Physical Activity by Self-Report: A Brief History and Future Issues

William L. Haskell

For the scientific domain of physical activity and public health research to advance its agenda of health promotion and disease prevention continued development of measurement methodologies is essential. Over the past 50 years most data supporting a favorable relationship between habitual physical activity and chronic disease morbidity and mortality have been obtained using self-report methods, including questionnaires, logs, recalls, and diaries. Many of these instruments have been shown to have reasonable validity and reliability for determining general type, amount, intensity, and bout duration, but typically do better for groups than individuals with some instruments lacking the sensitivity to detect change in activity. During the past decade the objective assessment of physical activity using accelerometer-based devices has demonstrated substantial potential, especially in documenting the pattern of light-, moderate-, and vigorous-intensity activity throughout the day. However, these devices do not provide information on activity type, location or context. Research that combines the strengths of both self-report and objective measures has the potential to provide new insights into the benefits of physical activity and how to implement successful interventions.

Keywords: questionnaires, sedentary behavior, measurements, energy expenditure

The Oxford English Dictionary defines science as “the intellectual and practical activity encompassing the systematic study of the structure and behavior of the physical and natural world through observation and experiment.” In a more restricted sense it is defined as “a branch of study which is concerned either with a connected body of demonstrated truths or with observed facts systematically classified and more or less colligated by being brought under general laws, and which includes trustworthy methods for the discovery of new truth within its own domain.” Thus, any scientific domain, including physical activity and public health, needs to consist of facts systematically classified based on trustworthy methods for establishing new truths. One major challenge presented by these definitions for the domain of physical activity and public health is how best to obtain valid and reliable data on habitual physical activity in diverse, free-living populations. Over the past half century self-report of physical activity (PA) has contributed significantly to the documentation of the clinical benefits attributed to a more physically active lifestyle throughout the lifespan. How has assessment of PA evolved over time and has this evolution been effective in answering more detailed questions about the complexities of PA behaviors of interest to scientists, practitioners, and the public?

The Science Base for Physical Activity and Public Health

The science base for the domain of physical activity and public health has been contributed to by a variety of professions using a diversity of scientific methods. These professions have included basic and applied exercise/activity biology, physical activity epidemiology, clinical science, athletic (sports) training, behavioral science, genetics and the evolving investigations of the biology, epidemiology, and psychology of inactivity or sedentary behavior. Much of this science base has been established by direct observation and measurement during experiments in basic and applied science laboratories and physical education/athletic research facilities and not dependent on self-report of activity that is not directly observed. However, to study what activities people do throughout their daily lives, direct observation was not practical and up until recently nor was objective measurement. Thus, self-report of PA became the default alternative even though the various limitations of this methodology have been well recognized. Despite these limitations, much of what has been learned over the last 60 years about the relations between habitual physical
activity and clinical events is based on data collected by self-report.

Early Documentation of the Relationship Between Physical Activity and Chronic Disease

The domain of physical activity and public health is rooted in the population-based research and writings of Professor “Jerry” Morris and colleagues in the 1950s. However, development of this science base began much earlier as man’s biological responses to exercise were being systematically measured in the 18th and 19th centuries. For example Sir John Floyer in 1707 invented a watch with a second hand (the “physician’s watch”) that permitted the first accurate measurement of pulse rate during exercise. In the 1850s, Dr. Edward Smith measured carbon dioxide production and heart rate in prisoners assigned to hard labor in London prisons—attempting to demonstrate the inhumane nature of the punishment of strenuous exercise. Throughout the first half of the 20th century, exercise scientists, especially physiologists, made major contributions to this science base regarding the acute and chronic responses of many tissues, organs, and systems to a wide variety of activity types, durations, and intensities. Much of this research involved exercise training with a primary focus being on improvements in the various components of physical fitness with much less emphasis on health outcomes.

Occupation as a Marker of Physical Activity

In 1953, based on the occupational classification of male employees of the London Transport Executive (drivers versus conductors of double-deck buses) and of London Civil Servants (mail carriers versus telephonists plus other employees with sedentary jobs), Morris and colleagues first proposed the “vigorous physical activity and coronary heart disease prevention” hypothesis. Their results, that men more active on the job had fewer and less severe acute coronary events than less active men, was followed by reasonably rapid publication over the next 2 decades of cross-sectional and longitudinal observational studies using occupation as the PA exposure variable and all-cause or disease-specific mortality as the primary outcome. By 1963, at least 13 additional studies had been published in peer-reviewed scientific journals exploring the relationship between physical activity and mortality with physical activity classifications based on occupation. No self-report information regarding PA or sedentary behavior during work or leisure time was included in these reports. For a summary of these studies see the review by Fox and Haskell in 1967. Also, during this time mortality rates were compared among athletes involved in different sports and nonathletes.

Studies using occupational classification to categorize subjects as active or inactive were conducted when in developed or developing countries there were still significant segments of the workforce who routinely performed moderate-to-vigorous intensity activity for some portion of their workday. In most studies where the physical activity differences in the contrasting samples were substantially different, mortality rates were significantly different between the groups. During this period most employees worked a “full day” at least 5 days/week so that the amount of moderate- or vigorous-intensity activity performed by the more active workers substantially exceeded current public health physical activity recommendations. Men in the more physically active jobs included lumberjacks, postmen, farmers, sharecroppers, construction workers, and semi- and unskilled laborers. These studies begin to shift the thinking prevalent at the time that jobs requiring hard labor were a significant cause of various chronic diseases. In the US, studies continued to be published through the mid-1970s using occupational classification as the indicator of PA exposure.

Of note, none of these studies included data on women.

Early Use of Self-Report Questionnaires to Document Physical Activity

In 1964, an article published in the Journal of the National Cancer Institute by Hammond reported on the association between habitual physical activity determined by self-report and all-cause mortality. The question on physical activity was included in an extensive multifactor evaluation of 442,094 men who then were followed for 34 months. The questionnaire was very short with low participant burden: How much physical activity do you get at work or play? Possible answers were: none, slight, moderate, heavy. The relative risk for age-standardized death rates between men reporting heavy versus none was 0.54 for nonsmokers and 0.64 for smokers. Following this publication there was a major shift away from using occupational classification to self-report as the main PA exposure variable. Self-report was used to identify dominant activities during occupation or leisure-time, total PA (occupational plus leisure-time PA) or just leisure-time activity (LTPA).

A good example of a very early project using a self-report questionnaire designed to obtain both occupational and leisure-time activity is the Health Insurance Plan of Greater New York study published by Frank and colleagues in 1966. In addition to questions about time spent walking and doing other active tasks on the job and during leisure time it also asked about the amount of time spent sitting on the job. These investigators created a PA score based on the assignment of points to each answer; higher points for more vigorous-intensity activities or for spending more time at a moderate- or vigorous-intensity activity. By adding up these points 3 activity categories were created: least, moderate, and most active. Compared with the least active men, the most active men had lower rates of CHD. No information was provided on the relationship between sitting time and CHD events.
The Limited Evolution of Physical Activity
Self-Report Over the Past 30 Years

For approximately 50 years, data collected using various types of self-report questionnaires have been critical in establishing a strong favorable relationship between the amount of PA habitually performed and major chronic disease morbidity and mortality. During these years more diseases have been added to this list, additional biomarkers identified that are positioned in the causal pathway between activity and clinical events and subjects other than predominately white men have been included, especially women. Also during this time it has been shown that a wide variety of activity types and intensities confer health benefits and that the dose-response relationship for many clinical outcomes extends from quite small amounts of PA above the least active (especially when the least active are very inactive) to very large amounts and high intensities of activity. While the general shape of this relationship is beginning to become somewhat clear, many details are still undefined, especially at the lowest and highest ends of the activity spectrum.15

While the science base for the domain of physical activity and public health has demonstrated enormous growth, the advancement of physical activity assessment by self-report has not been that robust. Many of the questionnaire formats and methodologies have changed little from the 1970s or ’80s when the Minnesota Leisure-Time Physical Activity Questionnaire, the College Alumni Physical Activity Survey, the Stanford 7-Day Physical Activity Recall, and other questionnaires were developed. In the last 30+ years, the accuracy, validity, and reliability of a number of these questionnaires have been determined in more diverse populations and some shown to be sensitive to change. However, only limited advancements have been made in providing details about other specifics of the PA, such as activity type, profile or pattern throughout the day, and context or location.

As the importance of habitual physical activity in the health of the public has become more established through research, new questions about these relationships are being asked. The answers to some of these questions will be best answered using objective measures, some with self-report measures and some only with a combination of measures. So far most objective measures using motion or physiological sensors only provide information on PA intensity and bout duration and do not define a specific type of activity, context of the activity or location. The accelerometer-based units can tell when a person is inactive or active and give some indication of absolute intensity while some newer units can tell if the person is sitting or standing. Using wireless multisensor devices in the future it will be possible to obtain some classification of activity type, intensity, bout duration, and location of the activity and provide real-time feedback using mobile phones or similar devices. However, for measures of type, context, and location of activity, some combination of self-report and objective measurement will likely provide the most accurate, valid, and reliable information at low subject burden and reasonable investigator cost.16

There exist a number of physical activity self-report instruments that have demonstrated good validity and reliability in documenting leisure-time physical activity as well as total daily physical activity for groups (less well so for individuals) of men or women. In addition, some of these instruments do a reasonably good job at classifying activities into general absolute intensity categories (light, moderate, vigorous) and document time spent in well-defined activities such as walking. Difficulty increases when attempts are made to identify time spent in specific activities of daily living (especially those of light intensity), obtaining information about activities performed at earlier times in a person’s life (lifetime PA), the context and location that the activities were performed and change in PA over weeks, months, or years.

In addition to the continued development and evaluation of existing or new PA questionnaires, surveys, logs, or recalls, other approaches need to be considered as alternatives or adjuncts to these typical self-report measures. One possibility is implementing ecological momentary assessment (EMA) using mobile phones to collect real-time PA data in response to random time prompting Data from published time use surveys should be considered when attempting to put data from self-report into a broader perspective. Such data might be of particular value when evaluating PA change over extended periods.17 Use the following URL to obtain information on time use research (http://www.rntu.org).

Looking Ahead—Measurement of Sitting or Light-Intensity Activity

Is sitting or sedentary behavior a unique risk factor or is it just the low end of the activity spectrum? It is interesting to note that the Morris study of London double-deck bus drivers and conductors formed the foundation for physical activity epidemiology more than 50 years ago, while within the last decade this same study has been cited as the start of the investigation into the physiology of inactivity.19 Studies on rats where the hind limbs are off-loaded so their leg muscle are not contracting results in detrimental changes in fat and carbohydrate metabolism.19 This experimental condition has been likened to sustain sitting in humans and has been one stimulus for a great deal of speculation and an increasing amount of research on the independent effects of sitting—separate from being physically inactive—on health.20 The evidence to date supporting sitting or sedentary behavior as a chronic disease risk factor independent of PA is still quite limited compared with the enormous database supporting the protective effect of moderate-to-vigorous intensity PA.15

Answers to a vast number of questions that need to be asked about sitting as a risk factor separate from being physically inactive (which is the high risk comparison group for most studies of PA) will need advancements
in both self-report and objective measurement methods to separate sitting from other forms of light-intensity activity. Historically it has been demonstrated that using existing self-report tools the ability of participants at any age to accurately report time spent in low-intensity activities is very poor. For example, in developing the 7-day physical activity recall at Stanford University in the 1970s, it was determined that the accuracy of reporting time spent in light-intensity PA for the past week by adults was exceedingly poor. Thus, it was decided that rather than ask about time spent performing light-intensity PA it would be more accurate and expedient to obtain time spent sleeping and time spent in moderate, hard, and very hard-intensity activities, add these times together and subtract from 24 to get time spent in light-intensity PA. Also, to try and study the effects of sitting by asking about TV viewing time is likely to lead to erroneous conclusions when many adults spend much of a 40+ hour workweek sitting but not viewing television. It is encouraging that recently questionnaires have focused on total sitting time and have demonstrated reasonable validity.

Similar challenges exist for using objective measurement tools (primarily accelerometer-based movement detectors worn on the hip) for accurately identifying sitting time from other forms of light-intensity activity. The results of recent studies have shown that time spent being very inactive (lying, sitting, standing quietly) can be identified reasonably well compared with activities resulting in more movement, but that such devices cannot tell sitting from other postures. More recently developed equipment such as the Actual worn on the thigh does identify sitting from standing but not necessarily sitting from lying. Using wireless triaxial accelerometers placed on both the upper and lower body and more advanced analytical approaches will provide new information on posture, light-intensity activity, and activity type.

Katzmarzyk and colleagues reported on all-cause mortality associated with time spent reported sitting and classified men and women according to meeting physical activity recommendations. The results of recent studies have shown that time spent being very inactive (lying, sitting, standing quietly) can be identified reasonably well compared with activities resulting in more movement, but that such devices cannot tell sitting from other postures. More recently developed equipment such as the Actual worn on the thigh does identify sitting from standing but not necessarily sitting from lying. Using wireless triaxial accelerometers placed on both the upper and lower body and more advanced analytical approaches will provide new information on posture, light-intensity activity, and activity type.

Katzmarzyk and colleagues reported on all-cause mortality associated with time spent reported sitting and classified men and women according to meeting physical activity recommendations. The results of recent studies have shown that time spent being very inactive (lying, sitting, standing quietly) can be identified reasonably well compared with activities resulting in more movement, but that such devices cannot tell sitting from other postures. More recently developed equipment such as the Actual worn on the thigh does identify sitting from standing but not necessarily sitting from lying. Using wireless triaxial accelerometers placed on both the upper and lower body and more advanced analytical approaches will provide new information on posture, light-intensity activity, and activity type.

Closing the Gap on Self-Report of Activity and Inactivity

Standardization of Definitions

Especially at the low end of the PA intensity spectrum, standardization of definitions of what is being measured needs to be improved. Definitions need to be highly precise and consistent in their use. This is true for both self-report and objective measurements. Current discussions interchange postures, motion, intensity, energy expenditure, and muscle activation. For some, sedentary behavior means “sitting” while others include activities that produce less than 100 counts from an Actigraph worn at the hip or any activity considered to be less than 1.5 METs. How should lying down be classified and at what body angle does lying become sitting? While most people are sitting when they watch TV, some stand while working in the kitchen and others do much of their TV viewing in bed—so TV viewing time may not necessarily be a very accurate surrogate for sitting. We won’t even deal with those rare people who watch TV while riding a stationary cycle or walking on a treadmill at home or the local gym.

Absolute, Relative, and Perceived Intensity

Using the typical self-report is absolute, relative or perceived intensity of an activity being determined? When responses such as specific activities are classified into intensity categories (light, moderate, vigorous) using estimates of energy expenditure (frequently in METs), the investigator is assigning an absolute intensity to the activity. In this case the capacity of the person to perform the activity is not taken into account so that for a person with a low capacity the activity will be at a higher relative intensity (relative to their capacity) than a person with a higher capacity. It is generally assumed that perceived intensity (how hard they perceive the activity to be) is more closely related to relative than to absolute intensity when a subject is asked if an activity they perform (e.g., gardening) is light, moderate or vigorous intensity, is their response based on perceived, relative or absolute intensity? It is important to remember that nearly all data from observational studies in the general population using self-report or objective measures has categorized the activities by absolute intensity, but in most experimental studies the intensity of the activities assigned to intervention subjects has used relative intensity (e.g., % max heart rate, % max oxygen uptake). It may be possible to obtain relative PA intensity data from self-report by using a very standardized rating of perceived exertion scale.

Physical Activity Energy Expenditure and Self-Report

The rationale for using physical activity energy expenditure (PAEE) as a consolidated exposure variable for a
number of self-report instruments has been based on 2 objectives. First, PAEE was adopted as a means of having an overall index of how much PA a person reported. By assigning each activity an energy expenditure value (eg, METS), it is possible to combined time spent in different activities (eg, walking, cycling, swimming) and at different intensities into 1 exposure variable such as MET-minutes/week. In this case the goal was to have an overall indicator of PA and was not an attempt to establish a measure of energy balance. The other objective of estimating PAEE has been to try and determine the energy expenditure of the subject in relation to either their body weight (or body composition) or their energy intake. Unfortunately, one cannot determine caloric expenditure by determining self-report PAEE and dietary intake by self-report. Some PA self-report instruments can put people in correct categories of PAEE relative to each other but the absolute PAEE frequently is not correct.

**Being Objective Does Not Mean It Is Correct**

Just because an objective measure is used does not mean that either the data or its interpretation is correct. If objective and self-report results do not match with one another, the question seems to be—what is wrong with the self-report instrument? However, the first question that needs to be asked is “why do the results appear different from one another”? On closer observation of data from accelerometer-based devices, it appears they are detecting features of physical activity not routinely captured by self-report and vice versa. This may be particularly true when the exposure is closely linked to precise time and intensity criteria. For example, with accelerometer output one can precisely detect a ≥10-minute bout of activity above a specific threshold (eg, average number of counts estimated to equal 3.0 mph for a specific sample of subjects). If during this 10+ minute walk self-reported by the subject acceleration counts drop below the threshold for a minute or more this 10-minute bout is not counted. Or if the accelerometer placement or characteristics of the subject results in his/her accelerometer counts at 3.0 mph are less that the population threshold it is possible that the objective measure of this bout of activity will be ignored.

**Self-Report May Be Doing Better Than You Think**

Interest in and use of objective measures of PA in public health oriented research has rapidly accelerated over the past decade. Such measures are a valuable addition to PA measurement methodology and will continue to develop over the next several decades as new technologies and analytical approaches come on line. However, it needs to be remembered that nearly all of the PA exposure data supporting a link between habitual PA and chronic diseases and which all public health PA guidelines are based comes from self-report. If these guidelines are anywhere near correct then PA self-report has been useful and very likely will continue to be useful in the future. Closing the gap should continue to make data collected by PA self-report a highly valuable research methodology.

**References**