

Running head: PHYSICAL ACTIVITY QUESTIONNAIRE VALIDITY

Convergent Validity and Test-Retest Reliability of the Oxford Physical Activity Questionnaire for
Secondary School Students

David R. Lubans

University of Newcastle

Kathy Sylva

University of Oxford

Zane Osborn

University of Newcastle

Corresponding author:

David Lubans

University of Newcastle

Faculty of Education & Arts

Callaghan Campus

NSW 2308

Australia

Email: David.Lubans@newcastle.edu.au

Telephone: +61249212049

Fax: +61249217407

Abstract

The purpose of this study was to examine the convergent validity and test-retest reliability of the *Oxford Physical Activity Questionnaire* (OPAQ), a self-administered questionnaire designed to assess the activity patterns of adolescents. The test-retest reliability of the OPAQ was assessed by administering the questionnaire on two occasions separated by one week ($N = 87$, mean age = $13.1 \pm .9$). Intraclass Correlation (ICC) was used to examine the test-retest reliability of the OPAQ. The convergent validity of the OPAQ was evaluated using *Caltrac* accelerometers worn consecutively for 4 days ($N = 51$, mean age = $12.6 \pm .5$). Spearman's Rank Order Correlation was used to examine the convergent validity of the questionnaire. ICCs ranged from .76 to .91 and reliability was higher for males ($r = .89$) than females ($r = .78$). Correlations between self-reported physical activity and *Caltrac* accelerometer counts were related to vigorous physical activity ($r = .33$, $p = .01$) and moderate to vigorous activity ($r = .32$, $p = .02$). The OPAQ has excellent test-retest reliability and acceptable validity in comparison to other measures of youth physical activity.

Convergent Validity and Test-retest Reliability of the Oxford Physical Activity Questionnaire for Secondary School Students

The measurement of physical activity is an important and challenging venture (Sallis & Saelens, 2000). Accurate assessment is necessary to identify current levels of activity, assess the effectiveness of intervention programs (Sirard & Pate, 2001) and identify key health associations (Kriska & Casperson, 1997). Physical activity can be defined as “*any bodily movement produced by skeletal muscles that results in energy expenditure*” (Caspersen, Powell, & Christenson, 1985, p. 26, p. 126). Based on this definition, direct observation should be regarded as the ‘gold standard’ of physical activity. However, doubly labeled water (DLW) method and indirect calorimetry are generally regarded as the criterion measures for physical activity because they measure energy expenditure (EE), “*a physiological consequence closely associated with physical activity*” (Sirard & Pate, 2001, p. 439). Other objective measures of physical activity include heart rate monitors, accelerometers and pedometers. Despite a recent proliferation in ‘second generation’ motion sensors and heart rate monitors (Troost, 2001), self-report continues to be the most feasible method of assessing physical activity in large-scale studies (Kohl, Fulton, & Caspersen, 2000; Ozdoba, Corbin, & LeMasurier, 2004; Sallis & Saelens, 2000).

Using questionnaires, respondents are asked to recall physical activity participation during a period of time (from 24 hr to 1 year) or they may be asked about their usual or habitual physical activity. Physical activity questionnaires are generally inexpensive for the test administrator, convenient for the respondents (Booth, Bauman, Owen, & Gore, 1997). Additionally, they do not alter the behaviour under study, they can

be used with a wide range of ages and adapted to fit the needs of a particular population or set of research questions (Sallis & Saelens, 2000). However, questionnaires have a number of limitations, including: the need for careful design to avoid confusion and misunderstanding (Fowler, 1993), and the difficulty children have in recalling information about their activity patterns (Baranowski et al., 1988). Furthermore, social desirability bias has a potential to lead to over-reporting of physical activity (Kelder, Perry, Klepp, & Lytle, 1994; Warnecke et al., 1997).

Measuring the activity patterns of adolescents is of particular importance because the steepest declines in activity occur during this period (Sallis, 2000). Increasing the proportion of adolescents who meet the physical activity guidelines has been identified as a health priority (Pate et al., 2006). Although there are many existing questionnaires designed to assess child and adolescent physical activity, few assess the four hypothesized dimensions of physical activity (type, frequency, intensity and duration)(Sallis & Saelens, 2000). Furthermore, most existing questionnaires are inappropriate for detecting activity which is unstructured (e.g. recess time), which is a frequent source of activity for youth (Kohl et al., 2000). Specific limitations of adolescent questionnaires include; excessive burden on participants (e.g. Garcia, George, Coviak, Antonakas, & Pender, 1997; e.g. Pate, Ross, Dowda, Trost, & Sirard, 2003), inability of instrument to provide estimates of physical activity behaviour that can be compared to public health guidelines (Aaron et al., 1993; Aaron et al., 1995; Godin & Shephard, 1985), and period of assessment too short to provide a representation of usual behaviour (e.g. Pate et al., 2003; Weston, Petosa, & Pate, 1997). The *Child/Adolescent Activity Log* (CAAL: Garcia et al., 1997) is an assessment tool that satisfies the criteria identified by

Sallis and Saelens (2000) and can be used to assess physical activity time, type, frequency and duration. However, because it is a 1-day or 3-day recall, multiple administrations are required to assess usual physical activity (Telford, Salmon, Jolley, & Crawford, 2004). Furthermore, this method requires students to keep a running log of their daily activities; a time consuming task which relies on the adherence of children. The *Adolescent Physical Activity Recall Questionnaire* (APARQ: Booth, Okely, Chey, & Bauman, 2002) is a promising measure of adolescent physical activity, however, it requires individuals to recall organised and non-organised activity from winter and summer periods, placing considerable burden on participants. More recently Gao and colleagues (Gao et al., 2006) evaluated a modified version of the *Godin-Leisure-Time Exercise Questionnaire*. The authors found the instrument to have poor reliability and validity. The recently evaluated *School Health Action, Planning and Evaluation System* (SHAPES: Wong, Leatherdale, & Manske, 2006) has good reliability and validity and appears to hold promise for the measurement of youth physical activity on a population level. However, respondents are not required to report type of activity (e.g. swimming, soccer etc) or time of day that activity occurs. This information is very important for the development and evaluation of intervention strategies to promote activity among youth.

As there is not an instrument that can effectively quantify the true level and pattern of an adolescent's activity behaviour (Department of Health, 2004), the researchers attempted to develop a reliable simple questionnaire with good content and convergent validity. The Oxford Physical Activity Questionnaire (OPAQ) is a time-based seven-day recall that asks students to report on moderate to vigorous physical activity (MVPA) from the previous week (Appendix A). The questionnaire takes

approximately 15 minutes to complete and was designed to be a self-report measure of physical activity for secondary school students. It shares characteristics with both time-based and activity-based questionnaires. Time-based questionnaires have each day divided into time blocks and respondents are asked to provide the dominant activity for the period chosen from a list of common activities. Activity-based questionnaires are structured around a list of activities with minimal reference to time of day (McMurray et al., 2004). The OPAQ provides respondents with a list of example activities but students are not required to record sedentary activity. The OPAQ was designed specifically for use in the school environment and is presented in a timetable format (similar to ones students often use for their weekly schedule). It has been suggested that memory recall can be enhanced through the use of memory cues, such as a segmented day procedure (e.g., recess, lunch, after school, evening)(Baranowski et al., 1988). These were incorporated into the OPAQ. This allows researchers to determine duration and frequency of activity during different parts of the day and provides the opportunity to compare weekday versus weekend activity. The first items of the OPAQ questionnaire require respondents to report their usual method of transportation to school. It is important for questionnaires to access this information because active transportation has the potential to help students achieve their daily physical activity recommendations but has been largely ignored in population surveys of youth physical activity (Tudor-Locke, Ainsworth, & Popkin, 2001). Students are then asked to list all physical activities and their duration, completed in physical education lessons, school sport and after-school clubs. Following this, students are asked to report on other physical activities played at lunch or other times during the day. To assist in their recall, students are provided with

examples of activities and provided with the following definition: “*Moderate to vigorous physical activity makes you breathe heavily and increases your heart rate*”.

The OPAQ satisfies the criteria for content validity, as identified by Sallis and Saelens (2000). Firstly, the OPAQ provides data that enables researchers to distinguish between time spent in various activities and their metabolic costs. Although the students are provided with examples, the OPAQ is not an activity-based questionnaire and therefore students are not restricted by activities listed by researchers. Secondly, the OPAQ allows for the calculation of total number of minutes spent in MVPA, and using the compendium of physical activities (Ainsworth *et al.*, 2000), enables researchers to obtain estimated metabolic expenditure for the week. These results can be compared to public health guidelines. Thirdly, using the OPAQ, students are able to record information about their physical activity behaviour from various contexts (e.g., before school, at recess and after-school). Finally, the OPAQ is simple and easy to complete. Weisberg, Krosnick and Bowen (1996) have noted that when considering questionnaire design, the proportion of forms returned uncompleted is directly related to the length of the form. As such, the OPAQ was kept short (two pages) to increase completion rates.

Method

Procedure

This paper describes two studies designed to evaluate the test-retest reliability and convergent validity of the Oxford Physical Activity Questionnaire (OPAQ). In the first study the OPAQ was administered to subjects on two occasions separated by one week. In the validity study, students were asked to wear accelerometers for four consecutive school days. Upon the returning the accelerometers students were asked to complete the

OPAQ.

Participants

For the reliability study, a convenience sample of students was recruited from one secondary school in Oxford, United Kingdom (N=87). Students involved in the study were predominantly white and from mixed socio-economic backgrounds. Information leaflets were distributed and informed consent was obtained from students and parents. All students were encouraged to participate in the study regardless of academic ability, ethnicity or state of physical fitness. The study methodology was approved by the University of Oxford's Department of Educational Studies Ethics Committee and informed consent was obtained from parents and students.

The validity study was completed in one independent school in Newcastle, Australia (N = 51). Students in the validity study were also predominately white and from mixed socio-economic backgrounds. The study methodology was approved by the University of Newcastle's Ethics Committee and informed consent was obtained from parents and students.

Measures

Accelerometry. Previous studies have used a variety of objective measures (e.g. doubly labelled water, heart rate monitors, accelerometers and direct observation) to 'validate' physical activity questionnaires. While these measures are often described as providing criterion-related validity, they should not be regarded as a true physical activity criteria (Sallis & Saelens, 2000). Such measures are best viewed as providing convergent evidence toward the construct of physical activity (Patterson, 2000). Due to their reliability, validity and feasibility, *Caltrac* monitoring was used to provide evidence for

the convergent validity of the OPAQ. The *Caltrac* (Muscle Dynamics, Torrance, CA) is an electronic single plane accelerometer that measures vertical acceleration. *Caltrac* accelerometers have been validated against various dimensions of physical activity, in both laboratory and field conditions (Allor & Pivarnik, 2001; Eisenmann et al., 2004; Sallis, Buono, Roby, Carlson, & Nelson, 1990). In an early study Sallis and colleagues (1990) found the *Caltrac* to have strong inter-instrument reliability ($r = .96$) for children aged 8 to 13 years in field conditions. The authors found the accelerometers to be moderately related to heart rate monitoring on day 1 ($r = .54$) and day 2 ($r = .42$). More recently, Eisenmann and colleagues (2004) examined the validity of the *Caltrac* accelerometer in estimating energy expenditure during different activities (sweeping, bowling and basketball). They found moderate to strong correlations between energy expenditure derived from *Caltrac* activity counts and indirect calorimetry.

The *Caltrac* should be attached firmly at the waist, allowing the device to detect movement (Montoye, Kemper, Saris, & Washburn, 1996). When in the 'calories used activity' mode (CALC USED/ACTM) the *Caltrac* output provides a prediction of resting metabolic rate (RMR) and calories used during activity (outcome = estimated calorie expenditure). However, the RMR which accounts for a majority of daily energy expenditure (EE) is predicted on adult equations (Eisenmann et al., 2004). It has been suggested that when using the *Caltrac* with children and adolescents the internal program should be overridden by setting the following values: age = 99, gender = 0, weight = 25, height = 36 (Sallis et al., 1990).

Previous studies have established that four days of consecutive activity monitoring are necessary to provide a reliable measure of habitual physical activity

among children using objective measures (Janz, Witt, & Mahoney, 1995; Trost, Pate, Freedson, Sallis, & Taylor, 2000; Vincent & Pangrazi, 2002). In the present study, *Caltrac* accelerometers were given to students during roll call on Monday morning (time 1) and collected at the same time on Friday morning (time 2). Students were asked to wear the *Caltracs* throughout the day except when swimming or engaging in contact sports. Once programmed to override the internal program, *Caltrac* devices were sealed with tape to discourage student tampering. At time 2, students completed the Oxford Physical Activity Questionnaire. Upon returning the *Caltracs*, students were asked to complete an accelerometer questionnaire indicating if they had worn their pedometer for all four days.

Data management

Using the *Compendium of Physical Activities* (Ainsworth, Haskell, Leon, Jacobs, & Montoye, 1993; Ainsworth et al., 2000), physical activities listed by the students were converted into three weekly totals; moderate physical activity (MPA), vigorous activity (VPA) and total activity (MVPA). Activity intensity was defined in terms of multiples of resting metabolic rate (METs). In adults, 1 MET is approximately equivalent to 3.5ml of $O_2 \cdot kg^{-1} \cdot min^{-1}$, which equates to the energy expenditure for sitting quietly (Ainsworth, Montoye, & Leon, 1994). It has been suggested that the determinants of moderate and vigorous physical activity are different and should be examined separately (Trost et al., 1997). Moderate activity was defined as having a MET value of 3.5-5.9 and vigorous activity was defined as activity requiring greater than or equal to 6 METs of energy expenditure (Booth et al., 2002). Activities reported by students that were less than 3.5 METs were not included in the OPAQ totals (e.g. playing snooker, throwing darts). To

achieve a total OPAQ score, the MET value of each listed activity was multiplied by the number of minutes reported by the students.

Data analysis

The data were examined using the Statistical Package for the Social Sciences (SPSS) software version 12.0. An initial investigation of the data was made using histograms and visual checks. The reliability of the OPAQ measure was assessed using intraclass correlation coefficient (ICC) obtained from a two way analysis of variance (ANOVA) model. The Absolute Agreement model was selected and the average measures ICC is reported in the results. The classification system developed by Landis and Koch (1977) were applied to the findings: $< .40$ represented poor agreement, $.40$ to $.75$ represented fair to good agreement, and values $> .75$ indicated excellent agreement. The statistical significance of the ICC was not reported because in reliability estimation “*we are only interested in the actual value of the coefficient, not whether it is significantly different from zero*” (Patterson, 2000; p. 16). Wilcoxon Signed Rank was also used to examine the reliability of the OPAQ.

As the data were not normally distributed, Spearman’s Rank Order Correlation was used to examine the relationship between self-reported physical activity and accelerometer activity counts. *Caltrac* activity counts were divided by the number of days the student had worn the device to determine an estimate of daily energy expenditure (e.g. if a student forgot to wear the *Caltrac* for 1 day, their score was divided by 3). As *Caltrac* devices only provide a total of activity counts and do not differentiate between light, moderate or vigorous physical activity, total *Caltrac* activity counts were compared with self-reported MPA, VPA and MVPA.

Results

A total of 87 students completed both administrations of the OPAQ (Table 1). The mean age of the students was $13.1 \pm .9$ and there were roughly equal numbers of girls and boys. Table 2 shows the physical activity means and the test-retest reliability of moderate physical activity (MPA), vigorous physical activity (VPA) and moderate to vigorous physical activity (MVPA).

Table 1 Here

From the 1st survey the average number of METs reported by students for MPA, VPA and MVPA were 2177, 1612 and 3807 respectively. The test-retest reliability of the OPAQ to assess total physical activity was excellent (ICC = .91). More specifically, the reliability of recall of vigorous physical activity (ICC = .80) was better than the recall of moderate activity (ICC = .76). The reliability of boys recall was better than that of girls for MPA (ICC = .76 v ICC = .35), VPA (ICC = .80 v ICC = .65) and MVPA (ICC = .90 v ICC = .84). There was a statistically significant difference between scores for MPA and MVPA from the 1st and 2nd surveys, but there was no statistically significant difference between VPA from the two surveys.

Table 2 Here

Fifty one students wore accelerometers and completed the OPAQ in the convergent validity study (mean age $12.6 \pm .50$). The average number of METs reported by students for MPA, VPA and MVPA were 1834, 2034 and 3867 respectively. While VPA and MVPA were related to *Caltrac* accelerometer counts, $r = .33$ ($p = .02$) and $r = .32$ ($p = .02$), moderate activity was not $r = .01$ ($p = .94$).

Table 3 here

Discussion

This study evaluated the test-retest reliability and convergent validity of the OPAQ in two samples of secondary school students. Overall, the test-retest reliability estimates of the OPAQ were excellent ($ICC = .91$). Comparing self-report measures is difficult as there is little consistency in research design, time period assessed or samples used. However, it has been suggested that test-retest reliability coefficients of (.60) and above are acceptable for youth physical activity questionnaires (Sallis & Saelens, 2000). Because behaviour changes over time, reliability coefficients are generally higher when periods of assessment are smaller (e.g. one day recall as opposed to seven-day recall) and time between test and retest shorter. Furthermore, higher coefficients of reliability are generally found in simpler questionnaires which are easier to complete (Shephard, 2003). In the current study the time between test and retest was one week, the time period suggested by Sallis and Saelens (2000). When time between tests was one hour, the *Previous Day Physical Activity Recall* (PDPAR) was found to have reliability coefficients ranging from $r = .98$ to $.99$ (Weston et al., 1997). McMurray and colleagues (2004) found their three day time-based and activity-based questionnaires to have test-retest reliabilities of $r = .68$ and $r = .67$ respectively, when administrations were separated by one day. The test-retest reliability of the *Children's Leisure Activities Study Survey* (CLASS) ranged from .24 to .41 when assessed over a one week period (Telford et al., 2004). When the test-retest reliability of the *Modifiable Activity Questionnaire for Adolescents* was evaluated over a period of one year, the correlation coefficient ranged from .48 to .71 (Aaron et al., 1993).

The ICC reliability scores for vigorous physical activity were higher than those

for moderate activity. This findings corresponds with other reliability studies that have grouped activity by intensity (Fowler-Brown & Kahwati, 2004; Gao et al., 2006; Wong et al., 2006). People generally find it easier to remember more structured forms of physical activity, such as team sports or gym sessions (Fowler-Brown & Kahwati, 2004).

Furthermore, the low validity of moderate activity recall may be due to the poor ability of abstract thinking and memory of youth (Matthews, 2002).

The higher reliability scores for males, was consistent with the findings of some studies (Crocker, Bailey, Faulkner, Kowalski, & McGrath, 1997; Sallis, Buono, Roby, Micale, & Nelson, 1993) but not all (Aaron et al., 1995; Booth et al., 2002; McMurray et al., 2004). It is well established that males are more active than girls at school age and it has been argued that physical activity plays a more influential role in the lives of boys which may influence their cognitive recall of such activities (Sallis et al., 1993). Finally, it is important to note that behavioural changes may weaken the statistical interpretation of the reliability of a physical activity questionnaire (Kohl et al., 2000).

Difficulty in determining the validity of physical activity questionnaires is compounded by the multitude of existing measures and the lack of adequate criteria (Ainsworth et al., 1994). While the doubly labelled water method has been offered as the 'gold standard' of physical activity measurement (Ulf Ekelund et al., 2001), it is expensive and validation studies generally use other objective measures of physical activity (e.g. accelerometers, heart rate monitors, direct observation). Some studies have used cardiorespiratory fitness as a criteria for validation (Booth et al., 2002). However, the relationship between cardiorespiratory fitness (CRF) and habitual physical activity among youth is unclear, because very few children and adolescents achieve the volume

or intensity of activity necessary for the improvement of CRF (Armstrong & McManus, 1994). It has been suggested that the disparate findings relating physical activity to CRF in adolescents are due to inadequate physical activity measurement (U. Ekelund et al., 2001). Consequently, validity correlation coefficients have varied considerably for self-report measures (i.e. from .07 to .88). For the current study, *Caltrac* accelerometers were chosen, due to their reliability, validity and feasibility. Although moderate activity (MPA) was not related to *Caltrac* activity counts, vigorous activity (VPA) and moderate to vigorous activity (MVPA) were both significantly related to accelerometer data, $r = 0.33$ ($p < .05$) and $r = .31$ ($p < .05$), respectively. Compared to other self-report measures of adolescent physical activity these findings are acceptable. Studies comparing data from seven day physical activity self-reports with accelerometer counts have found statistically significant correlation coefficients ranging from $r = .30$ (Sallis et al., 1993) to $r = .44$ (Wong *et al.*, 2006). However, not all studies have found a significant correlation between self-reported activity and accelerometer counts. Time spent in moderate, vigorous and total activity were not related to accelerometer counts in the Children's Leisure Activities Study Survey (CLASS: Telford et al., 2004). Similarly, Gao and colleagues found their modified version of the *Godin-Leisure Time Exercise* Questionnaire to be unrelated to seven days of accelerometer monitoring (Gao et al., 2006).

Conclusion

A strength of the current study is that students wore accelerometers for a period of four days, which corresponded with the activity recall period. Previous validation studies have used a variety of monitoring time periods and monitoring has not always

corresponded with the period of activity recall (e.g. Telford et al., 2004). While it has been suggested that social desirability will lead to students overestimating physical activity (Warnecke et al., 1997), recently, researchers have found that students may underestimate physical activity of moderate intensity (Riddoch et al., 2004; Telford et al., 2004). This could be due to the sporadic nature of children's activity patterns (Welk, Corbin, & Dale, 2000) and the failure to report brief periods of activity (Tudor-Locke & Myers, 2001).

Researchers are yet to develop a seven day physical activity questionnaire that is more than moderately related to a physical activity criterion. This study has demonstrated that the OPAQ is a reliable measure of adolescent physical activity that has similar convergent validity to other seven day measures youth activity. Future research should examine the qualities of the OPAQ with a larger sample more representative of the general population of secondary school students.

Limitations of this study include: the use of adult physical activity compendium values, the possibility of reactivity to the accelerometers and the non-random selection of study participants. The compendium of physical activities (Ainsworth et al., 1993; Ainsworth et al., 2000) used to convert listed physical activities into totals for weekly metabolic expenditure was obtained from studies describing physical activity patterns of adults. These studies measured the energy cost of specific physical activities in field settings. Energy expenditure at both rest and during activity varies by pubertal stage; however, by age 15 for girls and 16 for boys, adult compendium values appear satisfactory (Harrell et al., 2005). Although the energy cost of activities will be similar for adults and young people, differences may increase measurement error. Student

reactivity to the wearing of accelerometers is another possible source of measurement error. Despite this, Ozoba et al. (2004) concluded from their study of sealed and unsealed pedometers that reactivity does not affect measurement with grade four children. The authors argued that if participants cannot see the pedometer feedback, they will be less inclined to change their behaviour. This study involved adolescent students who may be expected to act differently from primary school children. The non-random selection of the classes may have influenced the results of the study.

Acknowledgements

The authors are grateful for the support and participation of the schools, students and staff involved in current study.

References

- Aaron, D. J., Kriska, A. M., Dearwater, S. R., Anderson, R. L., Olsen, T. L., Cauley, J. A., et al. (1993). The epidemiology of leisure physical activity among an adolescent population. *Medicine & Science in Sports & Exercise*, *25*, 847-852.
- Aaron, D. J., Kriska, A. M., Dearwater, S. R., Cauley, J. A., Metz, K. F., & La Porte, R. E. (1995). Reproducibility and validity of an epidemiologic questionnaire to assess past year physical activity in adolescents. *American Journal of Epidemiology*, *142*, 191-201.
- Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs, D. R., Jr., & Montoye, H. J. (1993). Compendium of physical activities: Classification of energy costs of human physical activities. *Medicine & Science in Sports & Exercise*, *25*, 71-80.
- Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., et al. (2000). Compendium of physical activities: An update of activity codes and MET intensities. *Medicine & Science in Sports & Exercise*, *32*, S498-S516.
- Ainsworth, B. E., Montoye, H. J., & Leon, A. S. (1994). Methods of assessing physical activity during leisure and work. In C. Bouchard, R. J. Shephard & T. Stephens (Eds.), *Physical Activity, Fitness and Health: International Proceedings and Consensus Statement* (pp. 145-159). Champaign, Ill: Human Kinetics.
- Allor, K. M., & Pivarnik, J. M. (2001). Stability and convergent validity of three physical activity assessments. *Medicine & Science in Sports & Exercise*, *33*, 671-676.
- Armstrong, N., & McManus, A. (1994). Children's fitness and physical activity: A challenge for physical education. *British Journal of Physical Education*, *25*, 20-

26.

- Baranowski, T., Cullen, K. W., Basen-Engquist, K., Wetter, D. W., Cummings, S., Martineau, D. S., et al. (1988). Validity of self-report of physical activity: An information processing approach. *Research Quarterly for Exercise & Sport*, *59*, 314-327.
- Booth, M. L., Bauman, A., Owen, N., & Gore, C. J. (1997). Physical activity preferences, preferred sources of assistance, and perceived barriers to increased activity among physically inactive Australians. *Preventive Medicine*, *26*, 131-137.
- Booth, M. L., Okely, A. D., Chey, T., & Bauman, A. (2002). The reliability and validity of the Adolescent Physical Activity Recall Questionnaire. *Medicine & Science in Sports & Exercise*, *34*, 1986-1995.
- Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health related research. *Public Health Reports*, *100*, 126-131.
- Crocker, P. R. E., Bailey, D. A., Faulkner, R. A., Kowalski, K. C., & McGrath, R. (1997). Measuring general levels of physical activity: Preliminary evidence for the Physical Activity Questionnaire for Older Children. *Medicine & Science in Sports & Exercise*, *29*, 1344-1349.
- Department of Health. (2004). *At least five a week*. London: Department of Health.
- Eisenmann, J. C., Strath, S. J., Shadrick, D., Rigsby, P., Hirsch, N., & Jacobson, L. (2004). Validity of uniaxial accelerometry during activities of daily living in children. *European Journal of Applied Physiology*, *91*, 259-263.
- Ekelund, U., Poortvliet, E., Nilsson, A., Yngve, A., Holmberg, A., & Sjostrom, M.

- (2001). Physical activity in relation to aerobic fitness in 14- to 15-year-old boys and girls. *European Journal of Applied Physiology*, 85.
- Ekelund, U., Sjostrom, M., Yngve, A., Poortvliet, E., Nilsson, A., Froberg, K., et al. (2001). Physical activity assessed by activity monitor and doubly labeled water in children. *Medicine & Science in Sports & Exercise*, 33, 275-281.
- Fowler-Brown, A., & Kahwati, L. C. (2004). Prevention and treatment of overweight in children and adolescents. *American Family Physician*, 69, 2591-2598.
- Fowler, J. F. J. (1993). *Survey Research Methods* (2nd ed.). Newbury Park, London: Sage.
- Gao, S., Harnack, L., Schmitz, K. H., Fulton, J., Lytle, L., Van Coevering, P., et al. (2006). Reliability and validity of a brief tool to measure children's physical activity. *Journal of Physical Activity & Health*, 3, 415-422.
- Garcia, A. W., George, T. R., Coviak, C., Antonakas, C., & Pender, N. J. (1997). Development of the child/adolescent activity log: A comprehensive and feasible measure of leisure-time physical activity. *International Journal of Behavioral Medicine*, 4, 323-338.
- Godin, G., & Shephard, R. J. (1985). A simple method to assess exercise behaviour in the community. *Canadian Journal of Applied Sport Science*, 10, 141-146.
- Harrell, J. S., McMurray, R. G., Baggett, C. D., Pennell, M. L., Pearce, P. F., & Bangdiwala, S. I. (2005). Energy costs of physical activities in children and adolescents. *Medicine & Science in Sports & Exercise*, 37, 8.
- Janz, K. F., Witt, J., & Mahoney, L. T. (1995). The stability of children's physical activity as measured by accelerometry and self-report. *Medicine & Science in Sports &*

- Exercise.*, 27, 1326-1332.
- Kelder, S. H., Perry, C. L., Klepp, K. K., & Lytle, L. L. (1994). Longitudinal tracking of adolescent smoking, physical activity, and food choice behaviors. *American Journal of Public Health*, 84, 1121-1126.
- Kohl, H., Fulton, J., & Caspersen, C. (2000). Assessment of physical activity among children and adolescents: A review and synthesis. *Preventive Medicine*, 31, S54-S76.
- Kriska, A. M., & Casperson, C. J. (1997). Introduction to a collection of physical activity questionnaires. *Medicine & Science in Sports & Exercise*, 29, S5-S9.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Matthews, C. E. (2002). Use of self-report instruments to assess physical activity. In G. J. Welk (Ed.), *Physical activity assessments for health-related research* (pp. 107-124). Champaign, Ill: Human Kinetics.
- McMurray, R. G., Ring, K. B., Treuth, M. S., Welk, G. J., Pate, R. R., Schmitz, K. H., et al. (2004). Comparison of two approaches to structured physical activity surveys for adolescents. *Medicine & Science in Sports & Exercise*, 36, 2125-2143.
- Montoye, H. J., Kemper, H. C. G., Saris, W. H. M., & Washburn, R. A. (1996). *Measuring Physical Activity and Energy Expenditure*. Champaign, Illinois: Human Kinetics.
- Ozdoba, R. S., Corbin, C. B., & LeMasurier, G. C. (2004). Does reactivity exist in children when measuring activity levels with unsealed pedometers? *Pediatric Exercise Science*, 16, 158-166.

- Pate, R. R., Davis, M. G., Robinson, T. N., Stone, E. J., McKenzie, T. L., Young, J. C., et al. (2006). Promoting physical activity in children and youth: a leadership role for schools: a scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Physical Activity Committee) in collaboration with the Councils on Cardiovascular Disease in the Young and Cardiovascular Nursing. *Circulation*, *114*, 1214-1224.
- Pate, R. R., Ross, R., Dowda, M., Trost, S. G., & Sirard, J. (2003). Validation of a 3-day physical activity recall in female youth. *Pediatric Exercise Science*, *15*, 257-265.
- Patterson, P. (2000). Reliability, validity and methodological response to the assessment of physical activity via self-report. *Research Quarterly for Exercise & Sport*, *71*, 15-20.
- Riddoch, C. J., Bo Anderson, L., Wedderkopp, N., Harro, M., Klasson-Heggebo, L., Sardinha, L. B., et al. (2004). Physical activity levels and patterns of 9- and 15-year-old European children. *Medicine & Science in Sports & Exercise*, *36*, 86-92.
- Sallis, J. F. (2000). Age-related decline in physical activity: A synthesis of human and animal studies. *Medicine & Science in Sports & Exercise*, *32*, 1598-1600.
- Sallis, J. F., Buono, M. J., Roby, J. J., Carlson, D., & Nelson, J. A. (1990). The Caltrac accelerometer as a physical activity monitor for school-aged children. *Medicine & Science in Sports & Exercise*, *22*, 698-703.
- Sallis, J. F., Buono, M. J., Roby, J. J., Micale, F. G., & Nelson, J. A. (1993). Seven-day recall and other physical activity self-reports in children and adolescents. *Medicine & Science in Sports & Exercise*, *25*, 99-108.
- Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report:

- Status, limitations, and future directions. *Research Quarterly for Exercise & Sport*, 71, 1-14.
- Shephard, R. J. (2003). Limits to the measurement of habitual physical activity by questionnaires. *British Journal of Sports Medicine*, 37, 197-206.
- Sirard, J. R., & Pate, R. R. (2001). Physical activity assessment in children and adolescents. *Sports Medicine*, 31, 439-454.
- Telford, A., Salmon, J., Jolley, D., & Crawford, D. (2004). Reliability and validity of physical activity questionnaires for children: The Children's Leisure Activities Study Survey (CLASS). *Pediatric Exercise Science*, 16, 64-78.
- Trost, S., Pate, R., Saunders, R., Ward, D., Dowda, M., & Felton, G. (1997). A prospective study of the determinants of physical activity in rural fifth-grade children. *Preventive Medicine*, 26, 257-263.
- Trost, S. G. (2001). Objective measurement of physical activity in youth: Current Issues, future directions - An examination of methodological issues related to the use of physical activity monitors in youth. *Exercise & Sport Sciences Reviews*, 29, 32-36.
- Trost, S. G., Pate, R. R., Freedson, P. S., Sallis, J. F., & Taylor, W. C. (2000). Using objective physical activity measures with youth: How many days of monitoring are needed? *Medicine & Science in Sports & Exercise*, 32, 426-431.
- Tudor-Locke, C., Ainsworth, B. E., & Popkin, B. M. (2001). Active commuting to school: An overlooked source of children's physical activity? *Sports Medicine*, 31, 309-313.
- Tudor-Locke, C. E., & Myers, A. M. (2001). Challenges and opportunities for measuring

- physical activity in sedentary adults. *Sports Medicine*, 31, 91-100.
- Vincent, S. D., & Pangrazi, R. P. (2002). Does reactivity exist in children when measuring activity levels with pedometers? *Pediatric Exercise Science*, 14, 56-63.
- Warnecke, R. B., Johnson, T. P., Chavez, N., Sudman, S., O' Rourke, D. P., Lacey, L., et al. (1997). Improving question wording in surveys of culturally diverse populations. *Annals of Epidemiology*, 7, 334-342.
- Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research and data analysis*. London: Thousand Oaks, Sage.
- Welk, G. J., Corbin, C. B., & Dale, D. (2000). Measurement issues in the assessment of physical activity in children. *Research Quarterly for Exercise & Sport*, 71, 59-73.
- Weston, A. T., Petosa, R., & Pate, R. R. (1997). Validation of an instrument for measurement of physical activity in youth. *Medicine & Science in Sports & Exercise*, 29, 138-143.
- Wong, S. L., Leatherdale, S. T., & Manske, S. R. (2006). Reliability and validity of a school-based physical activity questionnaire. *Medicine & Science in Sports & Exercise*, 38, 1593-1600.

Tables

Table 1: Demographic characteristics of participants included in the reliability (N = 87) and validity ($n = 51$) studies

		Reliability sample	Validity sample
		%(n)	%(n)
Age	≤ 12	33.3 (29)	41.2 (21)
	13	23 (20)	58.8 (30)
	≥ 14	43.7 (38)	—
Sex	Female	44.8 (39)	47.1 (24)
	Male	55.2 (48)	52.9 (27)

— denotes no value

Table 2: Test-retest reliability of the OPAQ to assess physical activity

	Mean (SD)		<i>p</i> -value ³	ICC (95% CI)		
	S1 ¹	S2 ²		Boys	Girls	Overall
Moderate activity	2177 (1592)	1827 (1401)	.009	.76 (.57 - .87)	.35 (-.23 - .66)	.76 (.63 - .84)
Vigorous activity	1612 (1358)	1555 (1407)	.301	.80 (.65 - .89)	.65 (.34 - .82)	.80 (.70 - .87)
Moderate to vigorous activity	3807 (2705)	3317 (2463)	.004	.90 (.84 - .95)	.84 (.70 - .92)	.91 (.87 - .95)

¹ 1st survey administration

² 2nd survey administration

³ Wilcoxon signed rank test for mean difference between the 2 surveys

Table 3: Mean (METs/ week) and correlation of self-reported physical activities with Caltrac accelerometer counts

	METs/week ¹	SROC ²	
	Mean (SD)	<i>r</i>	<i>p</i> -value
Moderate activity	1834 (2007)	.01	.94
Vigorous activity	2034 (1754)	.33	.01
Moderate to vigorous activity	3868 (2636)	.32	.02

¹ Means are reported as METs per week

² SROC is Spearman Rank Order Correlation

Appendix A