Sources and magnitude of variability in pedometer-determined physical activity levels of youth

So-Yeun Kim1,*, Hye-Sang Park2

1Department of Kinesiology and Physical Education, College of Education, Northern Illinois University, DeKalb, IL, USA
2Department of Sports Health Management, College of Physical Education, University of Suwon, Suwon, Korea

This study examined sources of variability in habitual physical activity of children, and the minimum number of days required for estimating children’s habitual physical activity levels using pedometers. A total of 31 children wore two pedometers during five weekdays and four weekend days. A two random facet completely crossed design was conducted with two-way analysis of variances across weekdays, weekends, and weekdays and weekend days combined. Moderate/high generalizability coefficients were estimated across all days. Primary sources of variability were variance components of the person and person by day interaction. Minimum numbers of days required for estimating habitual physical activity levels using a pedometer were five during weekdays. However, estimating habitual physical activity levels during weekends, and weekdays and weekend days combined was impractical.

Keywords: Reliability, Physical activity, Pedometer, Child, Measurement

INTRODUCTION

One of the challenges to determine habitual physical activity (HPA) levels of children using pedometers is intraindividual (Baranowski et al., 2008; Baranowski and de Moor, 2000; Levin et al., 1999) and interinstrument variability (Baranowski and de Moor, 2000). Intraindividual variability in HPA indicates differences in physical activity levels in a person from day to day while interinstrument variability occurs when physical activity measures of the same brand provide different data under the same set of circumstances. Intraindividual variability in the HPA of children has been estimated by identifying the minimum number of days of measurement required to get representative information of HPA levels in a given population. Interinstrument variability in pedometers was examined during controlled environments including (a) manual shaking tests (Abel et al., 2009), (b) walking with pedometers for predetermined distances (Bassett Jr et al., 1996), and (c) engaging predetermined physical activities during certain amount of time (Jago et al., 2006). Limited information is available for interinstrument variability of pedometers in daily life. One of the issues of current approaches to HPA levels in children is that intraindividual and interinstrument variability have been examined separately. To understand true changes of HPA, both intraindividual variability and interinstrument variability should be examined at the same time. Both intraindividual and interinstrument variability can be examined at the same time using generalizability theory (G-theory). Recently, researchers in the field of exercise sciences have proposed using G-theory in the measurement of physical activity studies (Baranowski et al., 2008). The G-theory approach extends classical test theory, and allows researchers to estimate multiple sources of error instead of one source (Goodwin, 2001). Using G-theory, optimal measurement protocols for the estimation of HPA can be provided by estimating what the reliability for a condition would be, if more or fewer items (such as the number of days and instruments in the case of physical activity studies) were added. G-theory techniques have not been used extensively in the measurement of HPA for children using pedometers. Although Wickel and Welk (2010) deter...
determined sources of measurement error required estimating long-term levels of physical activity for youth, source of errors related to instruments (pedometers) were not included in the design. Therefore, the purpose of this study was to examine sources and magnitude of variability in the measurement of HPA levels of children using pedometers, and provide the minimum number of days required for monitoring HPA levels to determine their typical ambulatory physical activity levels.

MATERIALS AND METHODS

Participants

Thirty-one middle school students (15 boys and 16 girls) in grade 6 (mean ± standard deviation, 12.2 ± 0.6 yr) was recruited from six different physical education classes in the rural area of Northwestern United States. Students and their parents, who signed written informed consent forms, participated in the study. This study was approved by the University Institution Review Board prior to recruiting participants.

Measures

Omron HJ-112 pedometers (Omron Healthcare Inc., 2003; Vernon Hills, IL, USA) were used in this study. The pedometer can store steps for seven days. To examine the proper calibration of pedometers, the “shake-test” developed by Vincent and Sidman (2003) was performed. Each pedometer was placed in a shipping box from the manufacturer, and shaken 100 times. These tests were performed twice for each pedometer before and after the study. No pedometers exceeded ±5% error.

Experimental procedures

All the participants were asked to wear two pedometers on their waists, in line with the middle of their right thigh, for 11 days including seven weekdays and four weekend days, from the time that they got up in the morning until bedtime, except while swimming and showering, since the pedometers were not waterproof. The first two weekdays were for familiarizing participants with wearing pedometers to decrease reactivity that might influence the participants’ physical behaviors. To help each participant wear two pedometers on the correct place, two pedometers were attached on a Velcro belt. In order to prevent inflating counts, all participants were instructed not to tamper with the instruments, and the pedometers were sealed.

Statistical analyses

G-theory was used in this study, including generalizability study (G-study) and decision study (D-study). A G-study is designed to provide estimates of the variance components associated with each facet and their interactions while a D-study is designed to make substantive decisions about a measurement protocol using the results of the G-study (Shavelson et al., 1989).

G-study

A two-facet fully crossed design was employed to answer the first research questions about the sources of variability of HPA. Two random facets were instruments and days. Three two-way analysis of variances were used to examine the sources of variability in HPA levels across weekdays, weekends, and combined weekdays and weekend days. Average steps per day were used as the dependent variable. Seven variance components were estimated using the VARCOMP procedure from SAS ver. 9 (SAS Institute Inc., Cary, NC, USA). The percentage of variance associated with each source of variability was calculated by dividing each variance estimate by the total variance. The seven sources of variability include variance associated with persons, days, and instruments, three two-way interaction including persons by days, persons by instruments, and instruments by days, and the residual term (three-way interaction plus error).

Generalizability coefficients, G and phi coefficients were also calculated from the estimated variance components. Generalizability coefficients indicate how accurate the generalization is from an individual’s observed score to his or her true universe score. G coefficients are calculated based on relative decisions (decision based on the relative standing of individual) while phi coefficients are calculated based on absolute decisions (based on the absolute level of performance) (Shavelson et al., 1989). In relative decisions, variance components of person by instrument, person by day, and residual term contribute to error while all variance components except person are considered an error in absolute decisions (Shavelson et al., 1989).

D-study

The second research question about the optimal number of days required for monitoring HPA levels was examined using a D-study. By increasing or decreasing the number of instruments and days, the minimum number of facet levels required to establish the desired generalizability was assessed.

RESULTS

Only participants with complete data during weekdays week-
ends, and weekdays and weekend combined were used. The average steps per weekday, weekend day, and weekday and weekend days were 9,432 ± 3,306 (n = 31), 6,344 ± 2,414 (n = 30), and 7,933 ± 2,851 (n = 30) steps, respectively.

**G-study**

Variance component estimates and relative magnitudes for HPA levels are presented in Table 1. During weekdays, the largest source of variability was persons (53.16%), and the second largest source of variability was persons by days interactions (44.04%). During weekends, the largest source of variability was the persons by days interaction (57.95%), and the second largest source of variability was the persons (39.06%). During weekdays as well as weekends, the other variance components were very low. When data of weekdays and weekend days were combined, the primary sources of variability were the persons by days interaction (52.61%) and the persons (32.93%) while the days facet was also associated with 13.71% of total variance.

The estimated generalizability coefficients for the average data from two pedometers across five weekdays was high (G = 0.85, phi = 0.85). Moderate generalizability coefficients (G = 0.73, phi = 0.72) was estimated for the average data of two pedometers across four weekend days. When data from weekdays and weekend days were combined, high generalizability coefficients (G = 0.85, phi = 0.82) was also estimated for the average data from two pedometers across nine days.

**D-study**

To achieve sufficient reliability (G ≥ 0.80, phi ≥ 0.80) in the measurement of HPA during weekdays, measurements need to be obtained on at least five days using a pedometer. To determine HPA levels of children during weekend days, at least nine weekend days of measurement using a pedometer was estimated to be required. When data from weekdays and weekends were combined, at least 14 days of measurements using three pedometers were needed. Fig. 1 shows estimated phi coefficients across days using one pedometer.

**DISCUSSION**

In this study, the sources and magnitude of variability in ambulatory based HPA of children were examined using G-study. The results of G-study indicated that variability in HPA of children during weekdays, weekends, and weekdays/weekends was related to true physical activity changes rather than measurement errors. The source of variability contributing the most to the total variance was the persons and the persons by days interaction components across weekdays, weekends, and weekdays/weekends. This suggested that variability of HPA levels is related to differences among participants’ physical activity levels as well as inconsistencies from one day to another in particular person’s physical activity levels. Participating in sports on certain days during the week...

*Table 1. Variance component estimates and relative magnitudes*

| Variable          | Estimated variance components | Relative magnitude (%) |
|-------------------|-------------------------------|------------------------|
| **Weekdays**      |                               |                        |
| Persons (p)       | 4,608,384                     | 49.13                  |
| Instruments (i)   | -391.32                       | 0                      |
| Days (d)          | 55,071.40                     | 0.59                   |
| p×i               | 39,502.70                     | 0.42                   |
| p×d               | 4,508,880                     | 48.07                  |
| i×d               | 1,566.80                      | 0.02                   |
| Residual (p×i×d, e) | 166,355.30               | 1.77                   |
| **Weekends**      |                               |                        |
| Persons (p)       | 3,820,066.60                  | 32.33                  |
| Instruments (i)   | 6,576.70                      | 0.06                   |
| Days (d)          | 80,477.80                     | 0.68                   |
| p×i               | 73,066.80                     | 0.62                   |
| p×d               | 7,232,793.20                  | 66.21                  |
| i×d               | 1,343.20                      | 0.01                   |
| Residual (p×i×d, e) | 11,043.30               | 0.09                   |
| **Weekdays/weekends** |                           |                        |
| Persons (p)       | 3,060,364.40                  | 25.13                  |
| Instruments (i)   | 874,260.80                    | 7.18                   |
| Days (d)          | 91,772.70                     | 0.75                   |
| p×i               | 606,530.50                    | 4.98                   |
| p×d               | 2,061,636.40                  | 16.93                  |
| i×d               | 1,193,803.00                  | 9.80                   |
| Residual (p×i×d, e) | 4,290,172.80               | 35.23                  |

Note: Relative magnitude was calculated using estimated variance divided by the total variance. Negative variance components were set to zero in subsequent calculations as suggested by Morrow (1989).
could be reasons for the large persons by days interaction. Some participants regularly participated in a certain physical activity, two or three times a week. Also, during physical education classes, participants often had to decide on a particular physical activity from several activity choices including archery, track and field, basketball, baseball, rock climbing, and weight training. Thus, depending on what kind of physical activity the participants chose during physical education, some participants might have higher activity counts on some days.

Relatively small variance components of the instruments, the persons by instruments interactions, and the instruments by days interactions indicate that low interinstrument reliability, and consistent low differences between steps of two pedometers across days and persons. The results suggest that Omron HJ pedometers have high reliability evidence for the measurement of physical activity in children during free living settings. When data from weekdays and weekend days were combined, large residual components were associated with 35.23% of the total variance. This indicated that a substantial amount of the variance was related to the three-way interaction and to unexplained sources of error not measured in the study. The variance found could be due to different physical activity patterns between weekdays and weekends, which indicate physical activity levels during weekdays should not be interchangeable with physical activity during weekends. Further research should examine differences of daily physical activity patterns during weekdays and weekends. If there is a difference, “week” should be a unit of analysis rather than “day” (Baranowski and de Moor, 2000).

Minimum number of days required for monitoring HPA levels was examined using a D-study to achieve generalizability coefficients of 0.80. Different results were found across weekdays, weekends, and weekdays/weekends. During weekdays, the results of D-study indicate that at least five days of monitoring HPA levels using one pedometer, or four days using two pedometers was required to determine the HPA levels of children. To determine HPA levels of children during weekends, at least nine weekend days of monitoring physical activity using one pedometer were estimated to be needed. When weekdays and weekend days were combined, estimating HPA of children using one pedometer was problematic to achieve generalizability coefficients of 0.80 within reasonable length of time.

To decrease the minimum number of days monitoring necessary, different units of measurement of HPA should be considered. Monitoring only during the day from in the morning to afternoon, when children are expected to be physically active, might produce less variability than measurement throughout the weekend day. Further studies should examine physical activity patterns between boys and girls, and during weekdays and weekend days. If there are differences, then the minimum number of days required for estimating HPA of boys and girls, and during weekdays and weekends should be determined separately.

CONFLICT OF INTEREST

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

REFERENCES

Abel MG, Hannon JC, Eisenman PA, Ransdell LB, Pett M, Williams DP. Waist circumference, pedometer placement, and step-counting accuracy in youth. Res Q Exerc Sport 2009;80:434-444.

Baranowski T, de Moor C. How many days was that? Intra-individual variability and physical activity assessment. Res Q Exerc Sport 2000;71 Suppl 2:74-78.

Baranowski T, Masse LC, Ragan B, Welk G. How many days was that? We’re still not sure, but we’re asking the question better! Med Sci Sports Exerc 2008;40(7 Suppl):S544-S549.

Bassett DR Jr, Ainsworth BE, Leggett SR, Mathien CA, Main JA, Hunter DC, Duncan GE. Accuracy of five electronic pedometers for measuring distance walked. Med Sci Sports Exerc 1996;28:1071-1077.

Goodwin LD. Interrater agreement and reliability. Meas Phys Educ Exerc Sci 2001;5:13-34.

Jago R, Watson K, Baranowski T, Zakeri I, Yoo S, Baranowski J, Conry K. Pedometer reliability, validity and daily activity targets among 10- to 15-year-old boys. J Sports Sci 2006;24:241-251.

Levin S, Jacobs DR Jr, Ainsworth BE, Richardson MT, Leon AS. Intra-individual variation and estimates of usual physical activity. Ann Epidemiol 1999;9:481-488.

Omron Healthcare Inc. (2003). Omron instruction manual: Pedometer Model HJ-112 [Internet]. Vernon Hills, IL: Omron Healthcare Inc.; c2003 [cited 2015 Dec 14]. Available from: http://ec1.images-amazon.com/media/3d/d/01/A/man-migrate/MANUAL000039802.pdf.

Shavelson RJ, Webb NM, Rowley GL. Generalizability theory. Am Psychol 1989;44:922-932.

Vincent SD, Sidman CL. Determining measurement error in digital pedometers. Meas Phys Educ Exerc Sci 2003;7:19-24.

Wickel EE, Welk GJ. Applying generalizability theory to estimate habitual activity levels. Med Sci Sports Exerc 2010;42:1528-1534.