Arquivos de Ciências do Esporte
Archives of Sport Sciences

Original Article

Feasibility of using a continuous direct observation technique for assessment of free-living physical activity in young adults

Viabilidade da utilização de uma técnica de observação direta contínua para avaliação da atividade física de lazer em adultos jovens

Jeffer E. Sasaki[1]
Dinesh John[1]
Amanda Hickey[1]
Kate Lyden[1]
Todd A. Hagobian[1,2]
Patty S. Freedson[1,*]

Abstract

Objective: To demonstrate the feasibility and application of 'continuous focal sampling' direct observation (CFS DO) for physical activity (PA) measurement in free-living adults. *Methods*: Nine observers were trained to use CFS DO and completed two video-based examinations to evaluate observer reliability. We applied the method in free-living conditions by recording activity type and intensity among thirty college-aged students during 11.1 ± 1.0 hr observation periods. *Results*: Percent correct classification of activity type and intensity by the observers were 86.6 ± 6.5% and 76.1 ± 15.4%, respectively. Test-retest reliability coefficients for activity type and activity intensity were r = .79 and r = .78. Based on CFS DO measures, participants spent 57.4% and 15.5% of the time sitting and walking. Mean time spent in sedentary, light, moderate, and vigorous physical activity intensities were 359.6 ± 100.1, 178.8 ± 107.3, 85.4 ± 63.1, and 24.6 ± 24.6 min for the 11.1 ± 1.0 hr observation period. *Conclusion:* The CFS DO technique was reliable for assessment of free-living PA in the current study. Feasibility of CFS DO may be limited to shorter blocks of observation (2-3 hr).

Keywords: physical activity assessment, free-living conditions, criterion measure.

Author's Affiliation

[1]University of Massachusetts, Amherst, MA, United States.
[2]California Polytechnic State University, San Luis Obispo, CA. United States.

**Corresponding author*

Department of Kinesiology, University of Massachusetts, 30 Eastman Lane, 111 Totman Building, Amherst, MA 01003- 9258, United States.
e-mail: psf@kin.umass.edu

Conflict of interest

The other authors do not have any conflict of interest related to this paper.

Peer-review duration

Received: 29/05/2015
Approved: 05/08/2015

Introduction

Free-living physical activity (PA) is comprised of many behavioral and physiological characteristics. Behavioral variables include activity type (e.g., walking, sports) and activity domain (e.g., occupational, leisure, transportation) while physiological variables include energy expenditure (e.g., kcals/day) and activity intensity (e.g., moderate, vigorous). Assessing PA in free-living conditions is challenging because measuring or estimating these variables is complex.

Most free-living PA assessment methods estimate one or more features of PA behavior and their validity is dependent on the gold standard methods used to validate them. Currently, gold standard methods for validating questionnaires and accelerometer-based activity monitors are doubly labeled water and portable indirect calorimetry[1-3]. None of these gold standard methods assess behavioral aspects of human movement, which is essential for validation of indirect methods that estimate both behavioral and physiological variables of free-living PA.

A gold standard method to simultaneously assess physiological and behavioral variables of PA is direct observation (DO). Direct observation has been extensively used to measure PA in children and adolescents[4-10]. Both momentary time sampling (MTS) (e.g., SOFIT, SOPLAY) and continuous focal sampling techniques (CFS) (e.g., CARS, C-SOFIT) have been employed to assess PA in children and adolescents at school, during physical education class, and at home[7,9,11,12]. In MTS DO, an activity is observed during fixed time intervals (e.g., 20 sec) and recorded either using paper and pencil or a computer. There is an underlying assumption that activity does not change significantly from the observation interval until the end of the recording interval. Using CFS DO, we record the duration, frequency, type, and intensity of activity. A new event is recorded every time activity changes. This method is also referred to as 'duration coding method'[11,13-15]. Studies have shown that MTS DO produces similar results as CFS DO when assessing PA behavior in children and adolescents. However, these studies have mostly assessed PA behavior in specific settings such as in physical education classes or on the playground[9,11,12]. While progress has been made with DO for children and adolescents, no studies have proposed DO methods to specifically assess free-living PA in adults.

To record adult free-living PA, DO must capture behaviors that occur in a variety of settings as well as transitions that occur throughout the observation period. Given these requirements, CFS DO appears to be the most appropriate DO technique for assessing free-living PA in adults. However, we only know of one study using CFS DO to assess PA in adults and it was confined to a 'laboratory setting'. This study was conducted in our laboratory and investigated the validity of CFS DO in assessing PA intensity in adults[16]. To the best of our knowledge, no study has extended the use of CFS DO to measure free-living PA of adults in natural settings. This is partly due to time requirements for DO measurement, potential reactivity of participants, and subjectivity of observers[17]. It is not known if it is feasible to use CFS DO technique to observe free-living PA in adults. Therefore, the purposes of this study were (a) to create a custom CFS DO coding menu and train observers to use it, and (b) to test the feasibility of using the CFS DO technique to measure free-living physical activity in young adults.

Methods

Instrumentation

A hand-held personal digital assistant (PDA) (134 g; 8.1cm x 1.5cm x 11.4 cm) (Palm Tungsten E2, Palm Inc., Sunnyvale, CA.) with a touch screen was used to record PA. The PDA was programmed with custom software containing a dropdown menu of 24 commonly performed activities (Figure 1) and a sub-menu containing 4 numerical codes corresponding to an intensity category (1= sedentary [<1.5 MET], 2= light [1.50 – 2.99 METs], 3= moderate [3 – 5.99 METs], and 4= vigorous [≥6.0 METs]). A time stamp icon on the touch screen could be selected to record the start and stop times of activities, with the smallest recordable activity duration set to one second. Data were stored on a removable flash memory disk.

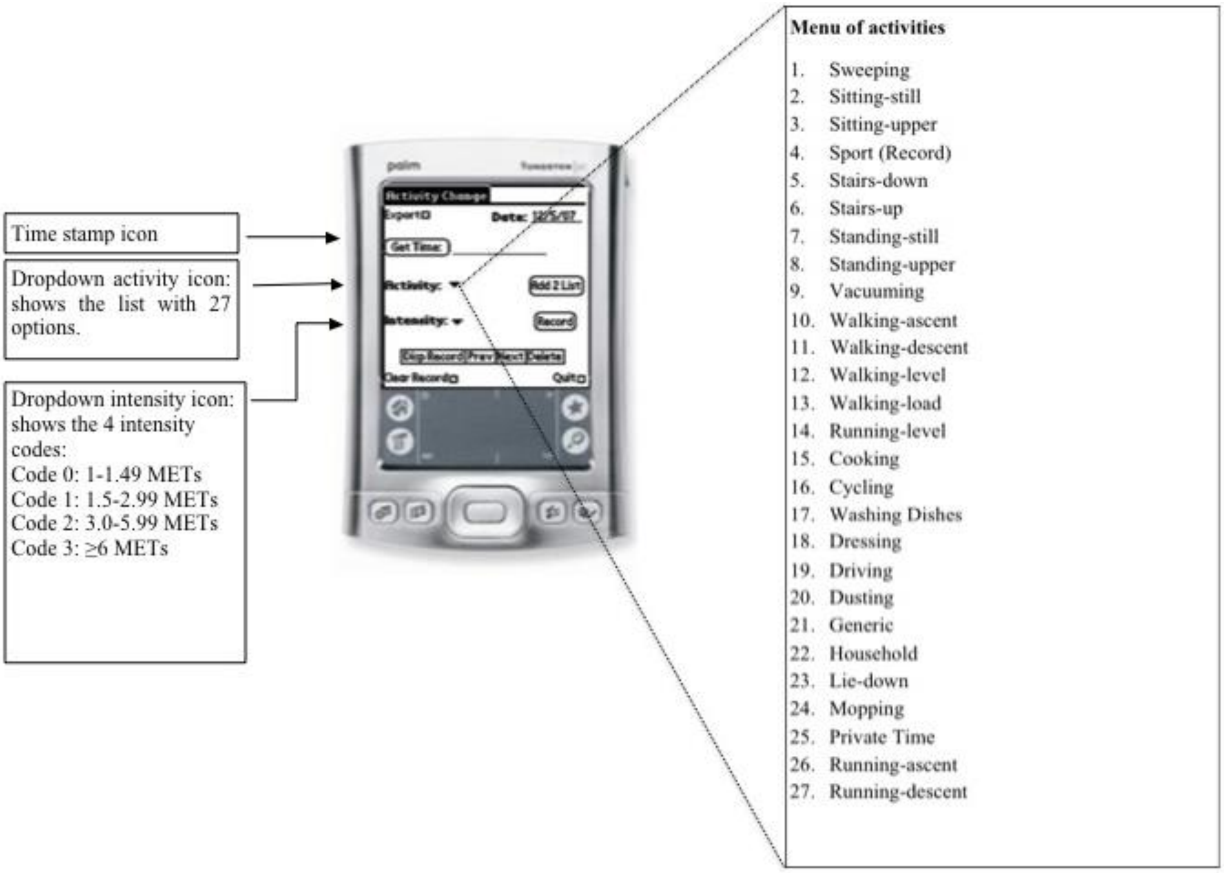

Figure 1. Personal digital assistant device, custom software and menu of activities used in the study.

Activity Menu

Two senior researchers (PSF and TAH) identified commonly performed day-to-day activities from the *Compendium of Physical Activities*[18] and created the activity menu (Figure 1). The options 'Sport' and 'Generic' were included to record activities absent from the activity menu. The 'Private Time' option was added to account for periods when participants required privacy (*e.g.,* bathroom breaks).

Observer Training

Nine observers were trained to use the direct observation system in 3 phases. In phase one, a senior researcher (TAH) provided observers with instructions on PDA operation and data collection procedures during two face-to-face sessions that were about one-hour long each. Some noteworthy instructions were given as follows: *"Avoid conversation (unless necessary) with participants and maintain a reasonable distance (5 to 20 feet) to minimize participant reactivity. Estimate activity intensity by observing both activity pace (e.g., approximate walking speed) and physiological responses of the participants (e.g., sweating, breathing hard). Account for increased energy expenditure if participants are carrying a load (e.g., backpack). If unclear about activity intensity, select the best possible intensity code and take notes describing activity intensity in a pocket memo-book. After data collection, consult senior researchers to discuss these notes and correctly classify activity intensity."*

In phase two, each observer first viewed a 25-minute training video comprising various free-living activities. The video included audio and subtitled instructions on classification of activity type, duration, and intensity. Observers then practiced DO for at least 5 to 6 hr on their own. Queries on the correct classification of activity type and intensity during these practice sessions were clarified in consultation with senior researchers (TAH and PSF).

The last phase of training simulated an actual free-living DO session. In this session, the trained observers recorded the activities of a senior researcher (TAH). Each observer performed DO for at least one hour. The senior researcher verified accuracy of the observed activity types and intensities.

Observer Testing

In the post-training period, observers completed 2 video tests separated by at least 24 hours. The same video was used on both occasions. The video test contained 20 video clips of different free-living activities requiring 52 coding entries: 3 were sedentary activities, 37 were light intensity activities, 9 were moderate intensity activities, and 3 were vigorous intensity activities. Observers recorded activity type and intensity for each video clip. Results of these trials were used to examine inter and intra-observer reliability and validity for activity type and intensity.

Free-living CFS DO

Thirty healthy students (mean values: age= 23.7 ± 6.0 years old; BMI= 24.4 ± 4.2 kg.m-2) from the University of Massachusetts Amherst volunteered to take part in this study. All participants were given a detailed explanation of the study and signed a written informed consent document that was previously approved by the institutional review board.

Direct observation sessions were conducted on weekdays and each participant was observed for at least 10 consecutive hours. Observers met participants at a predetermined time and location in the morning and continuously recorded activity type, intensity and duration on the PDA. Participants were instructed to maintain their habitual daily routine and ignore the presence of the observers. Observers were rotated in shifts lasting between one and three hours during the 10-hour DO session for each participant. It is important to note that the objective of this part of the study was to demonstrate the feasibility of assessing free-living PA using CFS DO and not to assess regular PA behavior, as this would demand assessment of more days.

Data reduction and statistical evaluation

Inter and Intra-Observer Analyses

Observer testing results were scored as percent correct classification of activity type and intensity using the coding of the expert observers as 'the truth' (TAH and PSF). For activity type, observer responses were scored as either correct (score of 1) or incorrect (score of 0). For activity intensity, observer answers were weighted in relation to the coding of the 'expert -observers', using the following scoring procedure:

1) Correct code (criterion-observers' coding)= 4
2) Incorrect code, one level higher or lower= 3
3) Incorrect code, two levels higher or lower= 2
4) Incorrect code, three levels higher or lower= 1

Percent correct classification of activity type and intensity for each observer were then calculated as described below:

%correctactivity type= (sum of correct coding instances/number of coding instances)*100

%correctactivity intensity = (sum of weighted scores/(number of coding instances*4))*100

Intra-observer test-retest reliability was computed using Spearman rank order correlation coefficients between coding from tests 1 and 2 from each observer. Kruskal-Wallis tests were used to examine differences between observer scores.

Direct Observation Data

Direct observation data from the PDA were exported to a laptop as individual spreadsheets for each subject. The spreadsheets contained information about observed activity type, corresponding intensity codes and time stamps for each activity. Observers consulted with senior researchers to clarify issues regarding activities classified as 'Sport' and 'Generic' and the entries were modified if necessary. 'Private-Time' entries were eliminated from all analyses.

Total time spent in each activity type and intensity category, and the number of bouts of 10 or more minutes of moderate-to-vigorous intensity PA was determined for each participant. Percent time spent in different activity types was calculated for the whole sample.

Results

Intra and Inter-Rater Statistics

Spearman rank order correlations indicated high intra-observer test-retest reliability between coding from observer tests 1 and 2 for activity type (r = .79) and intensity (r = .78) for the 52 classifications. The percent correct classifications among the nine observers for activity type and activity intensity were 86.6 ± 6.5% and 76.1 ± 15.4% based on criterion (coding by senior researchers TAH and PSF). No significant differences were observed for activity type and intensity between observer test scores (Kruskal-Wallis analysis). Average percent misclassification for each intensity category is presented in Table 1. Percent misclassification was higher for sedentary (22.3 ± 37.3%), followed by overestimation of light intensity PA (11.7 ± 7%) and underestimation of moderate intensity PA (11.1 ± 14.7%). Observers did not record 13.2 ± 8.5% coding instances.

Table 1
Average Percent Misclassification of intensity by observers.

	Sedentary	Light		Moderete		Vigorous	
		Underestimation	Overestimation	Underestimation	Overestimation	Underestimation	Missed
Mean (%)	22.2	0.9	11.7	11.1	8.6	3.7	13.2
SD	37.3	1.4	7.0	14.7	9.3	11.1	8.5

Direct Observation Data

The total free-living observation time was 334 hr, with an average of 11.1 ± 1.0 hr per participant. A minimum of three and a maximum of six observers observed each participant. The average shift duration for each observer was 2.5 ± 0.3 hr. Based on DO, participants spent an average of 359.6 ± 100.1 min in sedentary activities (sitting, driving, and lying down) and 178.8 ± 107.3 min in light, 85.4 ± 63.1 min in moderate intensity and 24.6 ± 24.6 min in vigorous intensity PA. Participants engaged in 2 ± 2 bouts of moderate-to-vigorous intensity PA lasting at least 10 min. Nine participants did not engage in any 10 min+ moderate-to-vigorous intensity PA bouts (Figure 2). Sitting still, sitting with upper body movement and walking accounted for more than 70% of total observation time (Figure 3).

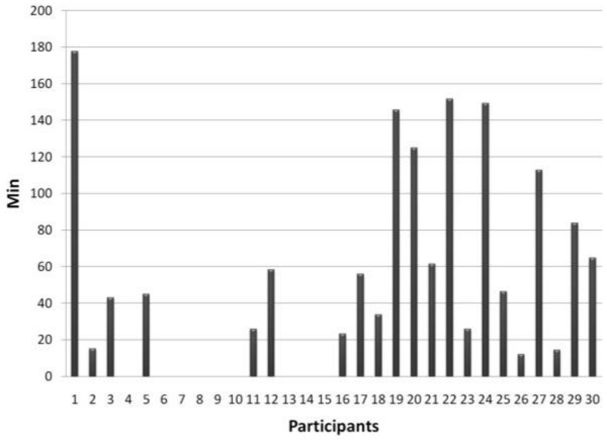

Figure 2. Total time spent in moderate or vigorous PA in bouts of 10 min+ for each participant.

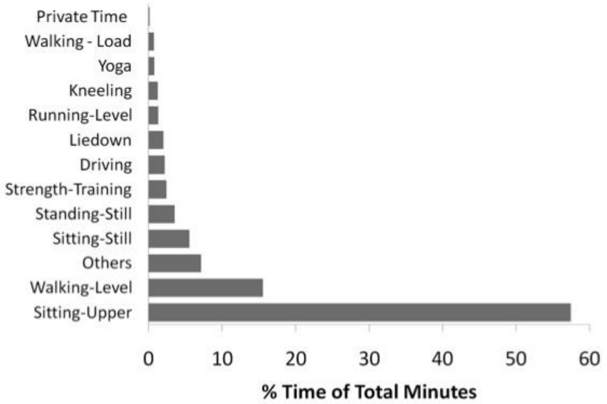

Figure 3. Percentage of the total time spent in different activities. Note: 1) 'Others' represents the sum of activities that were rarely performed and/or were very short in duration, 2) 'Sitting-upper' represents times when participants were sitting and performing upper body movement (*e.g.,* computer work, writing, organizing desk), 3) Some of the activities were not specifically listed in menu of activities but were coded under the 'Generic' option (*e.g.,* Strength training, Yoga).

Discussion

The present study showed that the CFS DO is a reliable and valuable technique to assess free-living PA in adults. Observers received standardized training and were tested to ensure accurate and reliable assessment of free-living PA with the CFS DO technique. Training and testing observers is labor intensive taking approximately 10-15 hours for each observer. Continuous Focal Sampling DO is a low cost method that only requires a 'recording device and a software that can be easily customized to meet specific needs (e.g., The Observer© (Noldus Information Technology, Wagenigen, Netherlands).

In the field, we showed that the CFS DO technique enables the continuous measurement of free-living PA and is not limited to a particular domain/setting. In contrast to MTS DO methods that use specific and sometimes long observation intervals (20 s to 5 min.) 4–10, the current CFS DO technique has a minimum sampling interval of one second. Observers, therefore, can record activities at the exact time they start and stop. This is important for assessing unstructured PA behavior in free-living conditions where most activities last for varied durations.

In this study, we used a convenience sample of graduate and undergraduate Kinesiology students, with most of the observations occurring on the campus of the university. Thus, extending our findings about the feasibility of CFS DO to the general population is not possible. Individuals who move to different locations during the day (e.g. from work to home to store) may be challenging when transitioning between observers. Although there are several challenges with observing individuals over extended periods of time (e.g. 11 hrs), we have recently used the CFS DO technique to validate accelerometer-based PA prediction models in free-living conditions for shorter time periods. For these studies, prolonged observation of habitual PA behavior is not necessary. Therefore, we use shorter blocks (2-3 hr) and provide instructions to participants to complete a wide array of sedentary behaviors and physical activities. Using this approach to obtain observer-based criterion measures of physical activity and sedentary behavior provides the necessary information to test accelerometer models that estimate activity time, type and intensity in free-living settings.

Unlike doubly labeled water and indirect calorimetry, CFS DO provides information about behavioral and physiological aspects of human movement. In the present study, we quantified PA type, intensity and minutes of MVPA lasting for 10 min or more (Figures 1-2) from direct observation. This is important to validate machine-learning models that estimate activity type, duration and intensity from accelerometer-based activity monitors (Bonomi et al., 2009; Freedson et al., 2011; John et al., 2011; Staudenmayer et al., 2009). These machine-learning models applied to the accelerometer signal features are accurate in classifying PA type and intensity using during controlled laboratory-based activity menus where activities are performed for fixed time intervals (e.g. one minute) 19–22. Nevertheless, activities in free-living conditions vary in duration and often occur randomly for short and varying time intervals. The CFS DO is an effective method for evaluating the performance of these machine-learning techniques in processing accelerometer data collected in free-living conditions. Similarly, researchers can use CFS DO to re-examine the validity of various accelerometer cut-points that classify free-living PA intensity on a minute-by-minute basis 23,24.

Continuous Focal Sampling DO may also be used to evaluate the accuracy of different measures of free-living sedentary behaviors. For example, CFS DO was used to examine the validity of different accelerometer cut-points in assessing free-living sedentary time 25. When using hip-worn activity monitors, it is challenging to distinguish between acceleration patterns of very light intensity activity (e.g., standing) and sedentary behaviors (sitting, lying). Our method accurately distinguishes such behaviors (Figure 2) and has potential to be a criterion measure for sedentary behavior.

Although we developed our own CFS DO application, there are specialized DO software in the market, such as the "The Observer© (Noldus Information Technology, Wagenigen, Netherlands)", that allows for easy customization of the coding menu. However, researchers might choose to use open-source software to create their own DO application. In this case, our study serves as a guide for creating and implementing a CFS DO method. The framework for creating the coding system, training observers, collecting and processing data, and applying CFS DO in the field was described in this manuscript. Researchers may tailor the CFS DO technique for their own purposes by customizing their training procedures and coding menu according to the variables of interest. We have been able to refine training that is delivered to observers. For example, we have observed that an important aspect to increase accuracy and precision of observers is to conduct discussion sessions after initial face-to-face and video training. These meetings allow observers to describe how they are coding activities and ask questions about specific issues. This process serves to minimize random and systematic coding errors among observers. It is also important to conduct practice observation sessions to compare expert and trainee results.

Strengths and Limitations

We reported the use of CFS DO technique to comprehensively assess free-living activity and sedentary behavior of adults during waking hours. The use of a lightweight, hand-held recording device enabled continuous recording of activities throughout the day. This allowed researchers to obtain information on PA behavior in different domains and settings. The novel use of continuous sampling allows for assessment of short duration activities that often takes place in free-living conditions. Additionally, the CFS DO technique herein described is low-cost and can be easily adapted to researchers' needs.

The current study has limitations. The presence of observers may have influenced participant behavior. However, this does not affect our findings since our aim was to show the feasibility of using CFS DO as a criterion measure to validate other free-living PA estimation methods and not as a technique to assess habitual PA behavior. A potential source of measurement error in the CFS DO technique is the subjective estimation of activity intensity. A recent study in our laboratory reported that observers are able to estimate activity intensity accurately in comparison to indirect calorimetry with biases for light intensity PA and MVPA of 2.1% and 4.9%, respectively 16. In addition, our results indicated high test-retest reliability for observers in classifying activity intensity (r = .78). A final limitation of the CFS DO is data processing, which requires considerable time and thus might not be appropriate for studies with a large sample size.

Conclusion

The present study demonstrated that CFS DO has limited feasibility in assessing free-living PA in young adults over an 'extended period of time'. Based on our experience, it is more feasible to use CFS DO during shorter-blocks of time in which participants perform behaviors of interest. The technique was reliable and may be used to simultaneously assess behavioral (e.g., activity type and domain) and physiological (e.g., activity intensity) aspects of free-living PA. The technique will be useful for validating and examining new and existing methods for predicting free-living PA from wearable sensors.

Referências

1. Bonnefoy M, Normand S, Pachiaudi C, Lacour JR, Laville M, Kostka T. Simultaneous validation of ten physical activity questionnaires in older men: a doubly labeled water study. J Am Geriatr Soc. 2001 Jan;49(1):28–35.
2. Ekelund U, Yngve A, Westerterp K, Sjöström M. Energy expenditure assessed by heart rate and doubly labeled water in young athletes. Med Sci Sports Exerc. 2002 Aug;34(8):1360–6.
3. Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. Obesity (Silver Spring). 2007 Oct;15(10):2371–9.
4. Brown WH, Pfeiffer KA, McIver KL, Dowda M, Almeida MJCA, Pate RR. Assessing preschool children's physical activity: the Observational System for Recording Physical Activity in children-preschool version. Res Q Exerc Sport. 2006 Jun;77(2):167–76.
5. Heath EM, Coleman KJ, Lensegrav TL, Fallon JA. Using momentary time sampling to estimate minutes of physical activity in physical education: validation of scores for the system for observing fitness instruction time. Res Q Exerc Sport. 2006 Mar;77(1):142–6.
6. Klesges RC, Coates TJ, Moldenhauer-Klesges LM, Holzer B, Gustavson J, Barnes J. The FATS: An observational system for assessing physical activity in children and associated parent behavior. Behavioral Assessment. 1984;6(4):333–45.
7. McKenzie TL, Sallis JF, Nader PR. SOFIT: System for Observing Fitness Instruction Time. J Teach Phys Educ. 1991 Jan 10;11(2):195–205.
8. O'Hara NM, Baranowski T, Simons-Morton BG, Wilson BS, Parcel G. Validity of the observation of children's physical activity. Res Q Exerc Sport. 1989 Mar;60(1):42–7.
9. Puhl J, Greaves K, Hoyt M, Baranowski T. Children's Activity Rating Scale (CARS): description and calibration. Res Q Exerc Sport. 1990 Mar;61(1):26–36.
10. Sit CHP, Lam JWK, McKenzie TL. Direct observation of children's preferences and activity levels during interactive and online electronic games. J Phys Act Health. 2010 Jul;7(4):484–9.
11. Keating XD, Kulinna PH, Silverman S. Measuring Teaching Behaviors, Lesson Context, and Physical Activity in School Physical Education Programs: Comparing the SOFIT and the C-SOFIT Instruments. Measurement in Physical Education and Exercise Science. 1999;3(4):207–20.
12. McKenzie TL, Marshall SJ, Sallis JF, Conway TL. Leisure-time physical activity in school environments: an observational study using SOPLAY. Prev Med. 2000 Jan;30(1):70–7.
13. Altmann J. Observational Study of Behavior: Sampling Methods. Behaviour. 1974;49(3-4):227–66.
14. Powell J, Martindale A, Kulp S. An evaluation of time-sample measures of behavior. J Appl Behav Anal. 1975;8(4):463–9.
15. Tacha TC, Vohs PA, Iverson GC. A Comparison of Interval and Continuous Sampling Methods for Behavioral Observations. Journal of Field Ornithology. 1985 Jul 1;56(3):258–64.
16. Lyden K, Petruski N, Mix S, Staudenmayer J, Freedson PS. Validation of direct observation for estimating physical activity and sedentary behavior in a free-living setting. J Phys Act Health. 2012;In press.
17. Vanhees L, Lefevre J, Philippaerts R, Martens M, Huygens W, Troosters T, et al. How to assess physical activity? How to assess physical fitness? European Journal of Cardiovascular Prevention & Rehabilitation. 2005 Apr 1;12(2):102–14.
18. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc. 2000 Sep;32(9 Suppl):S498–504.
19. Bonomi AG, Goris AHC, Yin B, Westerterp KR. Detection of type, duration, and intensity of physical activity using an accelerometer. Med Sci Sports Exerc. 2009 Sep;41(9):1770–7.
20. Freedson PS, Lyden K, Kozey-Keadle S, Staudenmayer J. Evaluation of artificial neural network algorithms for predicting METs and activity type from accelerometer data: validation on an independent sample. J Appl Physiol. 2011 Dec;111(6):1804–12.
21. John D, Liu S, Sasaki JE, Howe CA, Staudenmayer J, Gao RX, et al. Calibrating a novel multi-sensor physical activity measurement system. Physiol Meas. 2011 Sep;32(9):1473–89.
22. Staudenmayer J, Pober D, Crouter S, Bassett D, Freedson P. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. J Appl Physiol. 2009 Oct;107(4):1300–7.
23. Bassett DR Jr, Rowlands A, Trost SG. Calibration and validation of wearable monitors. Med Sci Sports Exerc. 2012 Jan;44(1 Suppl 1):S32–8.
24. Strath SJ, Bassett DR Jr, Swartz AM. Comparison of MTI accelerometer cut-points for predicting time spent in physical activity. Int J Sports Med. 2003 May;24(4):298–303.
25. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. Med Sci Sports Exerc. 2011 Aug;43(8):1561–7.
26. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc. 2008 Jan;40(1):181–8.