

## Original Article



## Validation of the Physical Activity Questionnaire for Older Children (PAQ-C) among Chinese Children \*

WANG Jing Jing<sup>1</sup>, BARANOWSKI Tom<sup>2</sup>, LAU WC Patrick<sup>1, #</sup>, CHEN Tzu An<sup>2</sup>,  
and PITKETHLY Amanda Jane<sup>1</sup>

1. Department of Physical Education, Faculty of Social Sciences, Hong Kong Baptist University, Hong Kong, China; 2. Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston 77030, Texas, USA

### Abstract

**Objective** This study initially validates the Chinese version of the Physical Activity Questionnaire for Older Children (PAQ-C), which has been identified as a potentially valid instrument to assess moderate-to-vigorous physical activity (MVPA) in children among diverse racial groups.

**Methods** The psychometric properties of the PAQ-C with 742 Hong Kong Chinese children were assessed with the scale's internal consistency, reliability, test-retest reliability, confirmatory factory analysis (CFA) in the overall sample, and multistep invariance tests across gender groups as well as convergent validity with body mass index (BMI), and an accelerometry-based MVPA.

**Results** The Cronbach alpha coefficient ( $\alpha=0.79$ ), composite reliability value ( $\rho=0.81$ ), and the intraclass correlation coefficient ( $\alpha=0.82$ ) indicate the satisfactory reliability of the PAQ-C score. The CFA indicated data fit a single factor model, suggesting that the PAQ-C measures only one construct, on MVPA over the previous 7 days. The multiple-group CFAs suggested that the factor loadings and variances and covariances of the PAQ-C measurement model were invariant across gender groups. The PAQ-C score was related to accelerometry-based MVPA ( $r=0.33$ ) and inversely related to BMI ( $r=-0.18$ ).

**Conclusion** This study demonstrates the reliability and validity of the PAQ-C in Chinese children.

**Key words:** Physical activity; Measurement; Children; Reliability; Validity

*Biomed Environ Sci, 2016; 29(3): 177-186*

*doi: 10.3967/bes2016.022*

*ISSN: 0895-3988*

*www.besjournal.com (full text)*

*CN: 11-2816/Q*

*Copyright ©2016 by China CDC*

### INTRODUCTION

There is conclusive evidence that regular physical activity (PA) is positively related to cardiovascular fitness, muscle strength, and lower risk of obesity and diabetes<sup>[1]</sup>. The World Health Organization (WHO) has identified physical inactivity as the fourth leading risk factor for global mortality causing an estimated 3.2 million or 6%

deaths globally<sup>[2]</sup>. PA and physical fitness track from childhood and adolescence into and throughout the adulthood<sup>[3]</sup>. The level of PA in childhood has been regarded as one of the best predictors for PA in later life<sup>[4]</sup>. Clearly, valid assessment is crucial to determine the relationships between PA and specific health benefits and to evaluate PA interventions for children and adolescents.

However, the accuracy of PA assessment is

\*The study was funded by the General Research Fund (GRF) from Research Grants Council of Hong Kong (to PWCL, project number: GRF 244913).

#Correspondence should be addressed to LAU WC Patrick, E-mail: wclau@hkbu.edu.hk; Tel: (852) 3411 5634; Fax: (852) 3411 5757.

Biographical note of the first author: WANG Jing Jing, female, born in 1985, PhD, specializing in physical activity and behavioural modification.

inversely related to practicality. The most accurate measures of PA (e.g., indirect calorimetry) are considered invasive and impractical for field-based studies. Accelerometry-based assessments are accurate, but expensive for use in larger populations, and encounter adherence issues (e.g. uncomfortable to wear, forgetting to wear the device, social embarrassment), especially among children<sup>[5]</sup>. Self-report questionnaires remain the most widely accepted and utilized methods in large populations as they provide low cost to investigators and low burden to participants. Moreover, contextual items on questionnaires provide information regarding various types of activities which is not available through objective measurement<sup>[6]</sup>.

Validated self-report PA measures for use in Chinese pediatric populations are limited. A Chinese 7-day physical activity recall questionnaire, tested among 92 4-6<sup>th</sup> grade children in Beijing, demonstrated acceptable test-retest reliability (kappa value ranged from 0.46 to 0.79) but moderate validity only among boys ( $r$  was 0.46, 0.38 for different activities)<sup>[7]</sup>. A modified Chinese version of the Children's Leisure Activities Study Survey (CLASS) determined reliable estimates of PA patterns among Hong Kong Chinese children aged 9 to 12 years<sup>[8]</sup>. However, the correlation with the accelerometer measure was non-significant for boys. In both these questionnaires reports of frequency (times) and duration (min) were required. However, children may have trouble recalling the frequency of activities and have limited ability to accurately report the duration of specific activities<sup>[9]</sup>. The memory and estimation biases in PA questionnaires have to be reduced to acceptable level for children<sup>[10]</sup>.

The Physical Activity Questionnaire for Older Children (PAQ-C) has been identified as a potentially valid instrument for use with children and adolescents<sup>[11]</sup>. The PAQ-C is a self-administered, 7-day recall questionnaire for children aged 8 to 14 years consisting of ten items, nine of which are structured to discern moderate-to-vigorous PA (MVPA). The scale uses a 5-point Likert scale with higher scores indicating higher PA levels<sup>[12]</sup>. The PAQ-C has been tested among several English speaking populations i.e. British, African American, European American, and Canadian<sup>[13-15]</sup>. Good internal consistency (Cronbach's  $\alpha=0.76$  to 0.84) and test-retest reliability ( $r=0.75$  to 0.82) have been documented. The construct validity of the PAQ-C has been tested against other questionnaires, as well as convergent validity which has been tested

against aspects of cardiovascular fitness<sup>[12,16]</sup>. Inconsistent validation findings suggest the PAQ-C requires refinement before use with diverse racial groups<sup>[15]</sup>. Language and cultural differences may also affect English language questionnaires when translated into Chinese<sup>[17]</sup>. Although the Chinese version of the PAQ-C has been applied to measure self-reported PA in China<sup>[18]</sup>, no existing studies have assessed the reliability and validity of the Chinese version.

The purpose of this current research was to provide reliability and validity for the Chinese version of the PAQ-C. We examined the general score psychometrics, the validity of the factor structure using confirmatory factor analysis (CFA), and convergent validity with body mass index (BMI) and an objective accelerometer measure of PA.

## METHODS

### *Participants*

Six Hong Kong primary schools that approved to participate in the study were included. The schools were located in two Hong Kong districts (New Territories and Hong Kong Island), which varied in student socio-economic status (SES). A total of 798 students (445 boys and 353 girls) aged 8 to 13 years who provided written informed consent were recruited from Grades 4-6 from May 2014 to February 2015. A subsample of 463 children (256 boys and 207 girls) participated in the 7-day accelerometer protocol. The study was approved by the Hong Kong Baptist University Committee on the Use of Human and Animal Subjects in Teaching and Research.

### *Measures*

**Physical Activity Measured by the PAQ-C** PA was assessed using the PAQ-C, which consists of nine computable items. The tenth item identifies whether sickness or other events may prevent the child from participating in regular PA and is not included in the calculation of activity scores. Of the nine computable PAQ-C items, the first provides a checklist of 22 common leisure and sport activities, followed by two supplemental blank spaces for participants to enter other activities not included in the list. The mean of all activities ('no' activity being 1, '7 times or more' being 5) on the activity checklist is calculated to form a composite score for item 1. The remaining eight questions assess activities conducted at particular

segmented times during the day [e.g. physical education (PE) class, recess, lunchtime, after school, evening, weekends] or day of week summary. The overall PAQ-C score is a composite value that calculates the mean of the nine item scores.

The translation of the questionnaires from English into Cantonese consisted of three separate forward translations by native speakers of the target language, and subsequently back translated by English speakers. Discussions with local experts in sport and exercise disciplines on the cultural adaptations to the list of activities resulted in 'ice skating' changed to 'in-line skating' and 'football' to 'soccer'. Uncommon activities were removed (street hockey, cross-country skiing and ice hockey/ringette), while five activities regularly conducted by Hong Kong Children [squash, tennis, table tennis, hiking, and martial arts (taekwondo, Judo, Kung Fu etc.)] were added. Prior to data collection, five Hong Kong Chinese students were invited to test the comprehensibility of the questionnaire<sup>[19]</sup> and minor wording revisions were made based on their feedback. The Chinese version of the PAQ-C is attached as the supplemental material.

#### **Physical Activity Measured by Accelerometer**

ActiGraph accelerometers GT3X (AG: Actigraph LCC, Fort Walton Beach, FL) were used to assess the convergent validity of the PAQ-C score. AGs have been widely used to objectively measure PA level and have demonstrated high reliability and validity among children<sup>[20]</sup>. The acceleration of PA is recorded by piezoelectric transducers and microprocessors into digital signals 'counts' at pre-selected epochs. In the present study, 5-sec epochs were set. Activity counts were summed as per minute interval. Based on recent recommendations<sup>[21]</sup>, cut-off points developed by Evenson et al.<sup>[22]</sup> were used to determine the intensity of moderate (MPA  $\geq 2296$  counts per min) and vigorous physical activity (VPA,  $\geq 4012$  counts per min) in children. Children were asked to wear AGs for 7 consecutive days. For analysis, extreme values ( $>20,000$  counts per min) were removed. No less than 8 hours of valid wearing time with no more than 20 minutes consecutive zeroes were recognized as a valid day. After one-week of wearing, children who could provide a minimum of 4 valid days (3 weekdays and 1 weekend day) were included in the final analyses<sup>[23]</sup>.

**Body Mass Index (BMI)** BMI was calculated as weight in kilograms divided by height in meters squared. Weight and height were taken from the

latest records which were measured by PE teachers in the middle of each semester. Height was measured to the nearest 0.1 cm and weight to the nearest 0.1 kg.

#### **Procedures**

The PAQ-C was delivered to students during school time in their classroom. Children completed the questionnaires under the supervision of the teachers and researchers. At the beginning of testing, a research assistant gave a brief explanation about the requirements for completing the PAQ-C. At least one research assistant was available to clarify any aspect of the questionnaires that were required at the time of questionnaire completion. Of all the participants, a subsample of 94 children (51 males and 43 females) was randomly selected to be assessed twice to explore the test-retest reliability of the PAQ-C score. The questionnaire completion was repeated as described above with 7-10 day interval, which was considered most feasible for all schools' schedules, and also considered a reasonable period to ensure that children could not remember the questionnaire in great detail<sup>[24]</sup>.

On the day of testing, children were gathered in the school hall where the PAQ-C was administered following the same procedures as described above. During the completion of the PAQ-C, a research assistant distributed the AGs to students who were asked to wear the device positioned on the right hip for 7 consecutive days during waking hours. The accelerometer could only be removed during water-related activities (swimming, showering, and bathing) and while sleeping, and any removal was to be recorded in the PA diary given to the students. The diary was used to improve compliance to wearing the accelerometers. Additionally, investigators created a WhatsApp group with the students' parents and asked for their assistance via the WhatsApp group, to remind their children to wear the device each day.

#### **Statistical Analyses**

The Kolmogorov-Smirnov test was performed to test the normality and outlier. The values of skewness and kurtosis were applied to determine whether the data transformation should be performed<sup>[25]</sup>. Means and standard deviations (SD) were calculated for the boys, girls, and combined samples on individual items and total PAQ-C scores. Cronbach's alpha coefficient (Cronbach's  $\alpha$ ) was computed for the reliability analysis, with values

greater than 0.70 deemed acceptable for general research purposes<sup>[26]</sup>. Along with the Cronbach's  $\alpha$ , the Composite Reliability ( $\rho$ ) value and Average Variance Extracted (AVE) value were also calculated to test the construct reliability of the scale. The  $\rho$  was used to measure the overall reliability of a collection of heterogeneous but similar items and was calculated as: (sum of the standardized loadings) / [(sum of the standardized loadings) + (sum of error variances)]. The AVE described the variance captured by measurement error as opposed to the variance attributed to the latent factors was calculated as: (the sum of squared standardized factor loadings) / [(the sum of squared standardized factor loadings) + (the sum of error variances)]. A composite reliability of 0.70 or above<sup>[27]</sup> and AVE of more than 0.50<sup>[28]</sup> are deemed acceptable. The item/scale relationships were examined by corrected item total correlations (CITCs), which calculated the correlation coefficients between the scores on the items and the sum of scores on all the other items. The CITCs should be over 0.20 to indicate a homogeneous scale<sup>[29]</sup>. The intraclass correlation coefficient<sup>[30]</sup> (two-way random model) was computed to determine test-retest reliability. Multivariate analysis of variance (MANOVA), adjusted for age, was used to examine any gender differences among items 1 to 9. Gender and age differences in the overall PAQ-C score were tested by an independent *t* test and analysis of variance (ANOVA), respectively. The Spearman correlation coefficient *r* was examined to evaluate the convergent validity of the PAQ-C score with BMI and the objective PA measures. All statistical analyses were performed using SPSS version 22.0 (Statistical Product and Service Solutions, developed by IBM corporation) and a two-tailed *P* value <0.05 was considered statistically significant.

CFA with maximum likelihood estimation was performed using Mplus (Version 7.2)<sup>[31]</sup> to confirm the single factor structure of the PAQ-C. Additionally, multiple-group CFAs was performed to examine the measurement invariance (e.g., factor-loadings and factor variances and covariances) between males and females. The model performance was evaluated by four widely used indicators: the Chi-square statistic ( $\chi^2$ ), the comparative-fit index (CFI), Tucker-Lewis index (TLI), and the root-mean-square effort of approximation (RMSEA). A small  $\chi^2$  relative to the degrees of freedom, resulting in a significant statistic, was considered as goodness of fit (even though it is sensitive to sample size). Criteria of model fit indices developed by Hooper and

colleagues<sup>[32]</sup> were applied in this study: CFI/TLI > 0.95 (great), >0.90 (good); RMSEA <0.05 (good), < 0.08 (acceptable).

## RESULTS

### *Descriptive Statistics*

Students with incomplete data, or who reported sickness or other events preventing them from participating in their usual activities, during the previous 7 days, were excluded. Twenty-one students (2.6%) did not provide complete data and 35 children (4.4%) reported sickness or other events which prevented them from participating in their usual activities during the previous week. No suspicious outliers were detected and no outliers were removed. This resulted in a final sample size of 742 children (412 boys and 330 girls) aged 8-13 years (8 yrs, *n*=12; 9 yrs, *n*=141; 10 yrs, *n*=166, 11 yrs, *n*=300; 12 yrs, *n*=112; 13 yrs, *n*=11; mean age 10.5±1.1 yrs). No gender ( $\chi^2=4.41$ , *P*=0.425) or age differences ( $\chi^2=6.87$ , *P*=0.842) were found between the excluded and retained participants. The Kolmogorov-Smirnov test revealed that the PAQ-C scores were not normally distributed (*P*=0.005). Considering the skewness (0.42) and Kurtosis (0.08) were much lower than the absolute value of 1.0, data transformation was not conducted in this relatively large sample and the original data was used for further analyses. Table 1 presents the descriptive statistics for the PAQ-C individual items, summary scores for males, females and the overall sample. Most items had adequate variance and their means were close to the center of range of values. Two items (checklist and lunchtime) had relatively low means with the values of 1.91 (SD: 0.78) and 1.69 (SD: 1.06). The means of the PAQ-C summary score for the whole sample was 2.62 (SD: 0.68). No age differences were detected in the PAQ-C score (mean (SD) at age ≤9 yrs: 2.73 (0.69); age at 10 yrs: 2.58 (0.70); age at 11 yrs: 2.60 (0.67); age ≥12 yrs: 2.59 (0.68); *F*<sub>(3)</sub>=1.74, *P*=0.158).

### *Scale Reliability*

The internal consistency coefficient (Cronbach's  $\alpha$ ) was 0.79 for the overall sample. The AVE value was 0.34, suggesting relatively poor reliability for the PAQ-C score. In contrast, a  $\rho$  value of 0.81 indicated that the reliability of the PAQ-C score was satisfactory. All CITCs were above the lower limit of 0.20, ranged between 0.29 and 0.72 (see Table 2).

### Factorial Validity in All Participants

The CFA model to test the fit in the sample was a single common-factor, simple-structure model. In this model a single latent PAQ-C construct was posited to fully account for the covariation among the 9 PAQ-C items. This model provided an acceptable fit to the data,  $\chi^2_{(27)}=140.51$ , TLI=0.904, CFI=0.928, RMSEA=0.068 (90% CI: 0.059-0.079) (Table 3). The standardized factor loadings ranged from 0.31 to 0.84 for nine items.

### Invariance and Difference Tests across Gender Groups

Gender differences, examined by the MANOVA, adjusting for age, were significant on PAQ-C item 1, 2, 3, 5 and 7 (Wilks Lambda =0.045,  $P<0.001$ ). Males were more active [mean (SD): 2.67 (0.70)] than females [mean (SD): 2.56 (0.66)] ( $t_{(740)}=2.10$ ,  $p=0.035$ )

(Table 1).

Results of the CFA among the gender subsamples are presented in Table 3, indicating that one dimensional PAQ-C structure was an acceptable fit for males ( $\chi^2_{(27)}=106.63$ , TLI=0.901, CFI=0.906, RMSEA=0.071 (90% CI: 0.063-0.082) and for females ( $\chi^2_{(27)}=65.44$ , TLI=0.921, CFI=0.948, RMSEA=0.058 (90% CI: 0.037-0.079). Then a sequential model testing approach was conducted by multiple-group CFAs to test whether the measurement was invariant across males and females. A baseline model (M1: no equality constraints) was established first, and the two increasingly constrained models (M2 with equality constraints on item loadings, and M3 imposing equality constraints on factor loadings, factor variances and covariances) were tested across gender groups<sup>[33]</sup>. The difference between two multi-sample models (i.e., M2 and M1, M3 and M2) was judged based on the difference of the CFI value

**Table 1.** Description of Item Scores for Male, Female, and the Combined Sample

Item	Boys (n=412)		Girls (n=330)		Overall (n=742)	
	Mean	SD	Mean	SD	Mean	SD
Checklist Q1*	1.98	0.85	1.83	0.69	1.91	0.78
PE class Q2*	4.12	0.96	3.93	0.99	4.04	0.98
Recess Q3*	2.21	1.27	2.03	1.14	2.13	1.21
Lunchtime Q4	1.74	1.13	1.63	0.98	1.69	1.06
Describes best Q5*	2.83	1.26	2.72	1.16	2.78	1.21
After schools Q6	2.96	1.14	2.90	1.03	2.93	1.09
Evenings Q7*	2.11	1.23	2.24	1.24	2.17	1.24
Weekends Q8	2.97	1.18	2.87	1.12	2.92	1.15
Weeks summary Q9	2.76	0.97	2.71	0.85	2.74	0.92
PAQ-C*	2.67	0.70	2.56	0.66	2.62	0.68

**Note.** SD, Standard deviation; \*, Significant differences between genders ( $P<0.05$ ).

**Table 2.** Corrected Item Total Correlations and Factor Loadings for the PAQ-C (n=742)

Item	Corrected Item Total Correlations	Standardized Factor Loading
Checklist Q1	0.52	0.59
PE class Q2	0.43	0.48
Recess Q3	0.29	0.32
Lunchtime Q4	0.29	0.31
Describes best Q5	0.60	0.67
After schools Q6	0.56	0.65
Evenings Q7	0.42	0.50
Weekends Q8	0.56	0.67
Weeks summary Q9	0.72	0.84

( $\Delta$ CFI). A value equal to or less than 0.010 indicates no difference between models and equality constraints<sup>[34]</sup>. Table 3 shows the fit statistics for the multistep CFA approach across gender groups. The unconstrained model (M1) displayed an acceptable fit to the data. The second model (M2) with the factor loadings constrained to be equal across the male and female subsamples indicated the satisfactory fit to the data. CFI did not show the substantial change between M2 and M1 (0.920 vs. 0.919), suggesting evidence of an invariance of factor loadings between males and females. The final model (M3, imposing equality constraints on factor loadings, factor variances and covariances) also indicated an acceptable fit to the data. In comparison of M3 against M2,  $\Delta$ CFI (0.919 vs. 0.911) was less than 0.010, which provided support for the invariance of the factor variances and covariances across gender groups. In sum, the findings of these models in Table 3 suggested the factor loadings and factor variances and covariances of 9-item one-structure PAQ-C were invariant in boys and girls.

#### Test-retest Reliability

Of 94 students who participated in the retest, one girl reported sickness and another girl did not provide complete data. The remaining 92 children (51 male, 41 females) were included in the final analysis. The test-retest reliability was evaluated by examining the intraclass correlation coefficient, which was 0.82 for the whole sample, 0.80 for males and 0.84 for females.

#### Convergent Validity

Table 4 shows the convergent validity by accessing the correlations between the PAQ-C score with BMI and objective PA measures. Of 463 children participated in the 7-day accelerometer protocol, 358 students (77.3%) provided valid data with average daily MVPA at 43.02 (SD: 13.72) minutes. The correlation between the summarized PAQ-C score and BMI was significant in males ( $r=-0.18$ ,  $P<0.01$ ), females ( $r=-0.19$ ,  $P<0.01$ ) and overall sample ( $r=-0.18$ ,  $P<0.01$ ). PAQ-C score were significantly correlated with MVPA measured by accelerometer in males ( $r=0.38$ ,  $P<0.01$ ), females ( $r=0.26$ ,  $P<0.05$ ) and all children ( $r=0.33$ ,  $P<0.01$ ).

#### DISCUSSION

To the best of our knowledge, this is the first PAQ-C validation study among Chinese children aged 8-13 years. Satisfactory internal consistency, good test-retest reliability, one-dimensional construct, and moderate convergent validity with accelerometer-based measures have provided evidence that the PAQ-C score is a valuable measurement tool for large PA assessment studies with Chinese children.

CITCs were performed to check if the individual item was measuring the same construct by evaluating the correlation between the corresponding item score with the others in the set of scale. All CITCs varied between 0.29 and 0.72, thus well above the lower limit of 0.20<sup>[29]</sup>. In the present

**Table 3.** Fit Indices for Cross-gender Samples of the PAQ-C

Variables	$\chi^2$	DF	P	CFI	TLI	RMSEA	90% CI	Model Comparison	$\Delta$ CFI
All participants	140.51	27	<0.001	0.928	0.904	0.068	0.059-0.079	-	-
Males (n=412)	106.63	27	<0.001	0.906	0.901	0.071	0.063-0.082	-	-
Females (n=330)	65.44	27	<0.001	0.948	0.921	0.058	0.037-0.079	-	-
Multi-group process (cross-gender)									
M1	172.06	54	<0.001	0.920	0.901	0.057	0.047-0.066	-	-
M2	190.92	62	<0.001	0.919	0.906	0.055	0.047-0.064	M2 vs. M1	0.001
M3	211.31	70	<0.001	0.911	0.909	0.055	0.046-0.063	M3 vs. M2	0.009

**Note.** M1, no parameters were constrained to be equal across gender groups; M2, factor loadings were constrained to be equal across gender groups; M3, factor loadings, factor variances and covariances were constrained to be equal across gender groups. DF, Degree of freedom; CFI, The comparative-fit index; TLI, Tucker-Lewis index; RMSEA, The root-mean-square effort of approximation; CI, Confidence interval;  $\Delta$ CFI, Change in the comparative-fit index.

study, the AVE value indicated that the PAQ-C scores explained only 34% of the variance in their items, lower than the recommended level (0.50). The result was similar with the construct reliability of other questionnaires (0.30 in the a modified Positive and Negative Affect Schedule including a direction scale)<sup>[35]</sup>, however, there is no comparable AVE data available for the PAQ-C. Cronbach's  $\alpha$  and  $\rho$  over 0.70 is usually considered an indication of a reliable questionnaire. The fact that Cronbach's  $\alpha$  was 0.79 and  $\rho$  was 0.81 for the overall sample suggested satisfactory scale consistency in Hong Kong Chinese children. Additionally, the current findings revealed ICCs higher than 0.80 for both genders, which is strong evidence to support the test-retest reliability of the PAQ-C score in this target population.

All participants presented a summarized PAQ-C score of 2.62 (SD: 0.68), which is lower than that in the studies among different racial groups of children, i.e., 3.49 (SD: 0.68) in British samples<sup>[14]</sup>, 3.36 (SD: 0.80) for European American children, and 3.37 (SD: 0.69) in African American children<sup>[15]</sup>. In two additional validation studies, summary scores over 3.20 were reported<sup>[12,16]</sup>. Consistent with previous PA level assessment studies, the finding from the current study further reveals that Hong Kong children demonstrated a lower PA pattern than their counterparts from other countries<sup>[36-37]</sup>. This finding calls for the imperative action to improve PA among Hong Kong Children.

Similar to the Pearson correlation, the standardized regression coefficient of each item was regarded as the predicting indicator for the construct of the questionnaire. The significant standardized regression coefficient was 0.32 for recess and 0.31 for lunch time PA, slightly lower than the suggested

loading value 0.40<sup>[38]</sup>, which were similar to the values in the study of Janz et al.<sup>[13]</sup> (0.25 for recess and 0.17 for lunch time). The low recess [mean (SD): 2.13 (1.21)] and lunch time PA scores [mean (SD): 1.69 (1.06)] may account for these lower loadings, which indicated that these two items did not add substantially to the PAQ-C score in Hong Kong Chinese children. In line with other empirical studies, Hong Kong primary school children were found to have low PA engagement during these school segments<sup>[36,39]</sup>. One observation study has pointed out that school recess-based activity in Hong Kong is limited with sitting accounted for 23.3% of the observation time, while standing accounted for 40.5% among students<sup>[39]</sup>. In Hong Kong, school policies do not encourage students to engage in activities during recess and lunch time. Students are asked to stay in class after their lunch to watch teaching videos together or to do their homework, and are not permitted to run during these segments in order to avoid possible injuries. Low PA level in Hong Kong children can also possibly be attributed to a short lunch time (approximately 25-30 min). Furthermore, schools in Hong Kong are generally small and are situated in high-density buildings, and a lack of outdoor play space (2 m<sup>2</sup> per student) may limit children's activity during recess<sup>[40]</sup>. This result highlights the necessity for further studies to quantify the PA levels of children during recess and lunchtime and to examine their contribution towards PA guideline.

CFAs were conducted on the test sample and cross-validation gender groups. The results of CFA in the overall sample suggest one-factor model was appropriate, which indicates that the questionnaire measures only one construct, on MVPA over the previous

**Table 4.** Correlations between the PAQ-C Score and PA Measures (n=358)

Variables	Description Mean (SD)			Correlation		
	Overall n=358	Males n=194	Female n=164	Overall n=358	Males n=194	Female n=164
BMI (kg/m <sup>2</sup> )	18.69 (3.69)	19.28 (3.96)	17.99 (3.21)	-0.18**	-0.18**	-0.19**
Objective PA (min/day)						
Total MPA	29.01 (9.49)	30.79 (9.37)	26.91 (9.22)	0.24**	0.34**	0.14
Total VPA	13.59 (6.49)	14.97 (6.84)	11.97 (5.65)	0.36**	0.36**	0.35*
Total MVPA	43.02 (13.72)	45.38 (14.20)	40.22 (12.62)	0.33**	0.38**	0.26*

**Note.** \*  $P < 0.05$ , \*\*  $P < 0.01$ . BMI, Body mass index; MPA, Moderate physical activity; VPA, Vigorous physical activity; MVPA, Moderate-to-vigorous physical activity.

7 days. The finding was consistent with Janz and colleagues<sup>[13]</sup>. Boys as a group reported the PAQ-C score at 2.67 (SD: 0.70) as well as girls at 2.56 (SD: 0.66). A series of multiple-sample CFAs on the measurement invariance analysis revealed that the factor loadings and factor variances and covariances of the PAQ-C measurement model were invariant across gender groups. This finding suggests that the PAQ-C measures the same structure, and boys and girls responded to the items in the same way.

It is known that PA is a key component in weight control. Substantial evidences have shown that overweight and obese children are less physically active than their peers with normal body weight<sup>[41]</sup>. In the current study, convergent validity was assessed by testing the extent to which the PAQ-C score related to BMI. This study found an inverse relationship between the PAQ-C score and BMI ( $r$  was -0.18 for males, -0.19 for females, and -0.18 for overall samples), which supports the fact that children with higher BMIs are likely to participate in less PA<sup>[42]</sup>. This finding had similar magnitude to the correlation results ( $r=-0.16$ ) in the PAQ-C validation study of Moore and colleagues<sup>[15]</sup>.

Accelerometers monitor PA by recording the acceleration of human movement and this measure has been used as the criterion reference to detect intensity and quantity of movement<sup>[43]</sup>. The convergent validity of the PAQ-C score was further evaluated by calculating the correlation between the PAQ-C score and accelerometer determined PA measures. Moderate correlations were observed between the PAQ-C summary score and MVPA for boys ( $r=0.38$ ) and for the overall sample ( $r=0.33$ ). A slightly lower but significant correlation ( $r=0.26$ ) was found for girls. The correlation coefficients were similar to previously reported associations between the PAQ-C score and PA measured by the Caltrac activity monitor ( $r=0.39$ )<sup>[12]</sup>. The magnitude of correlation that we report are also similar to other PA recall questionnaires for children presented when compared with objective PA measures. For example, in the study of Welk et al.<sup>[44]</sup>, the correlation coefficient was 0.24 for the Youth Media Campaign Longitudinal Survey (YMCLS) in the estimation of weekly PA. Sallis et al.<sup>[45]</sup> reported the validity coefficient of 0.33 and 0.29 with heart rate  $\geq 140$  bpm and  $\geq 160$  bpm for the Seven-Day Physical Activity Recall (PAR) among fifth grade students. For 1-day to 3-day recalls in children, the correlation was found with 0.32 for the previous day Self-Admin PA Checklist (SAPAC)<sup>[46]</sup>, 0.35-0.43 for the Previous Day

Physical Activity Recall (PDPAR)<sup>[47]</sup>, 0.27-0.46 for the 3-Day Physical Activity Recall (3DPAR)<sup>[48]</sup>, and 0.22 for the 3-Day Aerobic Recall<sup>[49]</sup>. The validity results in the current study are also comparable with other PAQ-C convergent studies evaluated using step tests in European American children ( $r=0.30$ )<sup>[15]</sup>, and the ½ mile walk-run test ( $r=-0.37$ ) in British children<sup>[14]</sup>.

Several limitations of this study warrant consideration. The PAQ-C was designed for children aged 8-14 years. However, despite the fact that the age range of children recruited in the current study was 8-13 years, the proportion of children at the 8 and 13 years old was small, which may limit the generalizability of the results for children. Additional studies are needed to examine the suitability of the PAQ-C score among the pediatric population in these age groups. Secondly, regarding the convergent validity of the PAQ-C score with BMI, the latest records of participants' height and weight from PE teachers were used rather than those measured by researchers during testing, this may slightly affect the findings. Additionally, even the participants in the study were recruited from two districts with different SES in Hong Kong, given the influence of SES in children's PA engagement and the diverse SES presenting in China, more studies are recommended to further confirm the wide application of the PAQ-C.

## CONCLUSION

This is the first reported validation study of the Chinese version of the PAQ-C. The findings provided support for the reliability and validity of the PAQ-C score in Hong Kong Chinese children aged 8-13 years. Although the PAQ-C is limited in its ability to provide information on PA frequency, intensity and duration, its ease of use and administration, low cost for investigators and low burden to participants make the PAQ-C applicable for use in large-scale PA studies with Chinese children.

## ACKNOWLEDGMENTS

We would like to thank all the participating children.

Received: December 7, 2015;

Accepted: March 10, 2016

## REFERENCES

1. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. *Can Med Assoc J*, 2006; 174, 801-9.



2. World Health Organization (2010). Global recommendations on physical activity for health. Retrieved from: <http://www.who.int/dietphysicalactivity/publications/9789241599979/en/>. [2015-01-05]
3. Malina RM. Physical activity and fitness: pathways from childhood to adulthood. *Am J Hum Biol*, 2001; 13, 162-72.
4. Telama R, Yang X, Laakso L, et al. Physical activity in childhood and adolescence as predictor of physical activity in young adulthood. *Am J Prev Med*, 1997; 13, 317-23.
5. Corder K, Brage S, Ekelund U. Accelerometers and pedometers: methodology and clinical application. *Curr Opin Clin Nutr Metab Care*, 2007; 10, 597-603.
6. Matthews CE, Welk G. Use of self-report instruments to assess physical activity. In Welk GJ, (Ed.) *Physical Activity Assessment for Health-related Research*, Champaign: Human Kinetics, 2002; 107-23.
7. Liu A, Ma G, Zhang Q, et al. Reliability and validity of a 7-day physical activity questionnaire for elementary students. *Chin J Epidemiol*, 2003; 24, 901-4. (In Chinese)
8. Huang Y, Wong SH, Salmon J. Reliability and validity of the modified Chinese version of the Children's Leisure Activities Study Survey (CLASS) questionnaire in assessing physical activity among Hong Kong children. *Pediatr Exerc Sci*, 2009; 21, 339-53.
9. Hussey J, Bell C, Gormley J. The measurement of physical activity in children. *Phys Ther Rev*, 2007; 12, 52-8.
10. Kremers SP, Visscher TL, Seidell JC, et al. Cognitive determinants of energy balance-related behaviours. *Sports Med*, 2005; 35, 923-33.
11. Chinapaw MJ, Mokkink LB, van Poppel MN, et al. Physical activity questionnaires for youth. *Sports Med*, 2010; 40, 539-63.
12. Kowalski KC, Crocker PR, Faulkner RA. Validation of the Physical Activity Questionnaire for Older Children. *Pediatr Exerc Sci*, 1997; 9, 174-86.
13. Janz KF, Lutuchy EM, Wenthe P, et al. Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. *Med Sci Sports Exerc*, 2008; 40, 767-72.
14. Thomas EL, Upton D. Psychometric properties of the physical activity questionnaire for older children (PAQ-C) in the UK. *Psychol Sport Exerc*, 2014; 15, 280-7.
15. Moore JB, Hanes Jr JC, Barbeau P, et al. Validation of the Physical Activity Questionnaire for Older Children in children of different races. *Pediatr Exerc Sci*, 2007; 19, 6-19.
16. Crocker P, Bailey DA, Faulkner RA, et al. Measuring general levels of physical activity: preliminary evidence for the Physical Activity Questionnaire for Older Children. *Med Sci Sports Exerc*, 1997; 29, 1344-9.
17. Deng HB, Macfarlane DJ, Thomas GN, et al. Reliability and validity of the IPAQ-Chinese: the Guangzhou Biobank Cohort study. *Med Sci Sports Exerc*, 2008; 40, 303-7.
18. Lau PW, Lam MH, Leung BW, et al. The longitudinal changes of national identity in Mainland China, Hong Kong and Taiwan before, during and after the 2008 Beijing Olympics Games. *Int J Hist Sport*, 2012; 29, 1281-94.
19. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *J Adv Nurs*, 2003; 42, 57-63.
20. Puyau MR, Adolph AL, Vohra FA, et al. Validation and calibration of physical activity monitors in children. *Obes Res*, 2002; 10, 150-7.
21. Trost SG, Loprinzi PD, Moore R, et al. Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc*, 2011; 43, 1360-8.
22. Evenson KR, Catellier DJ, Gill K, et al. Calibration of two objective measures of physical activity for children. *J Sports Sci*, 2008; 26, 1557-65.
23. Esliger DW, Copeland JL, Barnes JD, et al. Standardizing and optimizing the use of accelerometer data for free-living physical activity monitoring. *J Phys Act Health*, 2005; 3, 366-83.
24. Paiva CE, Barroso EM, Carnesecca EC, et al. A critical analysis of test-retest reliability in instrument validation studies of cancer patients under palliative care: a systematic review. *BMC Med Res Methodol*, 2014; 14, 8.
25. Blanca MJ, Arnau J, Lopez-Montiel D, et al. Skewness and Kurtosis in Real Data Samples. *Methodology*, 2013; 9, 78-84.
26. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951; 16, 297-334.
27. Raykov T. Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *Brit J of Math Stat Psy*, 2001; 54, 315-24.
28. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res*, 1981; 18, 39-50.
29. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and use*. 4th edition. Oxford: Oxford University Press; 2008.
30. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*, 1996; 50, 30-46.
31. Muthén LK, Muthén BO. (1998-2014). *Mplus (version 7.2) [computer software]*. Los Angeles, CA: Muthén & Muthén.
32. Hooper D, Coughlan J, Mullen MR. *Structural Equation Modelling: Guidelines for Determining Model Fit*. *Electronic Journal of Business Research Methods*, 2008; 6, 53-60.
33. Byrne BM. *Structural equation modeling with AMOS: basic concepts, applications, and programming*. 2nd eds. New York, NY: Routledge; 2010.
34. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling*, 2002; 9, 233-55.
35. Nicolas M, Martinent G, Campo M. Evaluation of the psychometric properties of a modified Positive and Negative Affect Schedule including a direction scale (PANAS-D) among French athletes. *Psychol Sport and Exerc*, 2014; 15, 227-37.
36. Gao Y, Wang JJ, Lau PW, et al. Pedometer-determined physical activity patterns in a segmented school day among Hong Kong primary school children. *J Exerc Sci Fit*, 2015; 13, 42-8.
37. Mak KK, Day JR. Secular trends of sports participation, sedentary activity and physical self-perceptions in Hong Kong adolescents, 1995-2000. *Acta Paediatr*, 2010; 99, 1731-4.
38. Koukkunen H, Penttilä K, Kempainen A, et al. C-reactive protein, fibrinogen, interleukin-6 and tumour necrosis factor- $\alpha$  in the prognostic classification of unstable angina pectoris. *Ann Med*, 2001; 33, 37-47.
39. Johns DP, Ha AS. Home and recess physical activity of Hong Kong children. *Res Q Exerc Sport*, 1999; 70, 319-23.

40. Education Department of Hong Kong Special Administrative Region Government (2000). Physical Space and Learning Environment of Primary and Secondary Schools. <http://www.legco.gov.hk/yr99-00/english/panels/ed/papers/e1693-03.pdf>. [2015-01-15]
41. Molnar D, Livingstone B. Physical activity in relation to overweight and obesity in children and adolescents. *Eur J Pediatr*, 2000; 159, S45-55.
42. Janssen I, Katzmarzyk PT, Boyce WF, et al. Comparison of overweight and obesity prevalence in school-aged youth from 34 countries and their relationships with physical activity and dietary patterns. *Obes Rev*, 2005; 6, 123-32.
43. Janz KF. Validation of the CSA accelerometer for assessing children's physical activity. *Med Sci Sports Exerc*, 1994; 26, 369-75.
44. Welk GJ, Wickel E, Peterson M, et al. Reliability and validity of questions on the youth media campaign longitudinal survey. *Med Sci Sports Exerc*, 2007; 39, 612-21.
45. Sallis JF, Buono MJ, Roby JJ, et al. Seven-day recall and other physical activity self-reports in children and adolescents. *Med Sci Sports Exerc*, 1993; 25, 99-108.
46. Sallis JF, Strikmiller PK, Harsha DW, et al. Validation of interviewer-and self-administered physical activity checklists for fifth grade students. *Med Sci Sports Exerc*, 1996; 28, 840-51.
47. Trost SG, Ward DS, McGraw B, et al. Validity of the Previous Day Physical Activity Recall (PDPAR) in fifth-grade children. *Pediatr Exerc Sci*, 1999; 11, 341-8.
48. Pate RR, Ross R, Dowda M, et al. Validation of a 3-day physical activity recall instrument in female youth. *Pediatr Exerc Sci*, 2003; 15, 257-65.
49. Janz KF, Witt J, Mahoney LT. The stability of children's physical activity as measured by accelerometry and self-report. *Med Sci Sports Exerc*, 1995; 27, 1326-32.