Obesity

Reliability and Validity of a Questionnaire for Physical Activity Assessment in South American Children and Adolescents: The SAYCARE Study

Marcus Vinícius Nascimento-Ferreira [1,2], Augusto César Ferreira De Moraes [1,3*], Paulo Vinícius Toazza-Oliveira[1], Claudia L.M. Forjaz[4], Juan Carlos Aristizabal[5], Alba M. Santaliesra-Pasías[3,6,7], Candela Lepera[8], Walter Viana Nascimento-Junior[9], Estela Skapino[10], Carlos Alberto Delgado[11,12], Luis Alberto Moreno[3,6,7], and Heráclito Barbosa Carvalho [1]*

Objective: The objective of this article is to test the reliability and validity of the new and innovative physical activity (PA) questionnaire.

Methods: Subsamples from the South American Youth/Child Cardiovascular and Environment Study (SAYCARE) study were included to examine its reliability (children: $n = 161$; adolescents: $n = 177$) and validity (children: $n = 82$; adolescents: $n = 60$). The questionnaire consists of three dimensions of PA (leisure, active commuting, and school) performed during the last week. To assess its validity, the subjects wore accelerometers for at least 3 days and 8 h/d (at least one weekend day). The reliability was analyzed by correlation coefficients. In addition, Bland-Altman analysis and a multilevel regression were applied to estimate the measurement bias, limits of agreement, and influence of contextual variables.

Results: In children, the questionnaire showed consistent reliability ($\rho = 0.56$) and moderate validity ($\rho = 0.46$), and the contextual variable variance explained 43.0% with −22.9 min/d bias. In adolescents, the reliability was higher ($\rho = 0.76$) and the validity was almost excellent ($\rho = 0.88$), with 66.7% of the variance explained by city level with 16.0 min/d PA bias.

[1] Youth/Child Cardiovascular Risk and Environmental Research Group, Department of Preventive Medicine, School of Medicine, University of São Paulo, São Paulo, Brazil. Correspondence: Augusto César Ferreira De Moraes (augustocesar.demoraes@usp.br) [2] Growth, Exercise, Nutrition and Development Research Group, Faculty of Health Sciences, University of Zaragoza, Zaragoza, Spain [3] Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA [4] Exercise Hemodynamic Laboratory, School of Physical Education and Sport, University of São Paulo, São Paulo, Brazil [5] Research Group in Physiology and Biochemistry, University of Antioquia, Medellin, Colombia [6] Agroalimentary Institute of Aragon, University of Zaragoza, Zaragoza, Spain [7] Biomedical Research Networking Center-Physiopathology of Obesity and Nutrition (CIBERObn), University of Zaragoza, Zaragoza, Spain [8] School of Nutrition, University of Buenos Aires, Buenos Aires, Argentina [9] Metabolic Diseases, Exercise, and Nutrition Research Group, Center for Health Sciences, Federal University of Piaui, Teresina, Piaui, Brazil [10] School of Nutrition, University of the Republic, Montevideo, Uruguay [11] Department of Pediatrics, School of Medicine, National University of San Marcos, Lima, Peru [12] Neonatal Unit, National Institute of Child Health, Lima, Peru.

Funding agencies: ACFDeM received a postdoctoral scholarship from the National Counsel of Technological and Scientific Development (CNPq; proc. 313772/2014-2) and the São Paulo Research Foundation (proc. 2014/13367-2 and 2015/14319-4). MVN-F received a PhD Student Internships Abroad scholarship from the National Counsel of Technological and Scientific Development (CNPq; proc. 200340/2015-8) and a PhD Student Brazilian scholarship from the São Paulo Research Foundation (FAPESP; proc. 2016/18436-8 and 2017/11732-3). LAM was given a visiting professor scholarship from the São Paulo Research Foundation (FAPESP; proc. 2015/11406-3). CLMF received an advanced scientist scholarship from the National Counsel of Technological and Scientific Development (CNPq; proc. 304003/2014-0). HBC received a research grant from the São Paulo Research Foundation (2014/11468-6, Research Grant for Scientific Publication Paper (2017/15560-2), and an advanced scientist scholarship from the National Counsel of Technological and Scientific Development (CNPq; proc. 300951/2015-9).

The South American Youth/Child Cardiovascular and Environment (SAYCARE) study was supported mainly by the Brazilian Government from the National Counsel of Technological and Scientific Development (CNPq; proc. 471266/2013-2) and the São Paulo State Government from the São Paulo Research Foundation (FAPESP; proc. 2014/11468-6). The SAYCARE study was also cofunded by other agencies in other countries as follows: (1) Collaborative Projects Fund (R.D. N 501-2015-INSN-DG-OEA) granted by the National Institute of Child Health, Lima, Peru; (2) Sustainability Strategy at the University of Antioquia 2014-2015, Research Group of Social and Economic Determinants of Health and Nutrition, and Demography and Health Research Group at the University of Antioquia, Medellin, Colombia, and Interuniversity Services Corporation from University of Antioquia; (3) Secretary of University Extension and Student Welfare, University of Buenos Aires, Buenos Aires, Argentina; and (4) European Regional Development Fund (MICINN-FEDER) to the Growth, Exercise, Nutrition and Development Research Group.

Disclosure: The authors declared no conflict of interest.

Author contributions: ACFDeM, MVNF, and CLMF designed this reliability and validity study. ACFDeM, MVN-F, LAM, and HBC supervised the data analysis and interpretation. ACFDeM, MVNF, PVT-O, WVN-J, JCA, CL, CAD, and ES contributed to the data collection. AMS-P contributed to the data interpretation. All authors were involved in writing the paper and had final approval of the submitted and published versions.

*Marcus Vinicius Nascimento-Ferreira and Augusto Cesar Ferreira De Moraes contributed equally to this work.

Additional Supporting Information may be found in the online version of this article.

Received: 16 June 2017; **Accepted:** 19 December 2017; **Published online 21 February 2018. doi:10.1002/oby.22116**

Conclusions: The SAYCARE PA questionnaire shows acceptable (in children) to strong (in adolescents) reliability and strong validity in the measurement of PA in the pediatric population from low- to middle-income countries.

Obesity (2018) **26**, S23–S30. doi:10.1002/oby.22116

Introduction

Recently, Lee et al. (1) reported that physical inactivity is responsible for 5.3 million deaths annually worldwide. However, physical activity (PA) is complex, comprising multiple domains, and it varies with age, sex, seasonality, day of the week, and time of day. Moreover, PA behaviors are influenced by biological, sociological, psychological, and environmental factors (2), which implies that they vary from one culture and/or country to another, making it difficult to assess self-reported PA by using a standardized questionnaire in different countries, although using such a questionnaire could improve the comparisons among countries and estimate the global PA prevalence. Thus, it is important to understand the advantages and disadvantages of the subjective instruments used to assess self-reported PA by analyzing their sensitivity to intra- and intercenter variability as well as their validity (3).

Studies have been developed to measure and compare the PA levels performed in different PA domains among some European countries (4). These studies have provided interesting insights into the actual low PA level in these populations as well as into the interaction of sociodemographic and environmental factors on this behavior in these countries. However, these results cannot be extrapolated to other countries and regions, such as South American countries, since socioeconomic status has effects on PA level trends (5), such as the variation in school physical education lessons, active commuting, and leisure-time activities among different countries (6).

Therefore, future studies should compare PA levels and their determinants among South American countries. However, because of the sociocultural characteristics of these countries, a valid and transcultural method for assessing PA in different domains should be developed. Three subjective methods have been developed in South American countries, and their validity has been tested (7-9); however, none of these questionnaires measure the total time of PA in the different domains and at different intensities (moderate and/or vigorous). Moreover, none of these questionnaires have verified the concordance of the estimate of the World Health Organization–recommended classification with objective methods.

Multicenter studies from Europe (10,11) and the United States (12) have developed specifically standardized questionnaires for assessing and comparing PA, but they cannot be directly applied for South American countries. Therefore, based on these previous questionnaires, the South American Youth/Child Cardiovascular and Environment (SAYCARE) study, which was described in the first paper of this supplement, developed a specific PA questionnaire for South American countries. The objectives of the present study are (1) to examine the reliability of this questionnaire in South American pediatric populations (aged 3-18 years old) and (2) to show its validity compared with accelerometers.

Methods

Study design

The participants of the present study were children and adolescents (3-18 years old) who were enrolled in preschool, primary school, and up to the third year of high school at both public and private schools. The exclusion criteria were pregnancy, inability to complete the questionnaires, and refusing to sign the informed consent. This consent was necessary for the parents or guardians and for the adolescents or children. In addition, the headmasters of the selected schools also gave their consent to collaborate with the study. The institutional research ethics committees of the involved countries approved the study protocol.

Data collection was performed in Buenos Aires (Argentina), Lima (Peru), Medellin (Colombia), Montevideo (Uruguay), Santiago (Chile), and São Paulo and Teresina (Brazil). In each city, schools were chosen by convenience for data collection. Their headmasters were contacted and received a formal invitation with detailed information about the study. For the schools that agreed to participate, an information letter and verbal explanation were provided for the potential participants and their parents or guardians. For those who agreed to participate, written informed consent had to be signed by a parent or guardian and by the participant if he or she was an adolescent.

For the purposes of the present study, the SAYCARE PA questionnaire was administered twice (first application of the questionnaire = Q1; second application of the questionnaire = Q2) with a 15-day interval, following the international protocols (13). For children from 3 to 10 years of age (preschool and primary school), the parents or guardians filled out the questionnaires (14) as a parental-proxy report of children's PA, whereas the adolescents (11-17 years) answered the questions by themselves. To complete the questionnaire, the responders received information about when and how Q1 and Q2 should be answered, and the return date was indicated on the questionnaire. In addition, to objectively measure the PA levels, some volunteers used an accelerometer for 7 days before completing the questionnaires. For the reliability analysis, data collected in seven centers (Buenos Aires [Argentina], Lima [Peru], Medellin [Colombia], Montevideo [Uruguay], Santiago [Chile], and São Paulo and Teresina [Brazil]) were used; for the validity analysis, data from four centers were included (Lima, Medellin, Teresina, and São Paulo).

Measures

SAYCARE PA questionnaire. The SAYCARE PA questionnaire addresses the routine activities (frequency and intensities) of the last week (7 days) and encompasses three PA domains (PA at school, PA at leisure time, and PA while commuting).

PA at school includes the following:

- Physical education classes: participate or not. If the child or adolescent participates, the total number of classes per week, the

duration of each class (in minutes), and intensity of classes were recorded.
- PA during school breaks: presence or absence of PA during the school break. If present, the duration of the break (in minutes) and intensity of PA were recorded.
- The volume of PA at school, expressed in minutes per day, was calculated by the sum of minutes of PA in physical education classes and minutes of PA during the breaks ([mean PA at physical education classes × number of class] + [mean PA during the school break × number of breaks]) / 5.

PA at leisure time includes the following:

- Supervised leisure-time PA (weekdays and weekend): participate or not. If the child or adolescent participates, the number of supervised PAs, the type of PA performed each day, the duration of PA (in minutes), and the intensity of each PA were recorded.
- Unsupervised leisure-time PA (weekdays and weekend): participate or not. If the child or adolescent participates, the type of PA performed each day, the duration of PA (in minutes), and the intensity of each PA were recorded.
- The time spent at PA at leisure time was assessed for the week (mean of weekdays) and weekend (mean of weekend days). The total time spent on PA at leisure was calculated as ([mean PA at leisure in the weekdays × 5] + [mean PA at leisure in the weekend days × 2]) / 7.

Active commuting includes the following:

- Active commuting: participate or not. If the child or adolescent participates, the type of PA performed (biking and/or walking), the number of days per week of each PA, the duration of PA (in minutes per day), and the intensity of each PA were recorded. The volume of active commuting in minutes per day was defined by the sum of duration of each commuting activity performed.

To cover these aspects (domains of PA, total number of minutes), the questionnaire contains 47 questions. The intensity was estimated by three levels of subjective fatigue: light intensity ("I do not sweat, and my breathing is normal"), moderate intensity ("I am a bit sweaty and with fast breathing"), and vigorous intensity ("I am quite sweaty and breathless") (15). Based on this scale, we stratified the amount of PA time according to each intensity category (moderate and vigorous), and we estimated the weekly total moderate-to-vigorous PA (MVPA), summarizing the time in both intensities. Following current PA guidelines (16,17), subjects were classified as active when they accumulated at least 60 min/d of MVPA. For this classification, we applied the cited scale.

Accelerometers. Accelerometer use and analysis followed the standardized protocol for youth populations (18). An Actigraph MTI accelerometer (model GT3X; Manufacturing Technology Inc., Fort Walton Beach, Florida) was used while attached to the waist of the volunteer for seven consecutive days for at least 8 hours. Parents or guardians as well as children and adolescents were instructed to put the accelerometer on in the morning as soon as they got out of bed and to remove it before going to sleep and during any aquatic activity (19). They also received a daily report with the accelerometer's instructions and were asked to complete this report with the day, time, and duration of any removal and/or replacement of the accelerometer; the activities performed during this period; and the beginning and end of any sleeping period or event.

Accelerometers were configured with 30 Hz, an epoch of five seconds, and a shaft for measures (7). For data analysis, periods of monitoring with zero counts per minute (CPM) for more than 20 minutes (19) and with more than 20,000 CPM (19) were excluded. The total time of use of the device was calculated by the total of the day subtracted from the nonuse period of time. Accelerometer data were considered valid only when they had been recorded for at least 8 hours per day and for a minimum of 3 days (at least two weekdays and one weekend day) (20), excluding the days of the delivery and return of the accelerometers (21).

PA was expressed as the overall PA (CPM per day) and as the average daily amount of time (in minutes) spent at light, moderate, and vigorous PA. We used validated cutoff points for each age group for these intensities, which were as follows: children (22): light (26-573 CPM), moderate (574-1,002 CPM), and vigorous ($\geq$ 1,003 CPM); adolescents (19): light (101-1,999 CPM), moderate (2,000-4,999 CPM), and vigorous ($\geq$ 4,000 CPM). According to recent guidelines (17), in both age groups, the subjects were classified as active when accumulating at least an average of 60 min/d of MVPA $\geq$ 574 CPM in children (22) and $\geq$ 2,000 CPM in adolescents (19).

Statistical analysis

Sample size calculations were performed to verify the reliability and validity of the PA questionnaire in the study population. For the reliability analysis, the sample size was calculated by using Cronbach's $\alpha = 0.75$, $\alpha = 5\%$, and $\beta = 80\%$ (23). For the validity analysis, the parameters were as follows: correlation coefficient = 0.50, $\alpha = 5\%$, and $\beta = 80\%$ (23). From these parameters, the necessary sample size estimated was 120 participants for reliability analysis and 83 for validity analysis. Considering the possible loss of participants, a 25% greater sample size was recruited for these analyses ($n = 150$ for reliability and $n = 103$ for validity) (24).

Descriptive analyses included the median and interquartile interval, percentages, and confidence intervals (95%). The reliability between Q1 and Q2 measurements was calculated for categorical variables by using Cohen's kappa statistic (k, agreement, and coefficient) and for quantitative variables by using the Spearman correlation coefficient (r_s). Reliability was analyzed by domain (PA at school, active commuting, and leisure PA), intensity (moderate and vigorous), and total (summary of the three domains).

For validity, PA times calculated by Q1 and accelerometer data were analyzed in the following two steps: (1) kappa coefficients and the Spearman correlation coefficient and (2) multilevel regression model with random intercept. We used the hierarchical conceptual model. A potential confounder was retained in the multivariate model when the significance level in the unadjusted analysis was $P \leq 0.20$ or any change of 10% in the variance estimation. The exception was city; the variable was retained in all adjusted models based on the theoretical assumption of the influence in the PA levels (25,26). In the multivariate analysis, potential confounders were allocated according to the following sequence: Model 0: city; Model 1: Model 0 plus school; Model 2: Model 1 plus sex and age (Table 3). The subjective measurements were plotted against

TABLE 1 Descriptive characteristics of children and adolescents

Children	Q1 ($n = 237$), %	Q2 ($n = 161$), %	Accelerometer ($n = 82$), %	P1	P2
Sex				0.632	0.46
Female	51.5	49.8	54.4		
Male	48.5	50.2	45.6		
Age				0.002	0.664
3-5 y	34.0	57.8	36.7		
6-10 y	66.0	42.2	63.3		
Maternal education level				0.916	0.262
Incomplete high school	23.7	18.2	13.6		
High school	14.2	15.2	20.5		
Technical education	9.5	9.1	4.5		
University degree	52.7	57.6	61.4		
School type				0.009	0.764
Public	48.0	61.9	46.2		
Private	52.0	38.1	53.8		
Adolescents	**Q1 ($n = 215$), %**	**Q2 ($n = 177$), %**	**Accelerometer, ($n = 60$), %**	**P1**	**P2**
Sex				0.135	0.512
Female	51.2	58.3	55.7		
Male	48.8	41.7	44.3		
Age				0.553	0.84
11-14 y	51.1	48.3	52.5		
15-18 y	48.9	51.7	47.5		
Maternal education level				0.163	0.01
Incomplete high school	22.3	11.9	11.4		
High school	24.9	16.7	15.9		
Technical education	11.4	14.3	4.5		
University degree	40.4	57.1	68.2		
School type				0.033	0.013
Public	46.0	36.2	29.0		
Private	54.0	63.8	71.0		

P1, χ^2 goodness of fit test for comparison between the samples in the first and second applications of the questionnaire; P2, χ^2 goodness of fit test for comparison between the first application of the questionnaire and accelerometer data; Q1, first application of the questionnaire; Q2, second application of the questionnaire.

respectively objective measurements to check for linearity and to confirm that the log transformation of variables was necessary. Unadjusted linear regression analyses were used to investigate the explained variance (calculated as the unadjusted R^2) of all subjective measurements and combinations of subjective measurements on behaviors derived from the objective measurements. The residuals of the models were exploited for calculating limits of agreement (± 1.96 SD). The city, school, sex, and age (gained by objective measurement) were regressed on the residuals to explore the influence of these variables on the prediction error. A full model was calculated by including all field measurements into the regression equation. A fitted (simplified) model was built by reducing the full model until only statistically significant ($P < 0.05$) variables were left in the model. This was performed manually by using several forward and backward steps to avoid possible bias introduced by automated procedures (27). Additionally, a sparse model with $R^2 \geq 0.90$ was calculated to identify a minimal set of variables for estimating PA with a certain degree of accuracy. The model-building process was performed again by using ridge regression to ensure that the processes were not biased by the multicollinearity present in the data. To investigate the fit of the data-driven models, the residuals were again analyzed by city. Biases (mean of residuals) and limits of agreement (± 1.96 SD) were calculated. Separate models for the investigated subgroups (children and adolescents) were fitted to the data and compared to the model fitted in the full sample. We developed a conceptual framework using the multilevel regression to investigate how disagreement varied by sociodemographic factors specific for multicenter studies. We used the sociodemographic variables (city, school, sex, and age) as predictors for disagreement between the two methods and modeled against disagreement for objective and subjective methods. Generally, we used multilevel analyses to account for the cluster design of the study (students within school classes and classes within schools). In a few of the analyses, the model did not converge, and these analyses were applied without the random effects of the school class and school that were available after applying the Durbin-Wu-Hausman test. Additionally, a Bland-Altman plot was used to investigate the influence of objective measurements on the residuals in the full sample (28).

TABLE 2 Descriptive results of PA and Spearman correlation coefficient to estimate the reliability and criterion validity of the SAYCARE PA questionnaire

PA (min/d)	Q1 - median (25th-75th)	Q2 - median (25th-75th)	Accelerometer - median (25th-75th)	Reliability (Q1 vs. Q2)	Validity (Q1 vs. accelerometer)
Children					
Active commuting	8.6 (2.0-37.1)	10.7 (2.9-21.4)		$r_s = 0.28$	
PA at school	22.0 (10.0-33.8)	20.0 (17.9-33.4)		$\mathbf{r_s = 0.31}$	
PA at leisure time	25.8 (17.1-51.4)	34.3 (17.1-47.1)		$\mathbf{r_s = 0.33}$	
Moderate intensity	48.1 (18.0-105.5)	33.0 (14.3-71.7)	47.8 (53.2-95.9)	$\mathbf{r_s = 0.37}$	$\mathbf{r_s = 0.61}$
Vigorous intensity	43.0 (35.1-132.9)	50.1 (17.7-168.0)	94.9 (56.6-166.4)	$\mathbf{r_s = 0.89}$	$r_s = 0.27$
Weekly total MVPA	103.5 (36.9-195.2)	60.0 (10.0-132.4)	117.3 (63.8-162.5)	$\mathbf{r_s = 0.56}$	$r_s = 0.44$
% of agreement with current PA guidelines, ≥ 60 min/d	50.9 (45.0-56.8)	65.7 (59.4-71.5)	77.4 (67.6-84.9)	**59.2% (k = 0.315)**	51.3% (k = −0.399)
Adolescents					
Active commuting	21.8 (10.0-30.0)	12.6 (4.2-17.1)		$\mathbf{r_s = 0.51}$	
PA at school	18.0 (9.0-25.2)	20.0 (12.0-30.0)		$\mathbf{r_s = 0.63}$	
PA at leisure time	32.9 (17.1-60.0)	25.7 (0.0-64.3)		$\mathbf{r_s = 0.68}$	
Moderate intensity	35.8 (14,3-43.3)	28.1 (12,0-30,0)	59.2 (41.7-79.4)	$\mathbf{r_s = 0.36}$	$r_s = 0.11$
Vigorous intensity	10.0 (7,4-35,6)	16.1 (0.0-35.6)	19.7 (10.0-29.8)	$\mathbf{r_s = 0.93}$	$\mathbf{r_s = 0.65}$
Weekly total MVPA	41.3 (12.0-94.1)	43.1 (8.0-56.6)	44.2 (28.3-63.2)	$\mathbf{r_s = 0.60}$	$\mathbf{r_s = 0.88}$
% of agreement with current PA guidelines, ≥ 60 min/d	28.4 (23.9-33.5)	24.9 (19.0-31.8)	25.8 (16.2-38.4)	**82.9% (k = 0.559)**	**52.7% (k = 0.509)**

Values are min/d, median (25th-75th). Moderate (or above) kappa agreement ($k \geq 0.40$) and Spearman correlation coefficient ($\rho \geq 0.30$) values are in bold.
MVPA, moderate-to-vigorous physical activity; PA, physical activity; Q1, first application of the questionnaire; Q2, second application of the questionnaire.

For validity, the intensity (moderate and vigorous) and total (summary of the three domains) were analyzed. A moderate Spearman correlation ($\rho \geq 0.30$) (16) and kappa coefficient ($k \geq 0.40$) (30) were considered acceptable. Stata software version 14.0 (StataCorp, College Station, Texas) was used for all statistical analyses. The criterion for statistical significance was a two-sided 5%.

Results

Two hundred and thirty-seven children and 215 adolescents made up the initial sample. From them, 82 children and 60 adolescents were randomly selected to measure PA with an accelerometer and thus to test the validity of the questionnaire. Due to other objectives of the research project, we overestimated the sample size. Nevertheless, some of the participants did not adequately complete the two questionnaires and the accelerometry procedure. As a result, 161 children and 177 adolescents for the reliability study and 82 and 60 for the validity study made up the final sample. The division of this sample across the countries is presented in Supporting Information Table S1. Their characteristics are presented in Table 1. For both analyses, the samples of children and adolescents were formed by males and females from public and private schools. In our sample, there were no significant differences for the sex distribution, different ages (only in adolescents), or maternal education (only in children) in either the reliability or validity studies.

Table 2 presents the analysis of PA by domain, intensity, weekly total MVPA, and proportion of individuals who complied with the current recommendation of sufficient PA. In children, the reliability coefficients were acceptable for all analyses and presented a strong value in vigorous intensity, whereas the coefficient of validity showed acceptable value. The kappa coefficients in both tests were moderate. In adolescents, the reliability and validity coefficients were strong ($\rho > 0.60$), and the kappa coefficients were acceptable.

The results of multilevel linear regression analyses are presented in Table 3. In children, the results in both intensity levels had negative biases, showing that the questionnaire underestimated the children's PA level. In adolescents, negative bias was observed only for moderate intensity. In both age groups, the influence of the contextual factor in the prediction of error and the proportional attributable to region-level variance were only significant for moderate intensity in children, MVPA in both age groups, and vigorous intensity in adolescents.

Discussion

Our PA questionnaire showed moderate-to-strong reliability coefficients in both age groups. The validity coefficient was acceptable in children, whereas it was strong and positive in adolescents. These results indicate that the SAYCARE PA questionnaire can be useful to estimate the PA level in the pediatric population because it is a

TABLE 3 Variance between regions and effect of adjustment by multilevel regression models on criterion validity coefficients

PA (min/d)	Adjusted variable for children	Region-level effects[a] Variance	P	PCV (%)	Bias[b]	95% LOA
Moderate intensity	Model 0: city	16.7	0.02	15.0		
	Model 1: Model 0 plus school	15.5	0.019	15.2		
	Model 2: Model 1 plus sex and age	−1.7	0.001	19.1	−13.6	−15.2 to 41.4
Vigorous intensity	Model 0: city	7.9	0.326	62.3		
	Model 1: Model 0 plus school	7.1	0.614	28.1		
	Model 2: Model 1 plus sex and age	36.4	0.128	61.4	−35.3	−36.8 to 56.1
Weekly total MVPA	Model 0: city	104.1	0.137	25.4		
	Model 1: Model 0 plus school	86.4	0.28	36.2		
	Model 2: Model 1 plus sex and age	7.1	0.017	43.4	−22.9	−24.6 to 19.9
PA (min/d)	**Adjusted variable for adolescents**	**Region-level effects[a] Variance**	**P**	**PCV (%)**	**Bias[b]**	**95% LOA**
Moderate intensity	Model 0: city	2.3	0.142	56.5		
	Model 1: Model 0 plus school	−26.6	0.198	54.8		
	Model 2: Model 1 plus sex and age	35.6	0.249	56.8	−19.5	−41.6 to 58.9
Vigorous intensity	Model 0: city	11.2	0.692	54.7		
	Model 1: Model 0 plus school	−1.4	0.778	54.3		
	Model 2: Model 1 plus sex and age	−28.6	0.002	41.6	18.3	−92.6 to 56.0
Weekly total MVPA	Model 0: City	27.1	0.296	70.6		
	Model 1: Model 0 plus school	−24.9	0.169	66.6		
	Model 2: Model 1 plus sex and age	−10.8	0.037	66.7	16.0	−14.2 to 17.4

[a]Significance testing for variance of random effect by likelihood ratio test.
[b]Bias: min/d average difference between methods (questionnaire first application and accelerometer).
LOA, 95% limits of agreement (min/d); MVPA, moderate-to-vigorous physical activity; PA, physical activity; PCV, proportional change attributable to region-level variance; Q1, first application of the questionnaire; Q2, second application of the questionnaire.

powerful tool in epidemiologic research, given that it can be used on a large scale and with high cost effectiveness (5).

Our results showed the important influence of the contextual factor (city) by the proportional attributable variance on the estimated PA level bias. These adjustments should be considered in validity multicenter studies because the contextual factors are associated with PA (25) and can influence the validation estimation, as French researchers have recently shown (4).

Previous studies in pediatric populations have reported similar results (23,30,31), but the limitation in these studies is that the authors evaluated only one city, whereas our study was conducted in seven South American cities for reliability and in four for validity. Questionnaire-based vigorous PA was strongly associated with test-retest reliability in both age groups and with concurrent validity against an accelerometer in adolescents. A possible explanation is that the vigorous PA was performed intentionally and/or at leisure time, which may reduce memory bias at the time of completing the questionnaire, although the intensities have been self-perceived (32).

On the other hand, the moderate intensity presented with moderate-to-acceptable reliability and validity coefficients. Recently, European researchers have shown a low magnitude of association between PA and accelerometers (15), which is possibly because PA at this intensity is less "controlled" and is difficult for both parents and children to report. The measurement of PA in pediatric populations is a challenge, and there is currently no consensus about the reliability and validity of subjective instruments, since there are many questionnaires available in the literature (5) and the measurement of this behavior is necessary to assist in the monitoring and construction of public policies to encourage its practice (33).

Rose, in his classic paper (34), reports that the prevalence of diseases and risk factors depends on the studied population. In a recent review, De Moraes et al. verified that this premise is also applicable to PA (35). This was the motivation for a multilevel analysis to estimate the effect of the contextual variable on the measurement bias of the questionnaire. The results confirmed the hypothesis that contextual variables influence the estimation of PA bias. For these reasons, we consider that country-specific characteristics and contextual variables can determine the PA level in children and adolescents (36).

There were several strengths of this study. It included the comparison of PA questionnaires against an objective measurement (accelerometers) in a multicultural youth group and in low- to middle-income countries, an area of research that has previously been limited; therefore, this study significantly adds to the body of literature in this area of research. Furthermore, adjustments were made for contextual factors, including the city and school level. This study also uses a comparison of PA questionnaires and accelerometry with commonly used protocols. Previously, the individuals used accelerometers for data to match the protocol for questionnaire data collection, and the individuals answered the questionnaire for a second time after 15 days to evaluate the test-retest reliability of the questionnaire. We decided to test the agreement using real-world protocols.

Despite its strengths, the study also has some limitations. This study used self-reports and proxy-reports that could have systematic errors associated. There was a considerable percentage of participants who were lost or refused to complete participation (53.6%); however, the final sample size was large enough to test the reliability in children and adolescents and the validity in children, and the final sample size was acceptable to perform validation analysis in adolescents (37). Another possible limitation was the selection of a sample by convenience. However, since the convenient selection comprised public and private schools, this kind of choice should not have influenced the validity and reliability results.

Conclusion

The SAYCARE PA questionnaire has moderate accuracy in youth. In children, the SAYCARE questionnaire shows acceptable reliability for PA measurement; conversely, there was moderate validity and systematic underreporting. In adolescents, the findings suggest that the questionnaire measure can provide a reliable, valid assessment of PA. O

Acknowledgments

All authors acknowledge the dean or chair of each school as well as the children, adolescents, and parents for their voluntary participation in the SAYCARE study. We also thank Mr. Andrew Mello Silva for managing the SAYCARE database. In Colombia, we thank the undergraduate students Carolina Alzate Echeverri, Luisa Fernanda Arroyave Zuleta, Sarah Lucía Ortiz Calderón, and Paola Zapata for helping collect information and for data entry. All authors thank the General Coordinator of the Growth, Exercise, Nutrition and Development Research Group, Luis A. Moreno (MD, PhD, full professor) from the Faculty of Health Sciences, University of Zaragoza, for the loan and help with the accelerometers for data collection.

References

1. Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012;380:219-229.
2. Bammann K, Sioen I, Huybrechts I, et al. The IDEFICS validation study on field methods for assessing physical activity and body composition in children: design and data collection. *Int J Obes (Lond)* 2011;35(suppl 1):S79-S87.
3. Chinapaw MJ, Mokkink LB, van Poppel MN, van Mechelen W, Terwee CB. Physical activity questionnaires for youth: a systematic review of measurement properties. *Sports Med* 2010;40:539-563.
4. Lawman HG, Horn ML, Wilson DK, Pate RR. A multilevel approach to examining time-specific effects in accelerometer-assessed physical activity. *J Sci Med Sport* 2015;18:667-672.
5. de Moraes AC, Fernandes CA, Elias RG, Nakashima AT, Reichert FF, Falcão MC. Prevalence of physical inactivity and associated factors in adolescents [in Portuguese]. *Rev Assoc Med Bras* 2009;55:523-528.
6. Crespo NC, Sallis JF, Conway TL, Saelens BE, Frank LD. Worksite physical activity policies and environments in relation to employee physical activity. *Am J Health Promot* 2011;25:264-271.
7. Bielemann RM, Reichert FF, Paniz VM, Gigante DP. Validation of the Netherlands physical activity questionnaire in Brazilian children. *Int J Behav Nutr Phys Act* 2011;8:45. doi:10.1186/1479-5868-8-45
8. Adami F, Bergamaschi DP, Hinnig PeF, Oliveira NS. Validity study of the "Physical Activity Checklist" in children [in Portuguese]. *Rev Saude Publica* 2013; 47:488-496.
9. Godard M C, Rodríguez N MeP, Díaz N, Lera M L, Salazar RG, Burrows AR. Value of a clinical test for assessing physical activity in children [in Spanish]. *Rev Med Chil* 2008;136:1155-1162.
10. Moreno LA, De Henauw S, González-Gross M, et al. Design and implementation of the Healthy Lifestyle in Europe by Nutrition in Adolescence Cross-Sectional Study. *Int J Obes (Lond)* 2008;32(suppl 5):S4-S11.
11. Ahrens W, Bammann K, de Henauw S, et al. Understanding and preventing childhood obesity and related disorders--IDEFICS: a European multilevel epidemiological approach. *Nutr Metab Cardiovasc Dis* 2006;16:302-308.
12. Trost SG, Fees BS, Haar SJ, Murray AD, Crowe LK. Identification and validity of accelerometer cut-points for toddlers. *Obesity (Silver Spring)* 2012;20:2317-2319.
13. Helmerhorst HJ, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *Int J Behav Nutr Phys Act* 2012;9:103. doi:10.1186/1479-5868-9-103
14. Lanfer A, Hebestreit A, Ahrens W, Krogh V, et al. Reproducibility of food consumption frequencies derived from the Children's Eating Habits Questionnaire used in the IDEFICS study. *Int J Obes (Lond)* 2011;35(suppl 1):S61-S68.
15. Verbestel V, De Henauw S, Bammann K, et al. Are context-specific measures of parental-reported physical activity and sedentary behaviour associated with accelerometer data in 2-9-year-old European children? *Public Health Nutr* 2015;18: 860-868.
16. Strong W, Malina R, Blimkie C, et al. Evidence based physical activity for school-age youth. *J Pediatr* 2005;146:732-737.
17. US Department of Health and Human Services. Key guidelines for children and adolescents. https://health.gov/paguidelines/guidelines/chapter3.aspx. Published 2008. Accessed August 11, 2016.
18. Ojiambo R, Cuthill R, Budd H, et al. Impact of methodological decisions on accelerometer outcome variables in young children. *Int J Obes (Lond)* 2011;35 (suppl 1):S98-S103.
19. Ruiz J, Ortega F, Martinez-Gomez D, et al. Objectively measured physical activity and sedentary time in european adolescents: the HELENA study. *Am J Epidemiol* 2011;174:173-184.
20. Hagströmer M, Bergman P, De Bourdeaudhuij I, et al. Concurrent validity of a modified version of the International Physical Activity Questionnaire (IPAQ-A) in European adolescents: the HELENA study. *Int J Obes (Lond)* 2008;32 Suppl 5:S42-S48.
21. Reis R, Petroski E, AS L. Medidas da atividade física: revisão de métodos. *Rev Bras Cineantropom Desempenho Humano* 2000;2:89-96.
22. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. *J Sports Sci* 2008;26:1557-1565.
23. Scott JJ, Morgan PJ, Plotnikoff RC, Lubans DR. Reliability and validity of a single-item physical activity measure for adolescents. *J Paediatr Child Health* 2015;51: 787-793.
24. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;45:255-268.
25. Ding D, Sallis JF, Kerr J, Lee S, Rosenberg DE. Neighborhood environment and physical activity among youth a review. *Am J Prev Med* 2011;41:442-455.
26. de Moraes AC, Fernández-Alvira JM, Rendo-Urteaga T, et al. Effects of clustering of multiple lifestyle-related behaviors on blood pressure in adolescents from two observational studies. *Prev Med* 2016;82:111-117.
27. Greenland S. Modeling and variable selection in epidemiologic analysis. *Am J Public Health* 1989;79:340-349.
28. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-310.
29. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.
30. Zaragoza Casterad J, Generelo E, Aznar S, Abarca-Sos A, Julián JA, Mota J. Validation of a short physical activity recall questionnaire completed by Spanish adolescents. *Eur J Sport Sci* 2012;12:283-291.
31. Hallal PC, Reichert FF, Clark VL, et al. Energy expenditure compared to physical activity measured by accelerometry and self-report in adolescents: a validation study. *PLoS One* 2013;8:e77036. doi:10.1371/journal.pone.0077036

32. Strath SJ, Bassett DR, Swartz AM. Comparison of the college alumnus questionnaire physical activity index with objective monitoring. *Ann Epidemiol* 2004;14:409-415.
33. Kelly P, Fitzsimons C, Baker G. Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. *Int J Behav Nutr Phys Act* 2016;13:32. doi:10.1186/s12966-016-0351-4
34. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;14:32-38.
35. de Moraes AC, Guerra PH, Menezes PR. The worldwide prevalence of insufficient physical activity in adolescents; a systematic review. *Nutr Hosp* 2013;28:575-584.
36. Sirard JR, Laska MN, Patnode CD, Farbakhsh K, Lytle LA. Adolescent physical activity and screen time: associations with the physical home environment. *Int J Behav Nutr Phys Act* 2010;7:82. doi:10.1186/1479-5868-7-82
37. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651-657.