Validity and reliability evidence of the physical activity level screening for preschoolers with developmental delays and/or disabilities

Changwoong Park¹, So-Yeun Kim^{2,*}, Lauriee L. Zittel², Marilyn Looney²

¹Department Kinesiology and Public Health Education, College of Education and Human Development, University of North Dakota, Grand Forks, ND, USA

²Department of Kinesiology and Physical Education, College of Education, Northern Illinois University, DeKalb, IL, USA

Psychometric properties of the physical activity level screening (PALS) for preschoolers with a developmental delay and/or a disability were examined in this study. The mean age of participants (16 boys and 14 girls) was 54.9 months (standard deviation, 10.4). Each participant wore one Actical accelerometer and was filmed for 20 min during the active play time. Intrarater and interrater reliability, and validity of the PALS

data were examined. The proportion of agreement for intrarater was 87.5% and 85.6% for interrater reliability. Intraclass correlation coefficients were 0.99 for intrarater and 0.96 for interrater. The PALS data showed moderate correlation with Actical data ($r=0.67$, $P<0.01$).

Keywords: Psychometric properties, Preschool, Systematic observation

INTRODUCTION

Insufficient daily physical activity (PA) in children has increased the risk for poor health (Nielsen et al., 2010). To understand PA behaviors of preschoolers and monitor their PA levels during the designated active play time in early childhood programs, it is important to measure PA levels of preschoolers. However, there is limited research on PA measurements (Oliver et al., 2007) for the preschoolers with and without disabilities, and many PA measurement instruments available today are not developed for preschoolers (Pate et al., 2008).

The physical activity level screening (PALS) is a systematic observation screening instrument designed to measure the PA levels of preschoolers during their active play time (Zittel et al., 2011). The PALS was developed to be practical and user friendly for early childhood practitioners, so they can use the tool to identify preschool-age children who are not physically active. However, reliability and/or validity evidence of the tool for preschoolers with developmental delays and/or disabilities has not been examined.

The purpose for the study, therefore, was to determine validity and reliability evidence of the PALS for preschoolers with developmental delay(s) and/or disabilities. Specific research questions include (a) what was the intrarater and interrater reliability of researchers' PALS data? (b) What was the validity evidence of the PALS data when it correlates with the accelerometer data?

MATERIALS AND METHODS

Participants

Thirty preschoolers (37–69 months; mean $\pm$ standard deviation [SD], 54.90 $\pm$ 10.4 months) with a developmental delay and/or a disability was recruited from an early childhood center after receiving the approval from a university Institutional Review Board. The participants' parents/legal guardians provided written informed consent. The demographic information of the participants is listed in Table 1. The participants had at least one or more disability diagnoses (speech impairment, developmental delay, learning disability, emotional disturbance, Down syndrome, autism, and/or other health impairments).

*Corresponding author: So-Yeun Kim <http://orcid.org/0000-0001-5779-7450>
Department of Kinesiology and Physical Education, College of Education,
Northern Illinois University, DeKalb, IL, USA
Tel: +1-815-753-7152, Fax: +1-815-753-1413, E-mail: soyeunkim@niu.edu
Received: December 25, 2016 / Accepted: March 28, 2017

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. Demographic information of the participants

Age (yr)	Boys (n = 16)			Girls (n = 14)		
	Caucasian	African American	Hispanic	Caucasian	African American	Hispanic
3	3	0	0	3	0	1
4	6	0	0	3	1	1
5	5	0	2	5	0	0
Total, n (%)	14 (87.5)	0 (0)	2 (12.5)	11 (78.6)	1 (7.1)	2 (14.3)

Instruments

Physical activity level screening

The PALS (Zittel et al., 2011) was developed to measure the PA levels of preschoolers during daily designated active play time, and uses a momentary time sampling method; coding occurs every 15 sec. The PALS uses modified PA contexts from the Children's Activity Rating Scale (CARS) (Puhl et al., 1990). The CARS has a 5-point scale and each activity level is coded once within a given minute. The PALS simplified the CARS point scale and reduced the PA levels down to 3 instead of 5 to be user friendly and practical for practitioners to screen activity levels. To record PA levels for the PALS, level 2, is the highest level used to record moderate and vigorous PA. Level 1 used to record light PA such as slow/light walking, or any torso movements to initiate light PA movements. Level 0 was for none (completely sedentary) to minimal activity that involve no torso movements.

Accelerometer

Actical accelerometers (Mini-Mitter, Bend, OR, USA) were used to validate the PALS data. The accelerometers are uniaxial devices, and have been validated to measure the PA levels of preschool-aged participants. McIver et al. (2005) reported a strong relationship ($r = 0.86$) between data from a portable metabolic analyzer (VO_2) and Actical accelerometers. Since the shorter epoch time is more appropriate to measure preschoolers' PA levels (Vale et al., 2009), a 15-sec epoch setting (the shortest setting for Actical accelerometers) was used for this study.

Procedures

For familiarization process, accelerometers and a video recording device (DCXR-350V, Sony, Tokyo, Japan) were also introduced to the participants during the classroom circle time. The participants wore an accelerometer at their waist level using Velcro belts and were asked to move around in the classroom for approximately 5 min, performing various locomotor skills, so they would be comfortable with the accelerometers and the belt for the actual data collection. During the actual data collection, each par-

ticipant was asked to wear one Actical accelerometer for at least 20 min during the designated active play time and was filmed. Participants wore an accelerometer at their waist level, in line with the middle of the right thigh using a Velcro belt.

Observer training protocol and data coding

Two investigators received systematic observation training from two experts on the systematic observation tools. The training consisted of reading about systematic observation and learning about the PALS. The PA levels of the PALS were extensively explained and discussed with the experts. The investigators met with the experts on two occasions to receive the PALS training and the average meeting lasted about 2–3 hr. The expert provided video examples of different PA levels to assure clear understanding of the different PA levels of the PALS.

For training to establish strong interrater reliability between two raters, a total of three video files (30 min each) of children from the pilot study were randomly selected, and PA levels were coded independently. The average proportion of agreements between the two investigators' PALS data was 90.7% (SD, 5.6%; range, 84.3%–94.7%).

Two investigators independently coded each participant's video data using the PALS. After 1 week of the first coding, the primary investigator recoded randomly selected 10 video files (about 30% of total film data) to examine intrarater reliability. For interrater reliability, another investigator coded randomly selected 10 video files.

Statistical analyses

The investigator's first and second PALS coding data (both total and average intensity level of PA data) were compared to examine intrarater reliability. A difference within the rater was calculated using the average proportion of agreement. Also, intraclass coefficients with a 95% confidence interval [CI] was performed using IBM SPSS Statistics ver. 20.0 (IBM Co., Armonk, NY, USA).

To examine interrater reliability, two investigator's PALS data (both total and average intensity level of PA data) were compared. A difference between the raters was calculated using the average

proportion of agreement and the average of the modified Cohen kappa. Also, an intraclass coefficient with a 95% CI was examined.

A Pearson correlation coefficients were used to determine criterion validity evidence of the PALS data by comparing total intensity levels of PA from the PALS data and total activity counts from Actical accelerometers. IBM SPSS Statistics ver. 20.0 (IBM Co., Armonk, NY, USA) was used to conduct the analysis.

RESULTS

The average proportion of agreements between the primary investigator's first and second PALS data was 0.875 (SD, 0.36; range, 0.791–0.902). The result of a one-way repeated measures analysis of variance for total PA levels showed significant difference, $F(1, 9) = 23.2$, $P = 0.001$, between the first (mean $\pm$ SD, 688.6 ± 264.2) and the second (mean $\pm$ SD, 649.6 ± 249.2) PALS coding data. Intraclass correlation coefficient for a single measure was 0.99 (95% CI: 0.98–0.99).

The average proportion of agreements between two investigator's data was 0.856 (SD, 0.0476; range, 0.789–0.915). The mean of modified Cohen kappa was 0.78 (SD, 0.7; range, 0.68–0.87). Intraclass correlation coefficient for a single measure was 0.96 (95% CI, 0.85–0.99).

A Pearson correlation coefficient of total PA levels from the primary investigator's PALS data (mean $\pm$ SD, 729.9 ± 284.9) and total activity counts (mean $\pm$ SD, $42,098.5 \pm 27,843.5$) showed a moderate correlation of $r = 0.665$, $P < 0.01$ (95% CI, 0.40–0.83).

DISCUSSION

Reliability and validity evidence of the PALS data on preschoolers with developmental delays and/or disability was examined in this study. There were strong intrarater and interrater reliability correlation coefficients as well as high proportion of agreements between raters and within a rater. Regarding validity evidence, the PALS data demonstrated a moderate correlation with Actical accelerometer data.

Both the average of the proportion of agreements of intrarater and interrater were higher than 85% for this study. Similar results of the proportion of agreement between raters have been reported with other systematic observation tools in PA for preschoolers without disabilities including the CARS for 84.1% (Puhl et al., 1990), the Observational System for Recording Physical Activity in Children Preschool Version (OSRAC-P) for 90% (Brown et al., 2006), and System for Observing Fitness Instruction Time for Preschoolers

(SOFIT-P) for 92% (Sharma et al., 2011). Intraclass correlation coefficients for both intrarater and interrater were high, above $r = 0.90$.

To further investigate the reliability of the PALS data, modified Cohen kappa was calculated to see if the similarity between raters occurred by chance. The modified Cohen kappa (Warrens, 2010) for this study was 0.78, supporting strong reliability evidence of the trained users' PALS data on preschoolers with a developmental delay and/or a disability. Watkins and Pacheco (2000) stated that the kappa values higher than 0.75 represent substantial agreements.

The PALS data demonstrated moderately positive criterion validity when it was compared with the data from Actical accelerometers ($r = 0.665$, $P < 0.01$). Similar results were reported in previous studies with preschoolers without disabilities. Sharma et al. (2011) examined validity evidence of the SOFIT-P on preschoolers without disabilities using Actigraph GT3X accelerometers. The authors reported a moderate positive correlation (moderate PA, $r = 0.53$; moderate-vigorous PA, $r = 0.54$) between the SOFIT-P and the data from accelerometers. Noland et al. (1990) found a positive correlation between the CARS data and Caltrac accelerometers ($r = 0.86$, $P < 0.001$) data on preschoolers.

The moderate relationship with the data from Actical accelerometers may be for a variety of reasons. Even though the accelerometers are validated to measure PA levels for preschoolers (McIver et al., 2005), one of the main limitations of the Actical and other accelerometers is that the device is worn at the hip level and may underestimate or fail to detect PA with limited torso movement (Oliver et al., 2007; Oliver et al., 2009). Actical accelerometers underestimated PA levels when the participants in this study engaged in climbing activities and other limited torso movement activities such as tricycle riding and sitting on a stationary toy horse.

In conclusion, the results of this study indicate that the PALS data demonstrated a moderate and positive correlation compared with Actical accelerometers. Also, acceptable intrarater and interrater reliabilities were found for the PALS data by trained observers although the training of observers for the PALS was relatively shorter than the other systematic observation tools measuring PA for preschoolers (e.g., approximately 7 days for the PALS vs. 7 weeks for OSRAC-P). The PALS would be a useful screening tool to monitor the PA levels of preschoolers with developmental delays and/or disabilities during active play time.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

- Brown WH, Pfeiffer KA, McIver KL, Dowda M, Almeida MJ, Pate RR. Assessing preschool children's physical activity: the Observational System for Recording Physical Activity in children-preschool version. *Res Q Exerc Sport* 2006;77:167-176.
- McIver KK, Pfeiffer KK, Almeida JJ, Dowda MM, Pate RR. Validity of the actigraph and actual accelerometers in 3- to 5 year-old children [abstract]. *Pediatr Exerc Sci* 2005;17:97.
- Nielsen G, Taylor R, Williams S, Mann J. Permanent play facilities in school playgrounds as a determinant of children's activity. *J Phys Act Health* 2010;7:490-496.
- Noland M, Danner F, DeWalt K, McFadden M, Kotchen JM. The measurement of physical activity in young children. *Res Q Exerc Sport* 1990;61:146-153.
- Oliver M, Schofield GM, Kolt GS. Physical activity in preschoolers: understanding prevalence and measurement issues. *Sports Med* 2007;37:1045-1070.
- Oliver M, Schofield GM, Schluter PJ. Accelerometry to assess preschooler's free-play: Issues with count thresholds and epoch durations. *Meas Phys Educ Exerc Sci* 2009;13:181-190.
- Pate RR, McIver K, Dowda M, Brown WH, Addy C. Directly observed physical activity levels in preschool children. *J Sch Health* 2008;78:438-444.
- Puhl J, Greaves K, Hoyt M, Baranowski T. Children's Activity Rating Scale (CARS): description and calibration. *Res Q Exerc Sport* 1990;61:26-36.
- Sharma S, Chuang RJ, Skala K, Atteberry H. Measuring physical activity in preschoolers: Reliability and validity of The System for Observing Fitness Instruction Time for Preschoolers (SOFIT-P). *Meas Phys Educ Exerc Sci* 2011;15:257-273.
- Vale S, Santos R, Silva P, Soares-Miranda L, Mota J. Preschool children physical activity measurement: importance of epoch length choice. *Pediatr Exerc Sci* 2009;21:413-420.
- Warrens MJ. A formal proof of a paradox associated with Cohen's kappa. *J Classif* 2010;27:322-332.
- Watkins MW, Pancheco M. Interobserver agreement in behavioral research: importance and calculation. *J Behav Educ* 2000;10:205-212.
- Zittel L, Kim SY, Park C. Assessing preschool children's physical activity using PALS. In: 40th National Adapted Physical Education Conference; 2011 Dec 17-19; Long Beach (CA), USA.