Likert scale while the current study utilized a 5-point Likert scale. Although a previous study found both 4-point and 5-point formats of a scale had the same structure,[45] future studies should conduct factor analysis of C-PBEPA and C-PBAPA using a 4-point format. Third, SEMs only includes benefits, barriers, and MVPA. Control variables should be added in SEM in future studies. Finally, participants' MVPA was measured based on their self-recall; more objective measurement of PA using motion sensors such as accelerometers should be used in future studies. Despite these limitations, to the best of our knowledge, this is the first study that empirically examined the validity and invariance of a Chinese version of perceived benefits and barriers to PA scales using a population-based sample. The C-PBEPA and C-PBAPA may serve as useful tools in the assessment of perceived benefits and barriers of PA among Chinese adolescents.

5. Conclusion

The study found that the revised 6-item C-PBEPA and 4-item C-PBAPA are unidimensional and demonstrated good model fit and predictive validity for Chinese adolescents. Additionally, the C-PBEPA exhibits invariance at the full scalar levels between genders and age groups. The C-PBAPA is also invariant at the full scalar level between genders. However, no scalar invariance test was established for C-PBAPA between the two age groups. Meaningful gender comparisons can be made based on both scales. While the C-PBEPA scale enables accurate comparisons of PA benefits between gender groups, scholars should exercise caution when using the C-PBAPA scale to compare levels of PA barriers between age groups.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Jun Dai reports financial support was provided by National Social Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Lee I-M, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*. 2012;380(9838):219–229. https://doi.org/10.1016/S0140-6736(12)61031-9.
2. Ng SW, Howard AG, Wang HJ, Su C, Zhang B. The physical activity transition among adults in China: 1991-2011. *Obes Rev*. Jan 2014;15(Suppl. 1):27–36. https://doi.org/10.1111/obr.12127.
3. Dumith SC, Gigante DP, Domingues MR, Kohl HW. Physical activity change during adolescence: a systematic review and a pooled analysis. *Int J Epidemiol*. 2011;40(3):685–698. https://doi.org/10.1093/ije/dyq272.
4. Kjønniksen L, Torsheim T, Wold B. Tracking of leisure-time physical activity during adolescence and young adulthood: a 10-year longitudinal study. *Int J Behav Nutr Phys Act*. 2008;5(1):69. https://doi.org/10.1186/1479-5868-5-69.
5. Janz NK, Becker MH. The health belief model: a decade later. *Health Educ Behav*. 1984;11(1):1–47. https://doi.org/10.1177/109019818401100101.
6. Sheeran P, Abraham C. The health belief model. In: Conner M, Norman P, eds. *Predicting Health Behavior*. 1995:23–61.
7. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. *Med Sci Sports Exerc*. 2000;32(5):963–975. https://doi.org/10.1097/00005768-200005000-00014.
8. Grubbs L, Carter J. The relationship of perceived benefits and barriers to reported exercise behaviors in college undergraduates. *Fam Community Health*. 2002;25(2):76–84. https://doi.org/10.1097/00003727-200207000-00009.
9. Reichert FF, Barros AJ, Domingues MR, Hallal PC. The role of perceived personal barriers to engagement in leisure-time physical activity. *Am J Public Health*. 2007;97(3):515–519. https://doi.org/10.2105/ajph.2005.070144.
10. Deflandre A, Antonini PR, Lorant J. Perceived benefits and barriers to physical activity among children, adolescents and adults. *Int J Sport Psychol*. 2004;35(1):23–36.
11. Wu TY, Pender N. Determinants of physical activity among Taiwanese adolescents: an application of the health promotion model. *Res Nurs Health*. 2002;25(1):25–36. https://doi.org/10.1002/nur.10021.
12. Robbins LB, Wu T-Y, Sikorskii A, Morley B. Psychometric assessment of the adolescent physical activity perceived benefits and barriers scales. *J Nurs Meas*. 2008;16(2):98–112. https://doi.org/10.1891/1061-3749.16.2.98.
13. Zhang J. Collectivism or individualism: an analysis of Chinese interactive culture. *American Review of China Studies*. 2000;1(2):57–65.
14. Zhao X, Selman RL, Haste H. Academic stress in Chinese schools and a proposed preventive intervention program. *Cogent Educ*. 2015;2(1), 1000477. https://doi.org/10.1080/2331186x.2014.1000477.

15. Robbins LB, Sikorskii A, Hamel LM, Wu TY, Wilbur J. Gender comparisons of perceived benefits of and barriers to physical activity in middle school youth. *Res Nurs Health*. 2009;32(2):163–176. https://doi.org/10.1002/nur.20311.
16. Gunnell KE, Brunet J, Wing EK, Bélanger M. Measuring perceived barriers to physical activity in adolescents. *Pediatr Exerc Sci*. 2015;27(2):252–261. https://doi.org/10.1123/pes.2014-0067.
17. Tergerson JL, King KA. Do perceived cues, benefits, and barriers to physical activity differ between male and female adolescents? *J Sch Health*. 2002;72(9):374–380. https://doi.org/10.1111/j.1746-1561.2002.tb03562.x.
18. Butt J, Weinberg RS, Breckon JD, Claytor RP. Adolescent physical activity participation and motivational determinants across gender, age, and race. *J Phys Act Health*. 2011;8(8):1074–1083. https://doi.org/10.1123/jpah.8.8.1074.
19. Arzu D, Tuzun EH, Eker L. Perceived barriers to physical activity in university students. *J Sports Sci Med*. 2006;5(4):615–620.
20. De Bourdeaudhuij I, Sallis J. Relative contribution of psychosocial variables to the explanation of physical activity in three population-based adult samples. *Prev Med*. 2002;34(2):279–288. https://doi.org/10.1006/pmed.2001.0979.
21. Dishman RK, Hales DP, Sallis JF, et al. Validity of social-cognitive measures for physical activity in middle-school girls. *J Pediatr Psychol*. 2010;35(1):72–88. https://doi.org/10.1093/jpepsy/jsp031.
22. Koehn S, Amirabdollahian F. Reliability, validity, and gender invariance of the exercise benefits/barriers scale: an emerging evidence for a more concise research tool. *Int J Environ Res Public Health*. 2021;18(7):3516. https://doi.org/10.3390/ijerph18073516.
23. Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organ Res Methods*. 2000;3(1):4–70. https://doi.org/10.1177/109442810031002.
24. World Health Organization. *Young People's Health-A Challenge for Society. Report of a WHO Study Group on Young People and "Health for All by the Year 2000*. 1986. *WHO Technical Report Series 731*.
25. Brislin RW. Back-translation for cross-cultural research. *J Cross Cult Psychol*. 1970;1:185–216. https://doi.org/10.1177/135910457000100301.
26. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB.
27. Cattell RB. The scree test for the number of factors. *Multivar Behav Res*. 1966;1(2):245–276. https://doi.org/10.1207/s15327906mbr0102_10.
28. Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas*. 1960;20(1):141–151. https://doi.org/10.1177/001316446002000116.
29. Bollen KA. A new incremental fit index for general structural equation models. *Sociol Method Res*. 1989;17(3):303–316. https://doi.org/10.1177/0049124189017003004.
30. Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychol Methods*. 1998;3(4):424–453. https://doi.org/10.1037//1082-989x.3.4.424.
31. Hair J, Babin B, Anderson R, Tatham R. In: *Multivariate Data Analysis*. sixth ed. New York: Prentice Hall; 2006.
32. Hair J, Black W, Babin B, Anderson R. In: *Multivariate Data Analysis*. seventh ed. 2010.
33. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res*. 1981;18(1):39–50. https://doi.org/10.2307/3151312.
34. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling*. 2002;9(2):233–255. https://doi.org/10.1207/s15328007sem0902_5.
35. Brown TA. In: *Confirmatory Factor Analysis for Applied Research*. 2 ed. New York: Guilford Press; 2015.
36. Jöreskog K, Long JS. Introduction. In: Bollen KA, Long JS, eds. *Testing Structural Equation Models*. Newbury Park, CA: Sage; 1993.
37. Muthén LK, Muthén BO. In: *Mplus User's Guide*. sixth ed. Los Angeles, CA: Author; 2010.
38. Garcia AW, Broda MAN, Frenn M, Coviak C, Pender NJ, Ronis DL. Gender and developmental differences in exercise beliefs among youth and prediction of their exercise behavior. *J Sch Health*. 1995;65(6):213–219. https://doi.org/10.1111/j.1746-1561.1995.tb03365.x.
39. Robbins LB, Gretebeck KA, Kazanis AS, Pender NJ. Girls on the move program to increase physical activity participation. *Nurs Res*. 2006;55(3):206–216. https://doi.org/10.1097/00006199-200605000-00007.
40. Anderson JC, Gerbing DW. Structural equation modeling in practice: a review and recommended two-step approach. *Psychol Bull*. 1988;103(3):411–423. https://doi.org/10.1037//0033-2909.103.3.411.
41. Landis RS, Edwards BD, Cortina JM. On the practice of allowing correlated residuals among indicators in structural equation models. In: Lance CE, Vandenberg RJ, eds. *Statistical and Methodological Myths and Urban Legends: Doctrine, Verity, and Fable in the Organizational and Social Sciences*. New York, NY: Routledge; 2009.
42. Croasmun JT, Ostrom L. Using likert-type scales in the social sciences. *J Adult Educ*. 2011;40(1):19–22. https://eric.ed.gov/?id=EJ961998.
43. Leone LA, Ward DS. A mixed methods comparison of perceived benefits and barriers to exercise between obese and non-obese women. *J Phys Act Health*. 2013;10(4):461–469. https://doi.org/10.1123/jpah.10.4.461.
44. Lubans DR, Sylva K, Morgan PJ. Factors associated with physical activity in a sample of British secondary school students. *Aust J Educ Dev Psychol*. 2007;7:22–30. https://eric.ed.gov/?id=EJ815618.
45. Adelson JL, McCoach DB. Measuring the mathematical attitudes of elementary students: the effects of a 4-point or 5-point Likert-type scale. *Educ Psychol Meas*. 2010;70(5):796–807. https://doi.org/10.1177/0013164410366694.