

1-1-2019

# Measurement Invariance and Latent Mean Differences of the Chinese Version Physical Activity Self-Efficacy Scale Across Gender and Education Levels

Han Chen

*Valdosta State University*

Jun Dai

*Shanghai University of Sport*

Yong Gao

*Boise State University*

---

Original article

# Measurement invariance and latent mean differences of the Chinese version physical activity self-efficacy scale across gender and education levels

Han Chen<sup>a</sup>, Jun Dai<sup>b,\*</sup>, Yong Gao<sup>c</sup>

<sup>a</sup> Department of Kinesiology and Physical Education, Valdosta State University, Valdosta, GA 31698, USA

<sup>b</sup> School of Sport Leisure, Recreation, and Arts, Shanghai University of Sport, Shanghai 200438, China

<sup>c</sup> Department of Kinesiology, Boise State University, Boise, ID 83725, USA

Received 30 April 2016; revised 4 October 2016; accepted 20 December 2016

Available online 11 January 2017

## Abstract

**Background:** Self-efficacy has been identified as an important determinant of youth's behavior change including physical activity (PA) participation. However, the dimensionality check of a PA self-efficacy scale has rarely been conducted in China. The current study aims to examine (1) the unidimensionality of a shortened Chinese version of PA self-efficacy scale (S-PASESC); (2) the measurement invariance of S-PASESC across gender and levels of education; (3) the latent factor mean difference between gender and levels of education; (4) the direct effects of self-efficacy on PA by different gender and education levels; and (5) the comparisons of the direct effects of self-efficacy on PA across gender and education levels.

**Methods:** The participants were 5th through 11th grade public school students recruited from 7 cities located in different geographic regions of China. The final data include a total of 3003 participants (49.7% boys) who have completed the scales.

**Results:** Confirmatory factor analysis (CFA) test supported the unidimensionality of S-PASESC. The S-PASESC is invariant across gender and 3 levels of education at both configural, full metric, and full scalar levels. Findings from latent mean comparisons showed that boys reported higher PA self-efficacy than girls. Students' perceived PA self-efficacy tend to decrease from elementary to high school. Finally, self-efficacy positively related to PA by groups of different gender and education levels and the relationship between self-efficacy and PA is stronger among middle school boys than girls.

**Conclusion:** Findings suggest S-PASESC is a valid scale for measuring Chinese students' PA self-efficacy.

© 2019 Published by Elsevier B.V. on behalf of Shanghai University of Sport. This is an open access article under the CC BY-NC-ND license.

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Chinese adolescents; Education levels; Factor analysis; Gender; Latent mean comparison; Measurement invariance; Physical activity self-efficacy scale

## 1. Introduction

Physical inactivity has been one of the major concerns in public health around the world. A global study identified physical inactivity as the 4th leading risk factor for global mortality.<sup>1</sup> This is consistent with the fact that globally about 31.1% of adults were physically inactive, and approximately 80.3% of adolescents (13–15 years old) fail to accumulate 60 min of daily moderate-to-vigorous physical activity (MVPA).<sup>2</sup> According to a national survey in China, there are approximately only 8% of Chinese school-aged students (6–18 years old) who perform any type of MVPA after school due to the academic

pressure.<sup>3</sup> Another recent national survey found that only 9.4% of boys and 1.9% of girls met the goal of 60 min of MVPA each day.<sup>4</sup>

While the data strongly suggest that an intervention to reverse the physical inactivity trend is necessary, it is important to first gain a better understanding of the various determinants of physical activity (PA). As a central component of Bandura's Social Cognitive Theory, self-efficacy is defined as an individual's belief in his or her ability to execute specific behaviors and is believed to be an important determinant influencing human behavior.<sup>5</sup> Indeed, numerous studies have documented that individuals' perceived self-efficacy directly influences their PA participation<sup>6,7</sup> or mediates the relationship between other variables (e.g., social support, physical environment) and PA.<sup>8–11</sup> Experimental studies targeting the enhancement of perceived self-efficacy also resulted

Peer review under responsibility of Shanghai University of Sport.

\* Corresponding author.

E-mail address: [daijun19800703@126.com](mailto:daijun19800703@126.com) (J. Dai).

in the enhancement of PA enjoyment among middle school students<sup>12</sup> or PA increase among adolescent girls.<sup>13</sup>

One of the early PA self-efficacy scales (PASES) was developed by Saunders and colleagues<sup>14</sup> using 5th grade students as participants. The PASES was based on findings from 2 studies<sup>15,16</sup> and contains 17 indicators categorized into 3 factors (i.e., support seeking, barriers, and positive alternatives). Several subsequent studies, however, failed to support this multi-dimensional scale. For example, Motl and colleagues<sup>17</sup> recruited 8th and 9th grade students as participants for testing the factorial validity of a similar self-efficacy scale. They generated a shortened 8-item unidimensional model (shortened PA self-efficacy scale, S-PASES) that demonstrated invariances for factor structure, factor loadings, and factor variances across time and groups. Dishman and colleagues<sup>18</sup> also examined the S-PASES and found this unidimensional model demonstrated acceptable fit for both African-American and Caucasian middle school girls. In addition, Dishman et al.<sup>18</sup> also found the factor structure, factor loadings, and factor variances were invariant across race and ethnicity. Bartholomew and colleagues<sup>19</sup> reexamined the 17-item PASES with 4th and 5th grade students. Neither the 3-factor, 17-item model, nor the 1-factor, 17-item model exhibited acceptable fit. Interestingly, they generated the exact same 8-item S-PASES that Motl and colleagues<sup>17</sup> obtained and found the S-PASES resulted in acceptable fit. Additionally, findings showed metric and structural invariance for this unidimensional S-PASES across Caucasian and Hispanic groups.

Although it seems the S-PASES is better represented by a unidimensional model and the scale is invariant across time and race or ethnicity, research in this area is scarce and most of the relevant research on PASES was conducted in North America. A dimensionality check of the S-PASES has rarely been examined in a different culture. A study conducted by Liang and colleagues<sup>20</sup> is among the very few that examined factor structure and reliability of S-PASES among elementary school students from Hong Kong. Their findings support the unidimensionality and reliability of Hong Kong version S-PASES. Nevertheless, there are some culture differences between Hong Kong and the Mainland of China, as Hong Kong has been a British colony for many years. Besides, colloquial Cantonese is spoken in Hong Kong, which is different from simplified Chinese used by most of students from China. To our best knowledge, however, a simplified Chinese version of S-PASES has not been tested for its construct validity using students in China as participants.

The positive effect of PA self-efficacy on PA levels prompted scholars to examine gender difference in PA self-efficacy. In general, boys perceived greater self-efficacy than girls.<sup>21,22</sup> For example, 5th grade male students reported higher PA barriers efficacy than their female counterparts.<sup>23</sup> Additionally, PA levels tend to decrease from childhood to young adulthood.<sup>24,25</sup> It is therefore important to examine the developmental patterns of PA self-efficacy across students' school years. Research in this area, however, is limited. Most of the studies targeted adults or older adults and their findings reveal a decline in PA self-efficacy or PA barrier efficacy as individuals continue to age.<sup>21,26</sup> It is not

clear how students' perceptions of their PA self-efficacy progresses across 3 education levels (i.e., elementary, middle, and high schools). Such information would be useful in designing intervention strategies to increase school age students' PA.

The examination of the aforementioned gender and education levels differences in PA self-efficacy should be based on the establishment of measurement invariance of self-efficacy scale at the scalar level. Failure to do so would make it unclear if the observed difference is due to actual gender or grade differences or if the differences are caused by different measurement structures.<sup>27</sup> To our best knowledge, however, measurement invariances of S-PASES across gender and educational levels have not been examined elsewhere.

Another area that has rarely been addressed is the latent mean comparisons between gender and education levels. Previous studies examining differences in PA self-efficacy across groups have been dominated with the traditional technique of composite scores comparisons. Latent means comparison, however, can generate more accurate results than the composite scores comparison using *t* test or analysis of variance (ANOVA) as the latent variables are free of measurement errors.<sup>28</sup> Finally, although the relationship between PA self-efficacy and PA levels was established in previous studies, the magnitudes of the direct effects comparison between these 2 variables across groups (i.e., gender, education levels) has not been examined before. Findings of such a study would provide useful information regarding the different importance levels of self-efficacy in relationship to PA across various groups.

Thus, we applied the Chinese version of S-PASES (S-PASESC) to a population-based sample of school-aged children and adolescents. The first objective of our study was to examine the dimensionality of the S-PASESC. The second objective was to test the measurement invariance of the S-PASESC across gender and levels of education. Third, after establishing the measurement invariance of S-PASESC, we tested the latent factor means difference across gender and levels of education. Fourth, we examined the relationship between self-efficacy as measured by S-PASESC and PA separated by groups of different gender and education levels. Finally, we compared the direct effects of self-efficacy on PA across gender and different education levels.

## 2. Methods

### 2.1. Participants

The target population included 5th to 11th grade public school students recruited from 7 cities located in different regions of China (i.e., Shanghai (eastern China); Guangzhou (south central China); Xi'an (northwestern China); Ürümqi (western China); Yulin (North Shaanxi province); Chuzhou (East Anhui province); and Heihe (northeast region of Heilongjiang province)). A total of 37 elementary, middle, or high schools were contacted and 24 of them agreed to participate in the study. Within each participating school, 1 to 4 classes from the same or different grade levels were randomly chosen as the target classes. The head teacher of each grade then contacted the classroom teacher to invite all students to participate. Due to the busy schedule, we excluded 9th

and 12th graders. There were at least 2 graduate teaching assistants or seniors who were sent to each school for collecting data. Data collectors also helped check missing questions and reminded the participants to complete all questions if necessary. The missing data were small ( $n=28$ ) and excluded from the data. The final data include a total of 3003 participants (49.7% boys;  $n=626$  for elementary;  $n=1329$  for middle school;  $n=1048$  for high school students) whose ages were between 10 and 20 years ( $14.2 \pm 2.0$  years, mean  $\pm$  SD). We obtained permission to conduct the study prior to the investigation from the Institutional Review Board of Shanghai University of Sport, the schools, the participants, and their parents.

## 2.2. Measurement

### 2.2.1. S-PASES

The 8-item S-PASES measuring individuals' PA self-efficacy was adopted from previous studies.<sup>17–19</sup> This scale is a shorter version adapted from a previously validated PASES that included 17 items categorized into 3 factors.<sup>14</sup> We also followed the suggestion of Ward and colleagues<sup>29</sup> by adding playing sedentary videogames as one of the sedentary activities, and replacing the item "I have the coordination I need to be physically active during my free time on most days" with "I can do active things because I know how to do them." An example of one of the S-PASES items is "I can be physically active during my free time on most days no matter how busy my day is." The S-PASES used a 5-point Likert scale ranging from 1 (*disagree a lot*) to 5 (*agree a lot*). A higher score indicates a student perceived higher PA self-efficacy. This instrument has adequate reliability as measured by McDonald's Omega (i.e.,  $\omega=0.92$  for boys,  $\omega=0.94$  for girls;  $\omega=0.92$  for elementary,  $\omega=0.92$  for middle,  $\omega=0.93$  for high school groups).

All items of the S-PASES were translated from English to Chinese (the S-PASESC) and then back translated to English by 2 independent bilingual scholars. The back translated version was very close to the original version. Consensus discussions were conducted between the 2 scholars to resolve any minor discrepancies.

### 2.2.2. Leisure time PA

The participants leisure time PA was measured using a Chinese version of the International Physical Activity Questionnaire-short form (IPAQ-S).<sup>30</sup> The IPAQ-S asks students to self-recall different types of activities (i.e., moderate, vigorous) they performed during the past 7 days. The students were provided with descriptions and examples of each activity type and were asked to record the number of days and the lasting time (in minutes) for each day of the types of activity in which they participated. Each day's moderate or vigorous PA time was added up. The total weekly MVPA time was calculated by the weekly sum of both moderate and vigorous PA time. The reliability and validity of IPAQ-S across 12 countries have been established before.<sup>30</sup>

## 2.3. Data analysis

First, SPSS Version 20.0 (IBM Corp., Armonk, NY, USA) was used to conduct descriptive statistics (mean  $\pm$  SD) together

with the univariate normality check including assessment of skewness and kurtosis values. Second, we used AMOS Version 22.0 (IBM Corp.) to conduct confirmatory factor analysis (CFA) with maximum likelihood estimation to test the unidimensionality of the scale, in which we hypothesized that all measured indicators of the self-efficacy instrument were represented by a single latent factor (marker variable with one indicator fixed to 1). We chose CFA because the S-PASES has been developed before and the current study is to verify the dimensionality of S-PASESC. The covariance matrix was modeled in CFA. All standardized factor loading within this single factor should be larger than 0.6 and statistically significant<sup>31</sup> in order to support the unidimensionality. Model fit was assessed using chi-square ( $\chi^2$ ), degree of freedom ( $df$ ), comparative fit index (CFI), goodness of fit index (GFI), adjusted GFI (AGFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). The accepted cutoffs for the values of GFI, AGFI, TLI, and CFI should be greater than 0.90; the thresholds for RMSEA and SRMR should be less than 0.06 and 0.08, respectively.<sup>32</sup>

Third, we followed Vandenberg and Lance's<sup>27</sup> suggestion to test measurement invariance with progressively restrictive stages. In Step 1, we conducted a configural invariance test to establish a baseline model across groups. The configural invariance test allows factor loadings, intercepts, as well as residuals to be estimated freely. The establishment of a configural invariance test implies the conceptual framework is the same across groups. If the data do not support the configural invariance test, the measurement invariance test should be terminated. In Step 2, we tested the metric invariance model in which all factor loadings were constrained the same. The metric invariance is a weak invariance test and the establishment of this test means that different groups responded to the indicators in the same way. In Step 3, we conducted the scalar invariance model in which the factor loading and indicator intercepts were constrained to be the same across groups. The scalar invariance test is a strong invariance test and the establishment of this test is required before the latent means can be compared across groups. The metric and scalar invariances tests were examined by assessing the change of CFI value; scholars believe the  $\chi^2$  difference test is too strict when the sample size is big.<sup>33</sup> Therefore, lack of evidence for metric and scalar invariance will be determined if the CFI change is  $> 0.01$ .<sup>34</sup>

Fourth, we compared latent mean differences between gender and levels of education (i.e., upper elementary, middle, and high schools). Specifically, a full scalar invariance model was used as the baseline. To compare latent mean between genders, we constrained the boys' group latent mean to 0 and the latent means of the girls' group was free to estimate. To calculate latent mean differences among 3 levels of education, we first set the high school group's latent mean to 0 and allowed the elementary and middle school groups' latent mean to freely estimate, which generated the latent mean comparison between high and elementary schools and between high and middle schools. The middle school group's latent mean was then constrained to 0 and the elementary and high school groups' latent means were allowed to freely estimate. This approach was

Table 1  
Scale items, means  $\pm$  SD, and psychometric properties of Chinese version of physical activity self-efficacy scale based on the total sample ( $\omega = 0.93$ ).

Scale items	Mean $\pm$ SD	Min–Max	Skewness $\pm$ SE	Kurtosis $\pm$ SE
1. I can be physically active on most days of the week.	3.46 $\pm$ 1.15	1–5	–0.23 $\pm$ 0.04	–0.70 $\pm$ 0.09
2. I can ask my parent or other adult to do physically active things with me.	3.30 $\pm$ 1.18	1–5	–0.19 $\pm$ 0.04	–0.72 $\pm$ 0.09
3. I can be physically active during my free time on most days even if I could watch TV or play (sedentary) video games instead.	3.40 $\pm$ 1.17	1–5	–0.18 $\pm$ 0.04	–0.77 $\pm$ 0.09
4. I can be physically active on most days even if it is very hot or cold outside.	3.25 $\pm$ 1.17	1–5	–0.07 $\pm$ 0.04	–0.76 $\pm$ 0.09
5. I can ask my best friend to be physically active with me on most days.	3.40 $\pm$ 1.17	1–5	–0.22 $\pm$ 0.04	–0.72 $\pm$ 0.09
6. I can be physically active even at home.	3.60 $\pm$ 1.12	1–5	–0.41 $\pm$ 0.04	–0.48 $\pm$ 0.09
7. I can do active things because I know how to do them.	3.34 $\pm$ 1.12	1–5	–0.13 $\pm$ 0.04	–0.55 $\pm$ 0.09
8. I can be physically active during my free time on most days no matter how busy my day is.	3.33 $\pm$ 1.19	1–5	–0.14 $\pm$ 0.04	–0.80 $\pm$ 0.09

Note:  $\omega$  = McDonald's Omega.

intended to examine the latent mean difference between elementary and middle schools. We used the value of the critical ratio (CR) to assess latent mean differences. CR is calculated by parameter estimate divided by its standard error, which tests whether the coefficient is significantly different from 0. A CR value larger than 1.96 indicates statistically significant differences in the latent means.<sup>35</sup> A positive CR implies that the comparison group has higher latent mean than the reference group. Conversely, a negative CR suggests that the comparison group's latent mean is smaller than the reference group.

Finally, a naïve bootstrapping method in AMOS ( $n = 5000$  resample) was used to examine the direct effects of efficacy barriers on PA by gender and different education levels. The AMOS syntax was developed within the function of the user-defined estimate in order to compare the direct effects of efficacy barriers on PA between boys and girls as well as among different education levels. The naïve bootstrapping method could generate a 95% confidence interval (CI) for both percentile and bias-corrected (BC) bootstrapping in which the statistical significance was determined if the 95% interval does not include 0.<sup>36</sup>

### 3. Results

#### 3.1. Descriptive analysis of S-PASESC

Table 1 shows the characteristics of all items of the scale based on the total sample. As can be seen from Table 1, the

internal consistency ( $\omega = 0.93$ ) was very good. All items' skewness and kurtosis were between +2 and –2,<sup>37</sup> indicating univariate normality test of the data was assumed.

#### 3.2. Factorial validity of S-PASESC

Before CFA was conducted to test the factorial validity of the scale, we examined the multivariate normality test. Findings showed that the multivariate kurtosis CR (101.394) is bigger than the recommended value of 5.<sup>37</sup> The violation of multivariate normality would increase the value of  $\chi^2$ , which subsequently would negatively influence fit indices of the model. We then used the Bollen-Stine bootstrap method to adjust the inflated  $\chi^2$  value to improve the overall model fit.<sup>38</sup> The results of modification indices suggested that allowing 6 items to correlate would greatly decrease the  $\chi^2$  value (i.e., Items 2 and 4, Items 6 and 7, and Items 7 and 8) for total, 2 gender, and 3 education level groups. We therefore allowed these residuals to correlate. Table 2 lists all indicators' standardized factor loading by total sample and different groups. All items' standardized factor loadings were above 0.60, supporting the unidimensionality characteristics of this scale. As also can be seen from Table 2, after the Bollen-Stine bootstrap correction and allowing 6 items' residuals to correlate, fit indices for all models (i.e., total sample, different gender, and level of education groups) were very good. Table 3 includes covariance and correlation matrices of S-PASESC based on the total sample.

Table 2  
Standardized factor loading of indicator and fit indices for total and different groups' models.

	Standardized factor loading of indicators*								Fit indices of models <sup>#</sup>								
	1	2	3	4	5	6	7	8	$\chi^2$	<i>df</i>	<i>p</i>	<i>CFI</i>	<i>GFI</i>	<i>AGFI</i>	<i>TLI</i>	<i>RMSEA</i> (90%CI)	<i>SRMR</i>
Total	0.788	0.654	0.815	0.830	0.703	0.652	0.631	0.772	31.92	17	<0.001	1	1	1	1	0.02(0.05–0.06)	0.02
Boys	0.774	0.621	0.792	0.814	0.705	0.603	0.619	0.762	32.46	17	<0.001	1	0.99	1	1	0.02(0.04–0.07)	0.02
Girls	0.796	0.708	0.831	0.836	0.696	0.694	0.636	0.773	29.97	17	<0.001	1	1	1	1	0.02(0.05–0.07)	0.02
Elementary	0.784	0.678	0.802	0.827	0.661	0.642	0.630	0.728	28.98	17	<0.001	1	0.99	1	0.99	0.03(0.03–0.07)	0.02
Middle	0.777	0.657	0.795	0.826	0.685	0.644	0.595	0.767	30.28	17	<0.001	1	0.99	1	1	0.02(0.03–0.06)	0.02
High	0.784	0.609	0.835	0.824	0.737	0.656	0.672	0.788	33.23	17	<0.001	1	0.99	1	0.99	0.03(0.06–0.09)	0.02

Notes: elementary, middle, and high refer to Grades 5th–6th, 7th–9th, and 10th–12th, respectively.

\*  $p < 0.001$  for all indicators.

<sup>#</sup> Fit indices were adjusted after Bollen-Stine bootstrapping correction ( $n = 2000$ ) as well as residuals correlation of 6 items (i.e., 2 and 4, 6 and 7, and 7 and 8).

Abbreviations:  $\chi^2$  = chi-square; *AGFI* = adjusted goodness of fit index; *CFI* = comparative fit index; *CI* = confidence interval; *df* = degree of freedom; *GFI* = goodness of fit index; *RMSEA* = root mean square error of approximation; *SRMR* = standardized root mean square residual; *TLI* = Tuck-Lewis index.

Table 3  
Covariance and correlation matrices of Chinese version of physical activity self-efficacy scale based on the total sample.

	E1	E2	E3	E4	E5	E6	E7	E8
E1	1.33	0.54	0.65	0.64	0.54	0.52	0.47	0.61
E2	0.73	1.40	0.53	0.46	0.49	0.43	0.43	0.46
E3	0.88	0.73	1.37	0.70	0.55	0.51	0.49	0.62
E4	0.86	0.64	0.95	1.37	0.58	0.51	0.52	0.65
E5	0.73	0.68	0.76	0.80	1.37	0.48	0.47	0.55
E6	0.68	0.57	0.67	0.67	0.63	1.26	0.58	0.55
E7	0.60	0.57	0.64	0.68	0.61	0.72	1.24	0.60
E8	0.83	0.65	0.87	0.89	0.76	0.74	0.79	1.41

Note: Above diagonal is the correlation matrix.

### 3.3. Measurement invariance

Table 4 includes the results for measurement invariances of the S-PASESC across gender and 3 education levels. The progressive gender invariance test showed that configural invariance model across gender demonstrated acceptable model fit. The metric invariance test constrained factor loading to be equal across genders. Findings of this test showed that the model fits the data well. Additionally, the change of *CFI* between configural and metric invariance tests is within the threshold of 0.01, supporting the metric invariance across gender. Finally, the scalar invariance test also demonstrated that the indicators' intercepts were invariant across genders, as the *CFI* change between the scalar and metric invariance tests was

Table 4  
Measurement invariance tests of shortened Chinese version of physical activity self-efficacy scale across genders and education levels.

	$\chi^2$	<i>df</i>	<i>CFI</i>	<i>RMSEA</i> (90%CI)	$\Delta CFI$
<b>Gender (2 levels)</b>					
Configural invariance	210.030	34	0.986	0.042 (0.036–0.047)	
Full metric invariance	225.609	41	0.985	0.039 (0.034–0.044)	0.003
Full scalar invariance	363.879	49	0.975	0.046 (0.042–0.051)	0.007
<b>Education (3 levels)</b>					
Configural invariance	219.072	51	0.986	0.033 (0.029–0.038)	
Metric (measurement weight)	243.696	65	0.986	0.030 (0.026–0.034)	0.003
Scalar (intercepts)	361.955	81	0.977	0.034 (0.030–0.038)	0.004

Abbreviations: *CFI* = comparative fit index; CI = confidence interval; *df* = degree of freedom; *RMSEA* = root mean square error of approximation.

Table 5  
Direct effects of self-efficacy on physical activity by gender and education level as well as the direct effects comparisons among different groups.

Parameter	Point estimate	Bootstrapping					
		Bias corrected 95%CI			Percentile 95%CI		
		Lower	Upper	<i>p</i>	Lower	Upper	<i>p</i>
Self-efficacy → weekly MVPA (boys)	150.590	128.017	174.038	<0.001	127.680	173.838	<0.001
Self-efficacy → weekly MVPA (girls)	106.850	86.908	127.678	<0.001	86.815	127.610	<0.001
Boy and girl difference	43.740	13.552	74.856	<0.01	13.420	74.763	<0.01
Self-efficacy → weekly MVPA (Elementary)	128.613	90.997	167.671	<0.001	90.707	167.368	<0.001
Self-efficacy → weekly MVPA (Middle)	126.204	102.815	148.896	<0.001	103.276	149.210	<0.001
Self-efficacy → weekly MVPA (High)	114.850	95.197	136.443	<0.001	94.775	135.939	<0.001
Elementary and middle school difference	2.409	−40.260	45.388	0.921	−40.377	45.300	0.929
Elementary and high school difference	13.763	−28.177	59.030	0.507	−28.782	58.160	0.523
Middle and high school difference	11.354	−19.706	40.055	0.516	−18.954	41.025	0.469

Note: 5000 bootstrap samples.

Abbreviations: CI = confidence interval; PA = physical activity; MVPA = moderate-to-vigorous physical activity.

not greater than 0.01. The measurement invariance tests across levels of education also had similar findings. The configural model was acceptable. In addition, the metric and scalar invariance tests found that the factor loadings and the indicators' intercepts were invariance across 3 levels of education.

### 3.4. Latent mean differences

Based on the establishment of the full scalar invariance across both gender and education levels, we can compare the latent mean differences across these groups. Findings of the latent mean comparisons between genders showed that girls ( $3.19 \pm 0.86$ ) had significantly lower PA self-efficacy than boys ( $3.49 \pm 0.85$ ) ( $CR = -7.84$ ;  $p < 0.001$ ; Cohen's  $d = 0.35$ ). The results of the latent mean differences among education levels revealed that elementary students ( $3.69 \pm 0.9$ ) had a higher score than the middle school ( $3.31 \pm 0.84$ ) ( $CR = 6.75$ ;  $p < 0.001$ ; Cohen's  $d = 0.43$ ) and high school students ( $3.16 \pm 0.82$ ) ( $CR = 10.05$ ;  $p < 0.001$ ; Cohen's  $d = 0.61$ ); middle school students had a higher score than the high school students ( $CR = 4.53$ ;  $p < 0.001$ ; Cohen's  $d = 0.18$ ).

### 3.5. The relationship between self-efficacy and PA as well as the moderating effects of gender and education levels

Table 5 includes the direct effects of self-efficacy on PA by gender and education levels. Clearly, self-efficacy directly

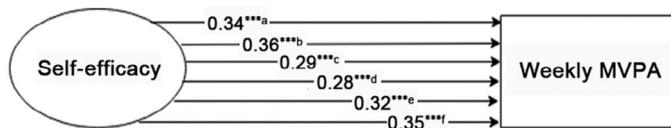


Fig. 1. Standardized regression coefficients for the direct relationship between self-efficacy and weekly MVPA separated by total, gender, and different grade levels. \*\*\*  $p < 0.001$ ; <sup>a</sup> total sample; <sup>b</sup> boys; <sup>c</sup> girls; <sup>d</sup> elementary school students; <sup>e</sup> middle school students; <sup>f</sup> high school students; MVPA = moderate-to-vigorous physical activity.

influenced PA for boys (e.g., 95% BC CI: 128.017–174.038), girls (e.g., 95% BC CI: 86.908–127.678), and elementary (e.g., 95% BC CI: 90.997–167.671), middle (e.g., 95% BC CI: 102.815–148.896), and high school (e.g., 95% BC CI: 95.197–136.443) students. Fig. 1 lists standardized regression coefficients of the relationship between self-efficacy and PA by total and different groups. In addition, as can be seen from Table 5, although self-efficacy is significantly related to PA for both genders, the direct effect of self-efficacy on PA is significantly greater for boys than for girls (direct effect gender difference: 95% BC CI: 13.552–74.856). The direct effect comparisons among different education levels revealed no significant difference (90%CI includes 0 for both BC and percentile).

#### 4. Discussion

According to Bandura's theory,<sup>5</sup> self-efficacy is a prerequisite for the behavior change. In this study, we administered the Chinese version of the S-PASES to a population-based sample of school-aged students in an effort to explore the scale dimensionality, measurement invariance, latent mean differences, relationship with PA, as well as the comparisons of the direct effects of self-efficacy on PA between genders and education levels. The 1-factor CFA test showed that this 8-item S-PASESC exhibited excellent model fit and demonstrated unidimensionality for the total sample as well as by different gender and education levels.

The PASES was initially developed by Saunders and her colleagues.<sup>14</sup> Their exploratory factor analysis generated 3 factors (i.e., support seeking, barriers, and positive alternatives) with a total of 17 indicators. However, some of the items' factor loadings were below 0.5 and only 3 items' factor loadings were above 0.7. Two studies then tested this scale again and both studies retained the same 8-item instrument (S-PASES) forming a single factor. In addition, these studies found the S-PASES demonstrated acceptable model fit.<sup>17,19</sup> Another study also found the S-PASES demonstrated acceptable model fit.<sup>18</sup> Our finding showed that this unidimensional S-PASES can be generalized to Chinese school-aged students. Specifically, factor loadings of all items of the S-PASESC were above 0.60 and the data fit the model very well. It seems that Chinese students' PA self-efficacy can be measured with this single-factor instrument. This unidimensional S-PASESC is especially useful in reducing respondent burden in future multivariable data collection study.

##### 4.1. Factor invariance

Measurement invariance testing for the S-PASES has been less studied. Scholars typically assumed that the instrument

being used operates the same way and contains the same construct across different groups.<sup>39</sup> To our best knowledge, only 3 studies have conducted measurement invariance tests for the S-PASES. Findings showed that the S-PASES was invariant across time, cohorts, or races.<sup>17–19</sup> Nevertheless, these studies either only recruited girls or a single education level of students (e.g., elementary or middle school) as participants. Thus, our large-scale population-based sample recruited from multiple sites in China and across 3 education levels are of particular importance. The current study established configural, full metric, and full scalar invariances across gender and levels of education. This finding implies that the S-PASESC has the same construct with each item associated equally with this factor for both genders and students from different education levels. Researchers employing S-PASESC in future studies can compare the PA self-efficacy meaningfully across gender and levels of education.

##### 4.2. Latent mean differences

The establishment of scalar invariance (equal factor loading and intercepts) across both gender and education levels indicated that the mean differences of the S-PASESC can be compared directly. Our findings revealed that boys exhibited higher latent means of PA self-efficacy than girls. This result is consistent with findings from previous studies testing the gender differences in PA self-efficacy using traditional *t* test or ANOVA.<sup>22,23</sup> Results may suggest the necessity to improve the PA self-efficacy among girls in order to increase their PA participation.

Findings also showed that high school students had lower latent means of self-efficacy than middle or elementary students, and middle school students also had lower latent means than elementary students in S-PASESC. The results also correspond with previous studies examining age differences in barrier efficacy to PA, efficacy to adhere to PA in various situations, or capabilities efficacy among adults or older adults using composite mean comparisons technique.<sup>21,26,40,41</sup> For example, Anderson-Bill and colleagues<sup>26</sup> found that the 40-year-old group reported higher barrier efficacy to PA than the 60-year-old group. Our study extended previous research by demonstrating that starting from the upper level elementary school years, students' perceived PA self-efficacy tend to decrease as they continue to age. While the decrease in PA self-efficacy among older adults might be due to the increase in functional limitations associated with aging, the mechanism of the negative relationship between ageing and self-efficacy among Chinese adolescents is not clear and needs further explorations.

##### 4.3. The relationship between self-efficacy and PA as well as the moderating effects of gender and education levels

Bandura<sup>42</sup> proposed that self-efficacy has a potential to influence exercise adoption, and numerous studies have established the relationship between the variables of gender and educational level.<sup>43,44</sup> For example, the composite scores of self-efficacy measured with Hong Kong version of S-PASES was found correlated with children from elementary schools.<sup>20</sup>

Findings from the current study revealed that self-efficacy positively related to leisure time PA for boys, girls, elementary, middle, and high school students in China. These findings demonstrate that self-efficacy is an important variable to Chinese school-age students that needs to be included in future intervention programs.

Although self-efficacy positively related to PA for both boys and girls, a moderating effect was found for gender such that the relationship between self-efficacy and PA was stronger for boys than for girls. Previous studies in this area generated conflicting findings. For example, Spence and colleagues<sup>45</sup> found the relationship between self-efficacy and PA was stronger in 7th–10th grade female students than their male counterparts. Our finding, however, is consistent with a study using high school students as participants.<sup>46</sup> Further examination of the moderating effect of gender within each education levels revealed that self-efficacy related to PA for both boys and girls within each education levels. Additionally, gender only demonstrated its moderating effect in middle school group such that the relationship between the 2 variables is stronger in boys than in girls (data not shown). It seems that self-efficacy is important among both Chinese boys and girls from all 3 education levels (i.e., elementary, middle, and high schools) in influencing their PA. Within middle school, this importance is especially prominent in boys. These findings warrant further investigation to better understand how gender could moderate the relationship between self-efficacy and PA across all levels of education.

Self-efficacy was found positively related to PA for elementary, middle, and high school students. No significant moderating effects of 3 education levels were found. Although students' perception in self-efficacy continues to decrease across 3 education levels, self-efficacy demonstrated equal importance for elementary, middle, and high school students. To account for gender difference, we examined the moderating effects of 3 education levels for boys and girls separately and the findings did not change (data not shown). To our best knowledge, this is the first study comparing the direct effects of self-efficacy on PA across 3 education levels. Findings provide evidence to necessitate the change of self-efficacy in all education levels of Chinese students in order to increase their PA.

Several limitations should be noted. First, the original 17-item PASES was not examined in the current study. Future studies should explore the factorial validity of the PASES with participants from different cultures. It would be interesting to test if the same 8-item result would be retained for a shortened unidimensional scale. Second, the participants in our study were all from China. Future study should examine measurement equivalence of the S-PASES across different cultures. Third, it is possible that culture differences may exist among the 7 different cities in which we recruited the participants. Culture differences of various regions and cities within China should be included in future validation study of the S-PASESC.

Despite these limitations, this research, as the first factorial validity and invariance tests of the S-PASESC with participants recruiting from different parts of China made several contributions. First, the study confirmed the unidimensionality

of the S-PASESC. Second, the S-PASESC was found invariant at the full scalar level across gender and different education levels. Third, based on the latent mean comparisons, boys demonstrated higher PA self-efficacy and students' PA self-efficacy tends to decrease from upper elementary to high school levels. Fourth, students' PA self-efficacy measured by S-PASESC positively related to their weekly MVPA by total sample and groups of different genders or education levels. Finally, compared with female students in middle school, PA self-efficacy seems to be especially important in their male counterparts as evidenced by the stronger positive relationship between PA self-efficacy and PA in boys.

## 5. Conclusion

The current study supports the unidimensionality of the S-PASESC. In addition, the S-PASESC is invariant across gender and 3 levels of education (i.e., upper elementary, middle school, and high school) at the configural, full metric, as well as full scalar levels. Researchers can compare the concept of PA self-efficacy meaningfully across gender and 3 education levels for Chinese students. Additionally, results from latent mean comparison indicated that boys reported higher PA self-efficacy than girls, high school students exhibited lower PA self-efficacy than middle or elementary school students, and middle school students reported lower PA self-efficacy than elementary school students. Further, self-efficacy as measured by S-PASESC was found related to individuals weekly MVPA by different gender and education levels. Finally, among middle school students, gender moderated the relationship between PA self-efficacy and PA such that this relationship is stronger in boys than in girls.

## Acknowledgment

This work was supported by a grant from the National Social Science Foundation of China (No. 13CTY031). We thank all the participating schools and students for their contribution to the study.

## Authors' contributions

HC drafted the article, conducted data analysis, and designed the study and interpreted the findings; JD collected data, designed the study, and interpreted the findings; YG provided critical revision of the study. All authors have read and approved the final version of the manuscript, and agree with the order of presentation of the authors.

## Competing interests

The authors declare that they have no competing interests.

## Appendix

Chinese version of shortened PASES (S-PASEESC)

1. 在每周大部分天数中, 我会锻炼身体。
2. 我可以要求父母或其他成年人和我一起锻炼身体。

3. 即使我可以看电视或玩电子游戏，我仍然可以在每周大部分天数中的空闲时间锻炼身体。
4. 即使室外非常炎热或寒冷，我仍然可以在每周大部分天数中锻炼身体。
5. 我可以要求我最好的朋友和我一起每周大部分天数中一起锻炼身体。
6. 即使在家里我仍然可以锻炼身体。
7. 我可以锻炼身体是因为我知道如何锻炼。
8. 不管我有多忙，我仍然可以在每周大部分天数中的空闲时间锻炼身体。

## References

1. World Health Organization (WHO). *Global recommendations on physical activity for health*. Geneva: WHO; 2010.
2. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, et al. Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet* 2012;**380**:247–57.
3. Tudor-Locke C, Ainsworth BE, Adair LS, Du S, Popkin BM. Physical activity and inactivity in Chinese school-aged youth: the China Health and Nutrition Survey. *Int J Obes* 2003;**27**:1093–9.
4. Wang C, Chen P, Zhuang J. A national survey of physical activity and sedentary behavior of Chinese city children and youth using accelerometers. *Res Q Exerc Sport* 2013;**84**(Suppl. 2):S12–28.
5. Bandura A. *Self-efficacy: the exercise of control*. New York, NY: W. H. Freeman; 1997.
6. Rhodes RE, Mark R. Social cognitive models. In: Acevedo EO, editor. *The Oxford handbook of exercise psychology*. New York, NY: Oxford University Press; 2012.p.273–94.
7. Trost SG, Pate RR, Saunders R, Ward DS, Dowda M, Felton G. A prospective study of the determinants of physical activity in rural fifth-grade children. *Prev Med* 1997;**26**:257–63.
8. Peterson MS, Lawman HG, Wilson DK, Fairchild A, Van Horn ML. The association of self-efficacy and parent social support on physical activity in male and female adolescents. *Health Psychol* 2013;**32**:666–74.
9. Silva P, Lott R, Mota J, Welk G. Direct and indirect effects of social support on youth physical activity behavior. *Pediatr Exerc Sci* 2014;**26**:86–94.
10. Motl RW, Dishman RK, Saunders RP, Dowda M, Pate RR. Perceptions of physical and social environment variables and self-efficacy as correlates of self-reported physical activity among adolescent girls. *J Pediatr Psychol* 2007;**32**:6–12.
11. Motl RW, Dishman RK, Ward DS, Saunders RP, Dowda M, Felton G, et al. Perceived physical environment and physical activity across one year among adolescent girls: self-efficacy as a possible mediator? *J Adolesc Health* 2005;**37**:403–8.
12. Hu L, Cheng S, Lu J, Zhu L, Chen L. Self-efficacy manipulation influences physical activity enjoyment in Chinese adolescents. *Pediatr Exerc Sci* 2016;**28**:143–51.
13. Dishman RK, Motl RW, Saunders R, Felton G, Ward DS, Dowda M, et al. Self-efficacy partially mediates the effect of a school-based physical-activity intervention among adolescent girls. *Prev Med* 2004;**38**:628–36.
14. Saunders RP, Pate RR, Felton G, Dowda M, Weinrich MC, Ward DS, et al. Development of questionnaires to measure psychosocial influences on children's physical activity. *Prev Med* 1997;**26**:241–7.
15. Reynolds KD, Killen JD, Bryson SW, Maron DJ, Taylor CB, Maccoby N, et al. Psychosocial predictors of physical activity in adolescents. *Prev Med* 1990;**19**:541–51.
16. Sallis JF, Simons-Morton BG, Stone EJ, Corbin CB, Epstein LH, Faucette N, et al. Determinants of physical activity and interventions in youth. *Med Sci Sports Exerc* 1992;**24**:248–57.
17. Motl RW, Dishman RK, Trost SG, Saunders RP, Dowda M, Felton G, et al. Factorial validity and invariance of questionnaires measuring social-cognitive determinants of physical activity among adolescent girls. *Prev Med* 2000;**31**:584–94.
18. Dishman RK, Motl RW, Saunders R, Dowda M, Felton G, Ward DS, et al. Factorial invariance and latent mean structure of questionnaires measuring social-cognitive determinants of physical activity among black and white adolescent girls. *Prev Med* 2002;**34**:100–8.
19. Bartholomew JB, Loukas A, Jowers EM, Allua S. Validation of the physical activity self-efficacy scale: testing measurement invariance between Hispanic and Caucasian children. *J Phys Act Health* 2006;**3**:70–8.
20. Liang Y, Lau PW, Huang WY, Maddison R, Baranowski T. Validity and reliability of questionnaires measuring physical activity self-efficacy, enjoyment, social support among Hong Kong Chinese children. *Prev Med Rep* 2014;**1**:48–52.
21. Netz Y, Raviv S. Age differences in motivational orientation toward physical activity: an application of social-cognitive theory. *J Psychol* 2004;**138**:35–48.
22. Wu SY, Pender N, Noureddine S. Gender differences in the psychosocial and cognitive correlates of physical activity among Taiwanese adolescents: a structural equation modeling approach. *Int J Behav Med* 2003;**10**:93–105.
23. Trost SG, Pate RR, Dowda M, Saunders R, Ward DS, Felton G. Gender differences in physical activity and determinants of physical activity in rural fifth grade children. *J Sch Health* 1996;**66**:145–50.
24. Kwon S, Janz KF, Letuchy EM, Burns TL, Levy SM. Developmental trajectories of physical activity, sports, and television viewing during childhood to young adulthood: Iowa Bone Development Study. *JAMA Pediatr* 2015;**169**:666–72.
25. Van Dijk ML, Savelberg HH, Verboon P, Kirschner PA, De Groot RH. Decline in physical activity during adolescence is not associated with changes in mental health. *BMC Public Health* 2016;**16**:300. doi:10.1186/s12889-016-2983-3.
26. Anderson-Bill ES, Winett RA, Wojcik JR, Williams DM. Aging and the social cognitive determinants of physical activity behavior and behavior change: evidence from the guide to health trial. *J Aging Res* 2011;**2011**:505928. doi:10.4061/2011/505928.
27. 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**:4–70.
28. Ployhart RE, Oswald FL. Applications of mean and covariance structure analysis: integrating correlational and experimental approaches. *Organ Res Methods* 2004;**7**:27–65.
29. Ward DS, Saunders RP, Pate R. *Physical activity interventions in children and adolescents*. Champaign, IL: Human Kinetics; 2007.
30. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;**35**:1381–95.
31. Anderson JC, Gerbing DW. Structural equation modeling in practice: a review and recommended two-step approach. *Psychol Bull* 1988;**103**:411–23.
32. Hu LT, Bentler PM. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol Methods* 1998;**3**:424–53.
33. Quintana SM, Maxwell SE. Implications of recent developments in structural equation modeling for counseling psychology. *Couns Psychol* 1999;**27**:485–527.
34. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling* 2002;**9**:233–55.
35. Byrne BM. *Structural equation modeling with AMOS: basic concepts, applications, and programming*. Abingdon-on-Thames: Routledge; 2013.
36. MacKinnon DP, Lockwood CM, Williams J. Confidence limits for the indirect effect: distribution of the product and resampling methods. *Multivariate Behav Res* 2004;**39**:99. doi:10.1207/s15327906mbr3901\_4.
37. Kline RB. *Principles and practice of structural equation modeling*. 4th ed. New York, NY: The Guilford Press; 2016.
38. Bollen KA, Stine RA. Bootstrapping goodness-of-fit measures in structural equation models. *Sociol Methods Res* 1992;**21**:205–29.
39. Byrne BM. Testing for multigroup invariance using AMOS graphics: a road less traveled. *Struct Equ Modeling* 2004;**11**:272–300.

40. McAuley E, Lox C, Duncan TE. Long-term maintenance of exercise, self-efficacy, and physiological change in older adults. *J Gerontol* 1993;**48**:218–24.
41. Wilcox S, Storandt M. Relations among age, exercise, and psychological variables in a community sample of women. *Health Psychol* 1996;**15**:110–3.
42. Bandura A. Human agency in social cognitive theory. *Am Psychol* 1989;**44**:1175–84.
43. Dwyer JJ, Chulak T, Maitland S, Allison KR, Lysy DC, Faulkner GE, et al. Adolescents' self-efficacy to overcome barriers to physical activity scale. *Res Q Exerc Sport* 2012;**83**:513–21.
44. Wilcox S, Sharpe PA, Hutto BE, Granner ML. Psychometric properties of the self-efficacy for exercise questionnaire in a diverse sample of men and women. *J Phys Act Health* 2005;**2**:285–97.
45. Spence JC, Blanchard CM, Clark M, Plotnikoff RC, Storey KE, McCargar L. The role of self-efficacy in explaining gender differences in physical activity among adolescents: a multilevel analysis. *J Phys Act Health* 2010;**7**:176–83.
46. Allison KR, Dwyer JJ, Makin S. Perceived barriers to physical activity among high school students. *Prev Med* 1999;**28**:608–15.