

## The Effect of Item Order on Physical Activity Estimates Using the IPAQ

Jodi Barnett<sup>1</sup>, Claudio Nigg<sup>1</sup>, Ilse De Bourdeaudhuij<sup>2</sup>, Christine Maglione<sup>1</sup>, Jay Maddock<sup>1</sup>

<sup>1</sup>*University of Hawaii at Manoa*

<sup>2</sup>*Ghent University, Ghent, Belgium*

### Abstract

**Objective:** To investigate the effect of item order on physical activity (PA) estimates using the International Physical Activity Questionnaire – short form (last 7 day) telephone interview (IPAQ-S7T). **Method:** The IPAQ-S7T was conducted for Time 1 (Hawai`i) and sample 1 (Belgium). Then item order was reversed for Times 2, 3 (Hawai`i), and sample 2 (Belgium). **Results:** In Hawai`i, vigorous activity levels were higher, walking levels lower for Time 1, compared to Times 2 and 3. Moderate activity was higher at Time 1 than Time 3. The Belgium study reported higher levels of vigorous activity for sample 1, compared to sample 2. Total volume of PA across time points was unchanged for the Hawai`i study, but decreased with reversed items in the Belgium study. **Conclusion:** Item ordering affects the levels and duration of reported PA using the IPAQ-S7T. This warrants further research on order efficacy and whether other PA measures are affected.

© 2007 Californian Journal of Health Promotion. All rights reserved.

*Keywords: survey design, item order, physical activity, self-report*

Physical inactivity is a global health problem that causes more than two million deaths each year making it one of the top 10 leading causes of death and disability ([WHO](#), 2007).

Two physical activity (PA) recommendations are endorsed by the American College of Sports Medicine (ACSM) and [Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion](#) (2007): Three to five days a week of vigorous activity for at least 20 minutes per session (ACSM, 1990; Pate et al., 1995); or at least 30 minutes of moderate intensity physical activity (PA) or more on most, preferably all, days of the week (Pate et al., 1995). Current estimates from the World Health Organization ([WHO](#), 2007) report that 60% to 85% of adults in countries around the world “are not active enough to benefit their health.” Accurate assessment of physical activity is essential for planning, implementing, and evaluating public health programs and policies addressing this issue (NCCDPHP, 2007; Sarkin, Nichols, Sallis, & Calfas, 2000).

Recently, the International PA Questionnaires ([IPAQ](#), 2007) were developed as a set of internationally comparable, valid, and reliable instruments (Craig et al., 2003) that examine self-reported PA and sedentary behaviors. Two versions exist: the long version which was created to provide a comprehensive assessment of physical activity practices (five activity domains asked independently) and the short form (four general PA items) which was designed for large population surveillance in which time and space is limited (Craig et al., 2003). Two formats of each version exist: the telephone format and the self-administered format.

The IPAQ – short form (last seven days) telephone interview (IPAQ-S7T) was employed as the behavioral indicator in a statewide psychosocial surveillance effort in Hawaii (Maddock, Marshall, Nigg, & Barnett, 2003). Results of the first assessment showed high levels of vigorous PA compared to other statewide surveillance results. As a result, IPAQ creators recommended switching the order of items, asking walking first, followed by

moderate, and then vigorous activity. Other studies have also reported high vigorous activity levels in other countries using the IPAQ-S7T (Rutten et al., 2003; Rzewnicki, Auweele, & De Bourdeaudhuij, 2002). Independently, similar data was collected in Belgium facilitating the item order comparison within different continents.

Item ordering effects are commonplace in academic tests (Lane, Bull, Kundert, & Newman, 1987; Neely, Springston, & McCann, 1994) and psychological surveys (Dahlstrom, Brooks, & Peterson, 1990). To date, no research has been conducted on item-ordering effects using the IPAQ questionnaires. Therefore, this report examines the impact of two different IPAQ-S7T item orders via three independent administrations in Hawai'i, and two independent administrations in Belgium, on the type and duration of self reported PA.

We hypothesized that changing the order of the items to ask vigorous PA last would result in a decrease in reporting of vigorous PA and an increase in reported moderate PA and walking.

## Methods

### Data Collection - Hawaii

Three cross-sectional surveys using random digit dialing procedures of Hawaii's non-institutionalized adult population were conducted in Spring 2002 (N = 4,706), Fall 2002 (N = 4,555), and Spring 2003 (N = 4,557). These sample sizes give us a precision estimate of + 2% with a 99% confidence interval for the state. The cooperation rate was 29.8%, 26.8%, and 32.3% at Times 1, 2, and 3 respectively. The person over 18 who had the last birthday was asked to complete the 10-minute interview, providing randomization within household. Interviews were guided by a computer aided telephone interview (CATI). Additional information on the surveillance system has been reported elsewhere (Maddock et al., 2003).

### Data Collection – Belgium

Two cross-sectional surveys of adults were conducted in Belgium, Europe. For the Sample 1, a random digit dialing national sample of Belgian adults, with characteristics generally

similar to the population was drawn (N=611). The cooperation rate was 38.7%. For Sample 2, a random sample of 1000 residents of Ghent, Belgium, was drawn for the age group 18-65 years old. A questionnaire and a prepaid and pre-addressed envelope were sent to the respondents' home. After six weeks a second letter to ask for participation was sent to all addresses. The final sample consisted of 521 participants (52.1% cooperation rate). Additional information on both surveys has been reported elsewhere (six and seven for Sample 1; and 11 for Sample 2).

## Measures

**Demographics.** Participants were asked a series of demographic questions, including age, sex, education attained, income level, and ethnic identification for the Hawai'i study and age, sex, education attained in the Belgium study.

International Physical Activity Questionnaire – short form telephone interview (IPAQ-S7T). This instrument assesses time spent engaging in vigorous PA, moderate PA, and walking per week. High reliability and good validity with CSA accelerometers have been demonstrated in 12 countries (Craig et al., 2003).

## IPAQ Procedures

For Time 1 in Hawaii and Sample 1 in Belgium, question order began with vigorous PA, followed by moderate PA, and then minutes of walking per week. This order was reversed for Times 2 and 3 in Hawaii and for Sample 2 in Belgium. Total activity per week is calculated as the sum of activity in all three domains.

The IPAQ was scored using the standardized protocol suggested by the Research Committee (IPAQ, 2007). Respondents were placed into one of three categories; (1) Insufficiently Active - No activity reported or some activity but not enough to meet Categories 2 or 3; (2) Sufficiently Active – three or more days of vigorous activity of at least 20 minutes per day, five or more days of moderate-intensity activity or walking of at least 30 minutes per day, or five or more days of any combination of walking, moderate-intensity or vigorous intensity activities achieving a minimum of at least 600

MET –min/week; (3) Highly Active – Vigorous-intensity activity on at least three days and accumulating at least 1500 MET –minutes/week, or seven or more days of any combination for walking, moderate-intensity or vigorous intensity activities achieving a minimum of at least 1500 MET –minutes/week.

All means and median levels of PA are presented using raw, un-capped data to obtain true sample estimates. Placement into categorical levels of PA was performed using capped levels of PA based on recommendations set forth by the scoring protocol. Activity was capped at 120 minutes per day in all three domains to prevent the over categorization of individuals into Highly Active while still identifying those who are Sufficiently Active.

### Analyses

Data were analyzed using SPSS 11.5. For Hawaii, the statewide representative data was weighted by gender and county to match 2000 US Census Bureau census estimates ([US Census](#), 2000). Data were not weighted for the Belgian samples. Descriptive data are displayed using means, standard deviations and frequencies. A criterion was set at  $p < .01$  *a priori* for all analyses to indicate significance due to the large sample sizes used. ANOVAs with follow-up Bonferonni tests [ $p < .01/3 = .0033$ ] were used to assess differences across time points for continuous variables. For categorical variables, chi-square tests were used.

## Results

### Sample Characteristics

Hawaii study: At Time 1, the sample consisted of 49.9% male, an average age of 46.1 years, 14.4 years of education, and a median income of \$40-49,000 per year. No significant differences were found across Times 2 and 3. The sample was ethnically diverse; Caucasian (33%), Japanese (15.6%), Hawaiian (21.3%) and Filipino (11.5%). Small differences in the ethnic distribution were displayed across time;  $\chi^2(10) 40.0, p < .001$ .

Belgian Study: Sample 1 consisted of 45.7% males, an average age of 48.6 years, and 11.7

years of education. Sample 2 consisted of 51.7% males, and average age of 41.0 years, and 12.6 years of education. No significant demographic differences were found between the Belgian samples ( $p > .05$ ).

### Change in Levels of Physical Activity

Reported PA levels for all time points in Hawaii are displayed in Table 1. Significantly higher levels of vigorous activity and lower levels of walking were reported during Time 1 when compared with Times 2 and 3. Moderate activity was significantly higher at Time 1 than at Time 3. However, total volume of PA across time points did not change.

Reported PA levels for the two Belgian samples are displayed in Table 2. Significantly higher levels of vigorous activity and of total volume of PA were reported in Sample 1 when compared with Sample 2. No significant differences in moderate activity and in walking were found between the two surveys.

Median levels of PA are also presented in Table 1 for Hawaii and in Table 2 for Belgium given the non-normal distribution of the sample. This was recommended by the IPAQ Research Committee. A large drop in median vigorous PA minutes was demonstrated from Time 1 (120) to both Times 2 (15) and 3 (20) in Hawaii. The large drop in median vigorous PA minutes was also found from Sample 1 (60) to Sample 2 (0) in Belgium.

### Change in Categorical PA Levels

Differences were evident in the percentage of individuals meeting Insufficiently Active, Sufficiently Active, and Highly Active criteria across time for Hawaii ( $\chi^2(4) 45.2, p < .001; \theta^2 = .048$ ) and for Belgium ( $\chi^2(2) 9.6, p = .008; \theta^2 = .009$ ). The largest differences were demonstrated in those meeting highly active criteria; in Hawaii Time 1 (61.6%), versus Time 2 (56.3%) and Time 3 (57.2%), and in Belgium Sample 1 (66.8%) versus Sample 2 (58.5%). Overall around 80% of participants were at least sufficiently active over time.

Table 1  
Change in levels of PA per week across time points in the Hawai'i study

Variable	Wave I n = 4,706	Wave II n = 4,555	Wave III n = 4,557	Effect Size
<b>IPAQ categories</b>	<b>Minutes</b> Mean (SD)	<b>Minutes</b> Mean (SD)	<b>Minutes</b> Mean (SD)	<b>eta<sup>2</sup></b>
Vigorous**	355.9 (586.9) <sup>2,3</sup>	185.0 (421.1) <sup>1</sup>	186.8 (429.2) <sup>1</sup>	.027
Moderate**	283.9 (530.9) <sup>3</sup>	308.3 (547.0)	330.8 (561.7) <sup>1</sup>	.001
Walking**	467.6 (732.6) <sup>2,3</sup>	587.7 (864.89) <sup>1</sup>	537.2 (804.0) <sup>1</sup>	.004
Total Activity	1107.3 (1299.4)	1081.1 (1291.2)	1054.8 (1267.8)	.000
<b>IPAQ categories</b>	<b>Minutes</b> (Median)	<b>Minutes</b> (Median)	<b>Minutes</b> (Median)	<b>eta<sup>2</sup></b>
Vigorous	120	15	20	-
Moderate	80	120	120	-
Walking	180	210	210	-
Total Activity	580	570	560	-

\*\* Significant Oneway ANOVA across time points. [Superscripts represent significant pair-wise comparisons between time points using a Bonferonni correction, p < .0033; each superscript indicates the time point it is significant from]

Table 2  
Change in levels of PA per week across samples in the Belgium study

Variable	Sample 1 n = 611	Sample 2 n = 521	Effect Size
<b>IPAQ categories</b>	<b>Minutes</b> Mean (SD)	<b>Minutes</b> Mean (SD)	<b>eta<sup>2</sup></b>
Vigorous**(p<0.001)	333.1 (671.2)	119.7 (303.2)	.038
Moderate	389.0 (658.0)	340.5 (639.9)	.001
Walking	880.2 (1180.6)	827.3 (1067.0)	.001
Total Activity**(p=0.002)	1602.3 (1768.1)	1287.6 (1539.9)	.009
<b>IPAQ categories</b>	<b>Minutes</b> (Median)	<b>Minutes</b> (Median)	<b>eta<sup>2</sup></b>
Vigorous	60	0	-
Moderate	120	90	-
Walking	420	420	-
Total Activity	930	780	-

\*\* Significant Oneway ANOVA across surveys using a Bonferonni correction, p < .0033.

### Discussion of Results

Our results indicate that item order of the IPAQ-S7T significantly affects the duration of vigorous PA reported. This is reflected in higher mean levels of vigorous activity and in the

percentage of individuals meeting Highly Active criteria when vigorous activity was investigated first during Time 1. This also corresponds with lower levels of moderate PA reported during the first administration compared with Time 3 and

lower levels of walking compared with Times 2 and 3 in the Hawai'i study. These patterns are in the direction hypothesized and correspond to changes in IPAQ item order between time points. Total volume of PA reported remained stable across time points regardless of item order in the Hawai'i study. The Belgian study replicated the vigorous activity conclusions. However, the overall quantity of PA decreased when walking was asked first. The high prevalence of walking in the Belgian samples may explain that no further increase in walking was seen upon the reordering of the items. It is recommended that keeping the order of items constant between survey periods is important to obtain the most reliable prevalence estimates of duration and intensity of PA.

A limitation should be noted. The studies use independent cross-sectional samples rather than administering the two-item orders to the same samples. However, two independent studies from two continents and a second replication time point in the Hawai'i study, increase our confidence to infer that results were associated

with question order rather than population or survey administration method differences (for the Belgian study).

A subject for concern is the high levels of PA reported in the current studies and other studies to date using the IPAQ-S7T, regardless of item order (Rutten et al., 2003; Rzewnicki, Auweele, & De Bourdeaudhuij, 2002). One explanation is that the IPAQ-S7T is an all-day PA indicator versus a leisure time physical activity indicator. Currently, recommendations only exist for leisure time PA.

In conclusion, this study found that item order affects reporting of the duration of different PA intensities, and may affect the total volume of PA reported but not the proportion of participants meeting national PA guidelines. Further research is warranted to determine if the findings reported here hold true when other PA indicators are used or the IPAQ-S7T is used in other populations and, most importantly, which order is more valid.

## References

- American College of Sports Medicine (ACSM). (1990). Position stand: The recommended quantity and quality of exercise of exercise for developing and maintaining cardiorespiratory and muscular fitness in healthy adults. *Medicine & Science in Sports & Exercise*, 22, 265-274.
- Craig, C. L., Marshall, A. L., Sjostrom, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E. et al. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, 35, 381-395.
- Centers for Disease Control and Prevention. (2007). National Center for Chronic Disease Prevention and Health Promotion home page. Retrieved March 1, 2007, from <http://www.cdc.gov/nccdphp/>
- Dahlstrom, W. G., Brooks, J. D., & Peterson, C. D. (1990). The Beck depression inventory: Item order and the impact of response sets. *Journal of Personality Assessment*, 55, 224-233.
- De Bourdeaudhuij, I., Sallis, J. & Saelens, B. (2003). Environmental correlates of physical activity in a sample of Belgian adults. *American Journal of Health Promotion*, 18, 83-92.
- International Physical Activity Questionnaire. (2007). IPAQ — international physical activity questionnaire. Retrieved March 1, 2007, from <http://www.ipaq.ki.se/>
- Lane, D. S., Bull, K., Kundert, D. K., & Newman, D. L. (1987). The effects of knowledge of item arrangement, gender, and statistical cognitive item difficulty on test performance. *Educational & Psychological Measurement*, 47, 865-879.
- Maddock, J., Marshall, C., Nigg, C., & Barnett, J. (2003). Development and first year results of a psychosocial surveillance system for chronic disease related health behaviors. *Californian Journal of Health Promotion*, 1(Sp Issue), 54-64.
- Neely, D. L., Springston, F. J., & McCann, S. J. (1994). Does item order affect performance on multiple-choice exams? *Teaching of Psychology*, 21, 44-45.
- Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C. 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.
- Rutten, A., Ziemainz, H., Schena, F., Stahl, T., Stiggelbout, M., Vanden Auweele, Y. et al. (2003). Using different physical activity measurements in eight European Countries. Results of the European Physical Activity Surveillance System (EUPASS) time series survey. *Public Health Nutrition*, 6, 371-376.
- Rzewnicki, R., Auweele, Y. V., & De Bourdeaudhuij, I. (2002). Addressing over reporting on the international physical activity questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutrition*, 6, 299-305.
- Sarkin, J. A., Nichols, J. F., Sallis, J. F., & Calfas, K. J. (2000). Self-report measures and scoring protocols affect prevalence estimates of meeting physical activity guidelines. *Medicine & Science in Sports & Exercise*, 32, 149-156.
- US Census Bureau. (2000). Census 2000 data for the State of Hawaii. Retrieved March 1, 2007, from <http://www.census.gov/census2000/states/hi.html>
- World Health Organization. (2007). Why move for health. Retrieved March 1, 2007, from <http://www.who.int/moveforhealth/en/>

### **Acknowledgements**

The Hawaii study was funded by the Hawaii Department of Health through the Tobacco Settlement Special Fund. Special thanks to QMARK Research and Polling for completing the telephone interviews. The Belgian studies were funded by the Ghent University, and by the Health Monitoring Programme of the European Union (EUPASS, European Commission, Luxembourg, agreement number VS1999/5133).

### Author Information

Jodi Barnett  
University of Hawaii at Manoa  
Department of Public Health Sciences  
John A. Burns School of Medicine  
University of Hawaii at Manoa

Claudio Nigg\*  
University of Hawaii at Manoa  
Department of Public Health Sciences  
John A. Burns School of Medicine  
University of Hawaii at Manoa  
1960 East-West Road, Honolulu, HI 96822  
Ph.: 808-956-2862  
Fax.: 808-956-5818  
E-Mail: [cnigg@hawaii.edu](mailto:cnigg@hawaii.edu)

Ilse De Bourdeaudhuij  
Ghent University, Ghent, Belgium

Christine Maglione  
University of Hawaii at Manoa  
Department of Public Health Sciences  
John A. Burns School of Medicine  
University of Hawaii at Manoa

Jay Maddock  
University of Hawaii at Manoa  
Department of Public Health Sciences  
John A. Burns School of Medicine  
University of Hawaii at Manoa

\* corresponding author