Research Quarterly for Exercise and Sport
©2000 by the American Alliance for Health,
Physical Education, Recreation and Dance
Vol. 71, No. 2, pp. 1-14

Assessment of Physical Activity by Self-Report: Status, Limitations, and Future Directions

James F. Sallis and Brian E. Saelens

Keywords: questionnaire, reliability, validity, prevalence, measurement

Self-report instruments continue to be the most widely used type of physical activity measure. Thus, it is important to identify the strengths and limitations of these measures as well as to continually strive to improve their quality. As more alternatives to self-reports are developed and come into wider use, the role self-reports plays in the mix of measurement options needs to be reconsidered.

In the present review, self-reports are defined as self-administered or interviewer-administered recall questionnaires, activity logs or diaries, and proxy reports (typically used to assess young children). The recognized benefits of self-report (and proxy report) measures are their ability to collect data from a large number of people at low cost. Recalls do not alter the behavior under study, and it is possible to assess all the dimensions of physical activity so patterns of behavior can be examined. Self-reports have been used in a range of ages, and measures can be adapted to fit the needs of a particular population or research question.

Numerous limitations of self-reports have been discussed (Ainsworth, Montoye, & Leon, 1994). Social desirability bias can lead to over-reporting of physical activity (Warnecke, Johnson, Chavez, Sudman, O'Rourke, Lacey, & Horm, 1997). Recalling physical activity is a highly complex cognitive task (Baranowski, 1988), and instruments vary in their cognitive demands. Children and very old adults are likely to have particular memory and recall skill limitations. Respondents and investigators must share understanding of ambiguous terms such as "physical activity," "moderate intensity," and "leisure time." Activity logs and diaries are limited by response rates and the extent to which participants follow instructions. Proxy reports are limited by the reporter's opportunity to observe the physical activity of the subject. Measures may not assess the primary modes of activity for certain gender, age, cultural, occupational, or income groups.

Because there are several published reviews of physical activity self-reports (Ainsworth, Montoye, & Leon, 1994; Kriska & Caspersen, 1997; Montoye, Kemper, Saris, & Washburn, 1996; Sallis, 1991; Washburn & Montoye, 1986), the present review is not intended to be comprehensive. The goal is to summarize findings on test-retest reliability, criterion-related validity, and content validity for physical activity self-reports that were developed in, or have been used in, the 1990s. Content validity has not been widely discussed related to physical activity measures, but it is critical to examine the extent to which an instrument measures the full range of physical activity characteristics.

Method

Self-report measures of physical activity were included for review if (a) there were published reports of validity testing of the measure using an objective assessment of physical activity, (b) the measure has been used in published studies during the 1990s, and (c) the time frame of recall was within the previous one year. Objective criterion-related validity measures that were considered to directly assess physical activity were doubly labeled water, accelerometers (and other motion sensors), direct observation, and heart rate monitoring. Indirect criterion measures, including physical fitness and body

James F. Sallis, Ph.D. and Brian E. Saelens, Ph.D. are with the Department of Psychology at San Diego State University.

composition, or diaries, were excluded, but these results can be found in previous reviews (Ainsworth et al., 1994; Kriska & Caspersen, 1997; Montoye, et al., 1996; Sallis, 1991; Washburn & Montoye, 1986).

Self-report measures meeting inclusion criteria are presented in Table 1 for children and adolescents (up to about age 18 years), Table 2 for young and middle-aged adults, and Table 3 for older adults (generally aged 60 years and over). Each entry includes the name of the measure and primary citation. The assessment format, administration time (if available), and measurement time frame are provided. Content validity is indicated by the (a) dimensions along which physical activity is assessed (i.e., type, frequency, intensity, duration), (b) ability of the instrument to produce estimates related to public health guidelines (i.e., sedentary, moderate, vigorous activities), and (c) ability to obtain scores for various contexts or purposes of physical activity (i.e., leisure, work, household, transportation). Test-retest reliability is presented as a mean of correlations across studies (where applicable) weighted by sample size. The validity criteria are listed along with validity correlations expressed as weighted means according to the sample size. Information presented in the tables was consolidated from (a) the 1997 *Medicine and Science in Sports and Exercise* supplement on physical activity questionnaires (Kriska & Caspersen, 1997), (b) the Montoye, Kemper, Saris, and Washburn (1996) text on physical activity measurement, and (c) articles from 1990 to 1999 obtained from Medline and PsychINFO database searches. Specific references used to compute each reliability and criterion validity mean are available from the first author. Database searches were conducted by entering as a keyword the name of each questionnaire and also by the keyword search of "physical activity" and "self-report" or "interview".

Results

Measures for Youth

Seventeen instruments for assessing child and adolescent physical activity were reviewed. Seven measures used interview only, or in combination with diary or self-administered survey. There were eight self-administered measures and two proxy reports. Self-administered and interview-based recalls were tested only on children aged 9 years and above. Recall time frames varied from one day to three months. One measure (Garcia et al., 1997) assessed all four dimensions of physical activity, six measures assessed three of the dimensions, and three did not allow assessment of any of the dimensions. Most of the youth measures included separate scoring for moderate and vigorous physical activity. Intervals between retests varied considerably, and some reliability correlations reflected both measurement error and true behavioral variation because recall time frames did not overlap. However, all reliabilities were acceptable (range=.60 to .98). All measures showed some evidence of validity. However, almost none of the youth measures were validated in more than one study. Although some of the validity correlations were nonsignificant or low (i.e., less than .20), 13 of the 27 correlations were .50 or above. Validity correlations ranged from .07 to .88 for self-administered surveys, from .17 to .72 for interviews, and from .40 to .77 for the two proxy reports.

Measures for Adults

Seven instruments were reviewed that were validated with young- to middle-aged adults. As shown in Table 2, the measures for adults had relatively high test-retest reliabilities, especially when intervals between tests were one month or less. Many reliability correlations reflected both measurement error and true variation in physical activity. Total and vigorous physical activity scores tended to have higher reliabilities than moderate physical activity scores.

Most measures for adults assessed the dimensions of frequency and intensity, but less commonly assessed type and duration. Moderate and vigorous physical activity scores were obtainable from most adult measures, but the assessment of sedentary behavior was rare. The assessment of "traditional" leisure time physical activities (e.g., biking, sports) was ubiquitous, but less common were assessments of activities in household and transportation contexts. Measurement of physical activity during work was also less common, perhaps due to the existence of specialized measures (e.g., Reiff et al., 1967).

Validity correlations for summary measures of adults' habitual or global physical activity were generally low, ranging from .14 to .36. However, possibly due to their short and specific recall timeframe, the Seven-Day Physical Activity Recall (PAR) had validity correlations of .50 and .53 with accelerometers, and the Modifiable Activity Questionnaire (MAQ) had high correlations with Caltrac accelerometer and doubly labeled water. Subscales assessing vigorous physical activity had higher validity correlations than those summarizing moderate physical activity.

Measures for Older Adults

Of the four measures developed for older adults, two were self-administered and two were interviews. Three of the instruments assessed all four dimensions of physical activity. All allowed scoring of moderate and vigorous physical activity, and coverage of activity contexts was generally good. Recall time frames were usual, typical

week, past seven days, and past year. One measure (DiPietro et al., 1993) had a reliability less than .60, but the rest were .75 and above. Correlations with doubly labeled water were high (r=.58 and .62), but correlations with Caltrac accelerometers were much lower. Only the Modified Baecke Questionnaire was validated in more than one study.

Discussion

There is a profusion of self- or proxy-report measures of physical activity for young people, adults, and older adults, most of which have some evidence of validity. Thus, investigators can often choose an existing measure that is appropriate for particular studies. The disadvantage of multiple available measures is that it is difficult or impossible to compare results across studies. There may be benefits of coming to a consensus on a small number of well-studied measures for each age group that can be used for a variety of research purposes.

Self-report instruments appear to be at different stages of development for the three age groups. There is substantial interest in physical activity measurement for young people, as indicated by the existence of 17 validated measures. However, the relatively large literature is merely producing additional measures rather than systematically exploring principles for measuring youth physical activity. For example, few studies have assessed the performance of measures in various age groups of young people (Crocker et al., 1997; Kowalski, Crocker, & Faulkner, 1997; Kowalski, Crocker, & Kowalski, 1997; Sallis, Buono et al., 1993). Validity appeared to be stronger for interview measures, because about 2/3 of validity correlations for interviews were above .50, compared to 1/3 of validity correlations for self-administered instruments. Although no validated parent *recalls* could be located, the two proxy reports that relied on daily logging showed consistent evidence of validity. Knowledge in the field is most likely to be advanced by studies that compare different measures (Janz et al., 1995; Sallis, Condon, et al., 1993) or different formats of similar measures (Sallis, Condon et al., 1993; Sallis et al., 1996), because such studies could lead to empirically-based principles for survey design.

The literature on adult self-report measures showed signs of greater maturity. Despite the development of many measures over several decades (Kriska & Caspersen, 1997; Montoye et al., 1996), a consensus seems to be developing based on the recent use of a smaller number of measures. Only seven validated measures have been used with some frequency in the past decade, and all but one of the measures was developed before the 1990s. In contrast to the youth literature, most of the adult measures have been evaluated in multiple studies. This literature has benefited from key studies that evaluated multiple measures simultaneously (Jacobs, Ainsworth, Hartman, & Leon, 1993).

The literature on physical activity self-reports for older adults is the newest and most limited. Four validated measures were identified, and three of them were developed or modified during the 1990s. All measures were specifically developed to be sensitive to physical activities common among older adults, and three of the four showed good evidence of reliability and validity.

It was notable that none of the measures used to collect U.S. national data on prevalence of physical activity for young people and adults (USDHHS, 1996) met the criteria for inclusion in this review, because they have not been validated. Although problems with U.S. national prevalence estimates were identified years ago (Slater, Green, Vernon, & Keith, 1987), measures remain unvalidated.

Requirements for Reliability Studies

Reliability is hampered by imprecise cognitive processing, memory errors (Baranowski, 1988), variability in levels of physical activity over time (Washburn & Montoye, 1986), and individual differences that predict recall biases (Buchowski, Townsend, Chen, Acra, & Sun, 1999; Jakicic, Polley, & Wing, 1998). A more thorough examination of the cognitive processes involved in recall is needed, with subsequent integration into recall formats and procedures that improve recall accuracy (Durante & Ainsworth, 1996).

Many test-retest reliability estimates for recalled physical activity reflect both measurement error and true variability in the behavior. Because types and amounts of physical activity are highly variable, retests need to cover the same time period as the initial test. Thus, one-day recalls must be repeated on the same day, and one week recalls should be repeated within one week. Recalls of "usual" physical activity should be less sensitive to the interval between tests. Booth, Owen, Bauman, and Gore (1996) argue that inclusion of many participants with no physical activity in analyses inflates reliability indices.

The relatively low reliability of recalls of moderate intensity activities, as compared to vigorous intensity activities, may be best explained by the low salience of moderate intensity activities (Durante & Ainsworth, 1996). It is unclear whether this problem can be overcome by improved recall procedures.

Content Validity

The importance of assessing various intensity levels has been increased by mounting evidence that sedentary behaviors (Andersen, Crespo, Bartlett, Cheskin, & Pratt,

1998; Eaton, Lapane, Garber, et al., 1995), moderate intensities and amounts of physical activity (Pate et al., 1995; USDHHS, 1996), and vigorous intensity activities (USDHHS, 1996) are related to various health outcomes. A content valid measure assesses all these currently recognized intensities.

Assessing other dimensions of physical activity may be essential for specialized studies. For example, no validated measures of muscle strengthening or flexibility exercises were identified. Weight-bearing characteristics of physical activity are important in studies of osteoporosis risk. Injury researchers may need assessments of contact sports.

Assessing the contexts within which people engage in physical activity is an important part of content validity. The lack of consideration of the context can lead to underestimates of prevalence. For example, Evenson and colleagues (1998) documented the underestimate of physical activity among African American women and men that occurred by relying only on measures of leisure-time physical activity. These populations were obtaining substantial physical activity in occupational settings, and women were frequently engaged in physically active household work.

Content validity will be impacted by ongoing changes in work and leisure environments. The introduction of new technologies has created new sedentary behavior options (e.g., internet use) while expanding opportunities for new types of physical activities (e.g., rollerblading, snowboarding).

Validation Criteria

Evidence of validity is limited by the quality of the self-report and the criterion measure. The present review included only validity data that were based on objective measures that directly measured movement or physical activity-related energy expenditure. Activity diaries were not considered direct validation criteria for this review because they are alternative self-reports and have not been well validated. Physical fitness measures are commonly used as validation criteria, but cardiorespiratory fitness is weakly related to physical activity, particularly in young people (Morrow & Freedson, 1994) and older adults (Tager, Hollenberg, & Satariano, 1998). However, limitations of all the direct criterion measures have been identified (Ainsworth et al., 1994; Montoye et al., 1996).

Research is needed to validate self-reports of the critical physical activity dimensions of type, frequency, intensity, and duration. Studies of youth and older adult self-report measures seldom reported attempts to validate these dimensions. However, most of the adult studies validated reports of moderate and vigorous intensity activities. Current accelerometers allow each minute of monitoring to be classified by intensity (Freedson, Melanson, & Sirard, 1998; Nichols, Morgan, Sarkin, Sallis, & Calfas, 1999), so future studies will be able to validate self-reported frequency, intensity, and duration.

Relative Vs. Absolute Validity

Self-reports with evidence of relative validity, shown by correlations with criteria, are useful in studies of associations with health outcomes and to assess outcomes in controlled intervention studies. However, absolute amounts of physical activity need to be assessed for epidemiologic research to define the dose-response associations between physical activity and health outcomes (Haskell, 1994) and for public health monitoring of the proportion of the population meeting recommendations for physical activity. Despite the importance of knowing the absolute amount of physical activity people are obtaining, virtually all of the validity data presented in the reviewed studies addressed only relative validity, expressed as correlations. We propose that assessment of absolute validity must become a major focus of validation efforts.

A few studies evaluated absolute validity, using heart rate monitors or accelerometers as criterion measures. One study of young people indicated that reports of vigorous physical activity were relatively accurate (Sallis, Buono et al., 1993). However, most studies indicated gross overestimation of physical activity on self-reports. Among fifth-grade students, the average overestimate of moderate to vigorous physical activity was 29 minutes with an interview format and 48 minutes with a self-administered format. Several of the overestimates were by more than 100% (Sallis et al., 1996). Adolescents over-reported days per week of doing 20 minutes of vigorous physical activity by 3.8 days, compared to accelerometer monitoring. However, reports for number of days with at least 60 minutes of moderate physical activity were more accurate (Prochaska, Sallis, Long, Lydston, & Wade, 1999).

Although two studies of adults showed similar amounts of physical activity assessed by self-report and objective measures (Bernstein, Sloutskis, Kumanyika, Sparti, Schutz, & Morabia, 1998; Racette, Schoeller, & Kushner, 1995), most indicated substantial over-estimation of physical activity. A short-term study found self-reports overestimated observed physical activity by 300% (Klesges, Eck, Mellon, Fulliton, Somes, & Hanson, 1990). Another study found that although reported and accelerometer-monitored total minutes of physical activity were almost identical, recalled vigorous physical activity was overestimated by 100%. It was notable that moderate intensity activities were *underestimated* by 35% (Hayden, Sallis, Armstrong, Whalen, & Sarkin, 1998). Physical activity was overestimated by 40% to 60% in women using daily records, showing that activity diaries are also subject to bias (Jakicic, Polley, & Wing, 1998). The most definitive study to date was conducted with 115 women and men who spent two days in a whole room indirect calorimeter (Buchowski et al., 1999). On exer-

cising days, participants over-reported high-moderate intensity activities (>4.5 METs) by about 90%, but low-moderate activities (2.5-4.4 METs) were reported accurately. Amount of over-reporting was positively correlated with percent body fat.

The weight of the current evidence supports a conclusion that young people and adults overestimate their physical activity, particularly vigorous intensity activities. These results suggest that prevalence estimates may be inaccurate, even at the national level, because prevalence data are virtually always based on self-reports (e.g., USDHHS, 1996). Self-reports may not be acceptable measures when *absolute* amount of physical activity needs to be assessed. A high research priority is to determine whether self-reports can be improved to yield more accurate estimates of absolute amount of physical activity.

Developing Self-Report Measures

Few studies cited in the present review described the rationale for the design of the self-report measure or the process of instrument development. Thus, the systematic procedures reported by a few investigators are instructive (Bernstein et al., 1998; Sallis et al., 1996; Washburn et al., 1993). Content validity was enhanced through consultation with experts, interviews with potential respondents, and statistical methods to identify types of activities to include. Extensive pilot testing of draft instruments was carried out, and revisions were made before formal psychometric evaluation. Different administration formats were evaluated. When new measures are developed, the rationale and procedures used to develop the instrument should be described.

More than 10 years ago Baranowski (1988) described the cognitive complexity of recalling physical activity and proposed a research agenda that would lead to improved measures. However, little if any progress has been made in conducting the proposed research, so the empirical basis for improving the accuracy of recalls is still absent.

Adaptation and Evaluation in Specific Demographic and Cultural Groups

There is a growing interest in examining the level of physical activity among diverse demographic, cultural, and ethnic populations (e.g., Sternfeld, Ainsworth, & Quesenberry, 1999). Investigators should be sensitive to such factors as differences in the range of physical activity performed within and between different populations (Ainsworth, Irwin, Addy, Whitt, Stolarczyk, 1999). Further studies are needed to understand cultural, demographic, and ethnic differences in reactivity to assessment methodologies (Warnecke, Johnson, Chavez, Sudman, O'Rourke, Lacey, & Horm, 1997).

Although some measures appear to perform adequately in specific cultural groups (Rauh, Hovell, Hofstetter, Sallis, & Gleghorn, 1992), there is considerable variability in the response of ethnically distinct groups to various measurement approaches. Ainsworth et al. (1999) found higher prevalence rates for meeting moderate physical activity guidelines among Native American and African American women when using 4-day activity diaries than when more structured assessments were used. Young, Miller, Wilder, Yanek, and Becker (1998) obtained higher prevalence rates among urban African Americans when including assessments of physical activity for transportation and work. Ainsworth et al. (1999) suggested that discrepancies were explained by self-report measures lacking physical activity content validity for a given population, so steps should be taken to ensure content validity of measures in target populations.

Warnecke et al. (1997) suggest the need to be aware of the differential impact of question wording on different populations. In an effort to develop a standardized approach to physical activity self-reports that can be used internationally, a group is developing and evaluating a format that is flexible enough to be tailored for individual countries yet still provides relatively comparable data (Pratt, Ainsworth, Booth, & Craig, 1999).

Conclusions

There are multiple self-report physical activity measures with adequate reliability, content validity, and relative criterion validity that can be used with youth, adults, and older adults. New measures should be developed only if they improve substantially on existing measures.

Among adults, relative validity was higher for reports of vigorous physical activity than for reports of moderate intensity activities.

Validated self-reports were not located that assessed strengthening exercises, flexibility exercises, weight-bearing activities, sedentary behavior, or other activity dimensions that are proposed to be related to health outcomes.

In general, interview measures had stronger psychometric characteristics than self-administered measures.

The literature provided few empirically-based principles for designing physical activity self-reports or selecting recall formats and strategies.

Most studies showed self-reports do not provide accurate estimates of the absolute amount of physical activity, so when absolute amounts of physical activity need to be estimated, objective measures should be used.

Few self-report measures have been developed for, or validated in, distinct demographic, ethnic, or cultural groups.

When used in combination with objective measures, the unique role of self-reports may be to assess the context and type of physical activities.

References

Ainsworth, B.E., Irwin, M.L., Addy, C.L., Whitt, M.C., & Stolarczyk, L.M. (1999). Moderate physical activity patterns of minority women: the Cross-Cultural Activity Participation Study. *Journal of Women's Health and Gender-Based Medicine, 8,* 805-813.

Ainsworth, B.E., Montoye, H.J., & Leon, A.S. (1994). *Methods of assessing physical activity during leisure and work.* pp. 146-159. In C. Bouchard, R.J. Shephard, & T. Stephens (Eds.), Physical activity, fitness, and health: International proceedings and consensus statement. Champaign, IL: Human Kinetics.

Andersen, R.E., Crespo, C.J., Bartlett, S.J., Cheskin, L.J., & Pratt, M. (1998). Relationship of physical activity and television watching with body weight and level of fatness among children: Results from the third National Health and Nutrition Examination Survey. *Journal of the American Medical Association, 279,* 938-942.

Baranowski, T. (1988). Validity and reliability of self-report of physical activity: An information processing perspective. *Research Quarterly, 59,* 314-327.

Bernstein, M., Sloutskis, D., Kumanyika, S., Sparti, A., Schutz, Y., & Morabia, A. (1998). Data-based approach for developing a physical activity frequency questionnaire. *American Journal of Epidemiology, 147,* 147-154.

Booth, M.L., Owen, N., Bauman, A.E., & Gore, C.J. (1996). Retest reliability of recall measures of leisure-time physical activity in Australian adults. *International Journal of Epidemiology, 25,* 153-159.

Buchowski, M.S., Townsend, K.M., Chen, K.Y., Acra, S.A., & Sun, M. (1999). Energy expenditure determined by self-reported physical activity is related to body fatness. *Obesity Research, 7,* 23-33.

Cale, L. (1994). Self-report measures of children's physical activity: Recommendations for future development and a new alternative measure. *Health Education Journal, 53,* 439-453.

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 and Science in Sports and Exercise, 29,* 1344-1349.

Durante, R. & Ainsworth, B.E. (1996). The recall of physical activity: Using a cognitive model of the question-answering process. *Medicine and Science in Sports and Exercise, 28,* 1282-1291.

Eaton, C.B., Lapane, K.L., Garber, C.A., Assaf, A.R., Lasater, T.M., & Carleton, R.A. (1995). Sedentary lifestyle and risk of coronary heart disease in women. *Medicine and Science in Sports and Exercise, 27,* 1535-1539.

Evenson, K.R., Rosamond, W.D., Cai, J., Brancati, F.L., & Ainsworth, B.E. (1998). Occupational activity and its relationship to leisure-time physical activity: The Atherosclerosis Risk in Communities Study. Poster presented at the 45th Annual Meeting of the American College of Sports Medicine. Orlando, FL.

Freedson, P.S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise, 30,* 777-781.

Garcia, A.W., George, T.R., Coviak, C., Antonakos, 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.

Gortmaker, S.L., Must, A., Sobol, A.M., Peterson, K., Colditz, G.A., & Dietz, W.H. (1996). Television viewing as a cause of increasing obesity among children in the United States, 1986-1990. *Archives of Pediatrics and Adolescent Medicine, 150,* 356-362.

Harro, M. (1997). Validation of a questionnaire to assess physical activity of children ages 4-8 years. *Research Quarterly for Exercise and Sport, 68,* 259-268.

Haskell, W.L. (1994). Health consequences of physical activity: Understanding and challenges regarding dose-response. *Medicine and Science in Sports and Exercise, 26* , 649-660.

Hayden, H.A., Sallis, J.F., Armstrong, C., Whalen, K., & Sarkin, J. (1998). Evaluation of the phone version of the seven day physical activity recall. *Medicine and Science in Sports and Exercise, 30,* S14.

Jacobs, D.R., Ainsworth, B.E., Hartman, T.J., & Leon, A.S. (1993). A simultaneous evaluation of ten commonly used physical activity questionnaires. *Medicine and Science in Sports and Exercise, 25,* 81-91.

Jakicic, J.M., Polley, B.A., & Wing, R.R. (1998). Accuracy of self-reported exercise and the relationship with weight loss in overweight women. *Medicine and Science in Sports and Exercise, 30,* 634-638.

Janz, K.F., Witt, J., & Mahoney, L.T. (1995). The stability of children's physical activity as measured by accelerometry and self-report. *Medicine and Science in Sports and Exercise, 27,* 1326-1332.

Klesges, R.C., Eck, L.H., Mellon, M.W., Fulliton, W., Somes, G.W., & Hanson, C.L. (1990). The accuracy of self-reports of physical activity. *Medicine and Science in Sports and Exercise, 22,* 690-697.

Kowalski, K.C., Crocker, P.R.E., & Faulkner, R.A. (1997). Validation of the physical activity questionnaire for older children. *Pediatric Exercise Science, 9,* 174-186.

Kowalski, D.C., Crocker, P.R.E., & Kowalski, N.P. (1997). Convergent validity of the physical activity questionnaire for adolescents. *Pediatric Exercise Science, 9,* 342-352.

Kriska, A.M., & Caspersen, C.J. (Eds.) (1997). A collection of physical activity questionnaires for health-related research. *Medicine and Science in Sports and Exericse, 29,* S1-S205.

Manios, Y., Kafatos, A., & Markakis, G. (1998). Physical activity of 6-year-old children: Validation of two proxy reports. *Pediatric Exercise Science, 10,* 176-188.

Montoye, H.J, Kemper, H.C.G., Saris, W.H.M., & Washburn, R.A. (1996). Measuring physical activity and energy expenditure. Champaign, IL: Human Kinetics.

Morrow, J.R., & Freedson, P.S. (1994). Relationship between habitual physical activity and aerobic fitness in adolescents. *Pediatric Exercise Science, 6,* 315-329.

Nichols, J.F., Morgan, C.G., Sarkin, J.A., Sallis, J.F., and Calfas, K.J. (1999). Validity, reliability, and calibration of the Tritrac accelerometer as a measure of physical activity. *Medicine and Science in Sports and Exercise, 31,* 908-912.

Pate, R.R., Pratt, M., Blair, S.N. et al. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of the American Medical Association, 273,* 402-407.

Powell, K.E. & Blair, S.N. (1994). The public health burdens of sedentary living habits: Theoretical but realistic estimates. *Medicine and Science in Sports and Exercise, 26,* 851-856.

Pratt, M., Ainsworth, B., Booth, M., & Craig, C.L. (1999). Can public health surveillance of physical activity be standardized internationally? *Medicine and Science in Sports and Exericse, 31,* S389.

Prochaska, J.J., Sallis, J.F., Long, B., Lydston, D., & Wade, D. (1999). Validity and reliability of physical activity screening tools for adolescents in primary care settings. *Medicine and Science in Sports and Exercise, 31,* S288.

Racette, S.B., Schoeller, D.A., & Kushner, R.F. (1995). Comparison of heart rate and physical activity recall with doubly labeled water in obese women. *Medicine and Science in Sports and Exercise, 27,* 126-133.

Rauh, M.J., Hovell, M.F., Hofstetter, C.R., Sallis, J.F., & Gleghorn, A. (1992). Reliability and validity of self-reported physical activity in Latinos. *International Journal of Epidemiology, 21,* 966-971.

Reiff, G.G., Montoye, H.J., Remington, R.D., Napier, J.A., Metzner, H.L., & Epstein, F.H. (1967). Assessment of physical activity by questionnaire. *Journal of Sports Medicine and Physical Fitness, 7,* 135-142.

Sallis, J.F. (1991). Self-report measures of children's physical activity. *Journal of School Health, 61,* 215-219.

Sallis, J.F., Buono, M.J., Roby, J.J., Micale, F.G., and Nelson, J.A. (1993). Seven-day recall and other physical activity self-reports in children and adolescents. *Medicine and Science in Sports and Exercise, 25,* 99-108.

Sallis, J.F., Condon, S.A., Goggin, K.J., Roby, J.J., Kolody, B., and Alcaraz, J.E. (1993). The development of self-administered physical activity surveys for 4th grade students. *Research Quarterly for Exercise and Sport, 64,* 25-31.

Sallis, J.F., Strikmiller, P.K., Harsha, D.W., Feldman, H.A., Ehlinger, S., Stone, E.J., Williston, B.J., and Woods, S. (1996). Validation of interviewer- and self-administered physical activity checklists for fifth grade students. *Medicine and Science in Sports and Exercise, 28,* 840-851.

Simons-Morton, B.G., Taylor, W.C., & Huang, I.W. (1994). Validity of the physical activity interview and Caltrac with preadolescent children. *Research Quarterly for Exercise and Sport, 65,* 84-88.

Slater, C.H., Green, L.W., Vernon, S.W., & Keith, V.M. (1987). Problems in estimating the prevalence of physical activity from national surveys. *Preventive Medicine, 16,* 107-118.

Sternfeld, B., Ainsworth, B.E., & Quesenberry, C.P. (1999). Physical activity patterns in a diverse population of women. *Preventive Medicine, 28,* 313-323.

Tager, I.B., Hollenberg, M., & Satariano, W.A. (1998). Association between self-reported leisure-time physical activity and measures of cardiorespiratory fitness in an elderly population. *American Journal of Epidemiology, 147,* 921-931.

Warnecke, R.B., Johnson, T.P., Chavez, N., Sudman, S., O'Rourke, D.P., Lacey, L., & Horm, J. (1997). Improving question wording in surveys of culturally diverse populations. *Annals of Epidemiology, 7,* 334-342.

Washburn, R.A., & Montoye, H.J. (1986). The assessment of physical activity by questionnaire. *American Journal of Epidemiology, 123,* 563-575.

Washburn R.A., Smith, K.W., Jette, A.M., & Janney, C.A. (1993). The Physical Activity Scale for the Elderly (PASE): Development and evaluation. *Journal of Clinical Epidemiology, 46,* 153-162.

U.S. Department of Health and Human Services. (1991). *Healthy People 2000: National health promotion and disease prevention objectives.* DHHS Pub. no. (PHS) 91-50212. Washington, DC: U.S.Government Printing Office.

U.S. Department of Health and Human Services. (1996). Physical activity and health: A report of the Surgeon General. Atlanta, GA: Centers for Disease Control.

Verschuur, R., & Kemper, H.C.G. (1987). Longitudinal changes in daily physical activity in girls and boys between age 12 and 18. In R. Verschuur (Ed.), *Daily physical activity and health* (pp. 61-94). Haarlem, The Netherlands: B.V. Uitgeverig de Vrieseboch.

Weston, A.T., Petosa, R., & Pate, R.R. (1997). Validation of an instrument for measurement of physical activity in youth. *Medicine and Science in Sports and Exercise, 29,* 138-143.

Young, D.R., Miller, K.W., Wilder, L.B., Yanek, L.R., & Becker, D.M. (1998). Physical activity patterns of urban African Americans. *Journal of Community Health, 23,* 99-112.

Table 1. Child and Adolescent Self-Report Measures of Physical Activity

PA Dimensions _____ Possible scores _____

Measure & Original Ref.	Assessment Format	Type	Frequency	Intensity	Duration	Sedentary	Moderate	Vigorous	Leisure	Work	Household	Transport	Re-test reliability, Weighted mean	Validity criterion	Validity, weighted mean
Seven-day P.A. Recall (Sallis, Buono, et al., 1993)	Interview; 15-20 min; 7 days		x	x	x		x	x	x	x			2-6 days .77	Heart rate (vigorous) observation	53; agreement on type = 46%; agreement on intensity = 75% (Wallace, 85)
Amsterdam Growth Study Q. (Verschuur & Kemper, 1987)	Combination of self-admin. & interview; 20 min; 3 months			x	x		x	x	x	x	x	x		Heart rate; pedometer (Montoye)	.17 .17
Weekly Activity checklist (Sallis, Condon, et al., (1993)	Self-admin; 5 min; 1-day	x	x	x			x	x					4 days .74	Caltrac	.30
Yesterday Activity checklist (Sallis, Condon, et al., 1993)	Self-admin; 10 min; 7-days	x		x			x	x					4 days .60	Caltrac	.28

Measure & Original Ref.	Assessment Format	Type	Frequency	Intensity	Duration	Sedentary	Moderate	Vigorous	Leisure	Work	Household	Transport	Re-test reliability, Weighted mean	Validity criterion	Validity, weighted mean
Four by one-day recalls (Calle, 1994)	Interview; 1-day, repeated four times		x	x	x		x	x	x				4 weeks .62	Heart rate; Observation	.61 .72
Physical Activity Interview (Simons-Morton, et al., 1994)	Diary & interview; 12 min; 1-day			x	x		x	x						Heart rate; Caltrac	.63 .55
Physical Activity Rating (Janz et al., 1995)	Self-admin; 1 min; typical												2 weeks .89 (Sall is 93)	Heart rate; CSA accel.	.33 (Janz 92) .07
3-Day Sweat Recall (Janz et al., 1995)	Interview; Last 3 days		x					x						CSA accel.	.47
3-Day Aerobic Recall (Janz et al., 1995)	Interview; last 3 days		x					x						CSA accel.	.22

Measure & Original Ref.	Assessment Format	Type	Frequency	Intensity	Duration	Sedentary	Moderate	Vigorous	Leisure	Work	Household	Transport	Re-test reliability, Weighted mean	Validity criterion	Validity, weighted mean
Self-Admin. P.A. Checklist (Sallis et al., 1996)	Self-admin; 35 min; previous day	x		x	x	x	x	x	x		x		Same day correlation with other format .65	Heart rate; Caltrac	.60 .32
Physical Activity Checklist Interview (Sallis et al., 1996)	Interview; 18 min; previous day	x		x	x	x	x	x	x		x		Same day correlation with other format .65	Heart rate; Caltrac	.52 .38
Child/ Adolescent Activity Log (Garcia et al., 1997)	Self-admin; multiple 1-day to 3-day recalls	x	x	x	x		x	x					Same day .95	Caltrac	.39
P.A.Q. for Older Children (Kowalski, Crocker, Faulkner, 1997) (Crocker, Bailey et al., 1997)	Self-admin; <20 min; 7-day recall												1 week .79	Caltrac	.39

Measure & Original Ref.	Assessment Format	Type	Frequency	Intensity	Duration	Sedentary	Moderate	Vigorous	Leisure	Work	Household	Transport	Re-test reliability, Weighted mean	Validity criterion	Validity, weighted mean
Physical Activity Question-naire for Adoles-cents (Kowalski, Crocker, & Kowalski, 1997)	Self-admin; 10-15 min; 7-day recall													Caltrac	.33
Previous Day Physical Activity Recall (Weston et al., 1997)	Self-admin; 1-day recall	x	x	x		x	x	x	x	x		x	Same day .98	Heart rate; Caltrac; pedometer	.53 .77 .88
Parent and teacher reports of Child Physical Activity (Haaro, 1997)	Parent & teacher log; 4 days of logging			x	x			x						Heart rate; Caltrac	.40 .53
Parent and teacher log of child P.A. (Manios et al., 1998)	Parent & teacher logs for 3-5 days		x	x	x	x	x	x						Heart rate	Teacher = .59; Parent = .77

Table 2. Adult Self-Report Measures of Physical Activity

Measure & Original ref.	Assessment Format	Type	Frequency	Intensity	Duration	Sedentary	Moderate	Vigorous	Leisure	Work	Household	Transport	Re-test reliability, Weighted mean	Validity criterion	Validity, weighted mean
Minnesota Leisure-Time Physical Activity Questionnaire (Taylor, Jacobs, et al., 1978)	Interviewer Administered; Past 12 months	x	x	x	x		x	x	x		x		1 month Total .89; Mod. .84; Vig. .87 1 year Total .69 Mod. .32 Vig. .71	Cal trac Energy Intake-REE	Total.21 Mod .23 Vig .22 Total .17
Paffenbarger Physical Activity Questionnaire (Paffenbarger, Wind, & Hyde, 1978)	Self- or interviewer administered; Past week or year	x	x	x	x		x	x	x				2 wks .34 1 mo .72 7-12 wks .58; 9 mo .47; 8 mo .34; 1 yr .73	Caltrac Energy intake-REE	Total .30 Total .32
Seven-Day Physical Activity Recall (Sallis, Haskell, et al., 1985)	Interviewer administered; Past week	x	x	x	x		x	x	x				Same day Total .86 2 week Total .69 Mod .28 Hard .32 Very hard .72 5 week Total .58 1 month Total .34 Mod .12 Hard .31 Very hard .33	Caltrac Tri-Trac	Total .50 Mod .25 Hard/Very Hard .30 Total .53 Mod .11 Hard .67 Very hard .66
CARDIA Physical Activity History (Jacobs, Hahn, et al., 1989)	Interviewer administered; Past year	x	x	x			x	x	x				2 weeks Total .84 Mod .77 Heavy .79 1 month Total .88 Mod .66 Heavy .91	Caltrac	Total .14 Mod .11 Heavy.31
Baecke Questionnaire of Habitual Physical Activity (Baecke, Burema, & Frijters, 1982)	Self-administered; Habitual activity with no time frame		x	x					x	x			1 month Total .89 Leisure .87; Sports .92 3 months Leisure .74; Sports .81 5 months Total .84 Leisure .79 Sports .80	LSI Caltrac Intake-REE	Leisure .18 Sports .12 Total .23 Leisure .02 Sports.30 Total .21

Table 2. Continued

Measure & Original ref.	Assessment Format	Type	Frequency	Intensity	Duration	Sedentary	Moderate	Vigorous	Leisure	Work	Household	Transport	Re-test reliability, Weighted mean	Validity criterion	Validity, weighted mean
Baecke et al., continued													11 months Total .79 Leisure .76 Sports .73		
Godin Leisure-Time Exercise Questionnaire (Godin & Shephard, 1985)	Self-administered; Habitual weekly		x	x			x	x	x				2 weeks Total .74 Mod .46 Strenuous .94 1 month Total .62 Mod .36 Strenuous .84	Caltrac	Total .36

Table 3. Older adult self-report measures of physical activity

Measure & original ref	Assessment Format	Type	Freq	Intensity	Duration	Sed	Mod	Vig	leisure	work	household	transport	Retest reliability; weighted mean	Validity Criterion	Validity, weighted mean
Modified Baecke Q. for Older Adults (Voorrips et al., 1991)	Interview; 30 min; past year	x		x	x		x	x	x		x	x	20 days .85	pedometer; Caltrac	.72 (Voorips) .22 (Pols 96)
Physical Activity Scale for the Elderly (PASE) (Washburn et al., 1993)	Self-administer; 7 days	x	x	x	x		x	x	x	x	x		3-7 weeks .75	doubly labeled water	.58 (Schuit 97)
Yale Physical Activity Survey (YPAS) (DiPietro et al., 1993)	Interview; 20 min; typical in past month	x	partial	x	x	x	x	x	x	x	x		2 weeks .56	Caltrac	overall =.20; vigorous= .14; walk =.31
Zutphen Physical Activity Q. (Caspersen et al., 1985)	Survey; 15 min; usual	x	x	x	x		x	x	x		x	x	3 months .93	doubly labeled water	.62