



FACULTY OF
BUSINESS &
ECONOMICS

HILDA PROJECT DISCUSSION PAPER SERIES No. 3/14, December 2014

The Measurement of Physical Activity in Wave 13 of the HILDA Survey

*Mark Wooden**

**The HILDA Survey Project was initiated, and is funded, by the
Australian Government Department of Social Services**

* The assistance of Simon Freidin with the preparation of statistics reported on in this paper is gratefully acknowledged.

Introduction

Every wave of the HILDA Survey conducted to date has included one simple question intended to measure the extent an individual engages in physical activity. Included each year in the self-completion questionnaire, it measures the frequency per week of participation in moderate or intensive physical activity. The question reads as follows:

In general, how often do you participate in moderate or intensive physical activity for at least 30 minutes?

Moderate level physical activity will cause a slight increase in breathing and heart rate, such as brisk walking.

Six pre-coded response categories are provided: “Not at all”, “Less than once a week”, “One to two times a week”, “Three times a week”, “More than 3 times a week (but not every day)”, and “Every day”.

The information collected from this question is very coarse. It does not measure the amount of time individuals actually spend in physical activity; it treats all types of activity (subject to satisfying the criteria that it must at least cause a slight increase in breathing and heart rates) the same, regardless of the intensity of that activity; and it only permits responses within bands.

Given the prominence of obesity as a public health issue, it has been suggested that the HILDA Survey should give greater priority to the collection of data that would measure both the extent of obesity within the Australian population and factors that contribute to obesity. As a result, data on (self-reported) height and weight began to be collected in wave 6 (and data on waist measurement in wave 13). In wave 9, a dedicated health module was designed that included, among other things, questions about diet. More detailed questions on physical activity, based on questions included in the Active Australia Survey (see AIHW 2003), were also trialled as part of the dress rehearsal for that survey wave. Interviewer feedback about both the quality of responses being provided and the relatively large imposition on respondent time, however, led us not to continue with these questions for the wave 9 main survey.¹

In wave 13, with the re-inclusion of the health module (which is scheduled to be included every 4 years), we again gave consideration to the inclusion of questions that would enable the construction of more reliable and informative measures of physical activity. For this wave we were drawn to the International Physical Activity Questionnaires (IPAQ), which over the last decade or so have increasingly been used for population surveillance of physical activity among adult populations (though not in Australia, where the Active Australia Survey is dominant, and implemented in large part in the 2011-12 Australian Health Survey conducted by the Australian Bureau of Statistics [ABS]).²

This paper: introduces the IPAQ and the types of summary measures it generates; describes its implementation in wave 13 of the HILDA Survey; and reports brief summary statistics from the wave 13 data describing the distribution of responses on the key outcome variables and associations with other health-related variables.

¹ A major difference between the version trialled in wave 9 and the Active Australia Survey instrument was the use of the term “usual week” in the HILDA wave 9 dress rehearsal, rather than the “last 7 days”. It was this distinction that we suspect accounted for much of the uncertainty in responses of many dress rehearsal sample members.

² Examples of the use of IPAQ within a cross-national context include Sjöström et al. (2006) and Bauman et al. (2009).

Survey Instrument

The IPAQ

The IPAQ is a self-reported measure of physical activity that is intended for the assessment of population levels of activity within and across countries. Unlike some other measures, it is designed to cover activity that occurs in all domains of life, rather than just during leisure time.³

It actually comprises a set of four different instruments: long and short versions for either self-administration or administration by telephone. All instruments identify three specific types of activity assessed over a “last 7 days” recall period. The long version differs from the short version in asking details about physical activities that occur within each of four domains: leisure time; domestic and gardening activities; work-related; and transport-related.

The short-form essentially involves two questions for each of three types of physical activity: (i) walking; (ii) moderate-intensity activity; (iii) and vigorous-intensity activity. The two questions measure first, the number of days (in the last week) on which the activity occurred, and second, the amount of time spent on one of those days (with the intent being to identify a usual or typical day during that week). In addition, there is one further question measuring the usual time spent sitting on a weekday.

Both versions have been reported to have acceptable measurement properties – good test-retest reliability and adequate criterion validity (Craig et al. 2003) – with the short format recommended for use in national prevalence studies. That said, there is some dispute around this. Like all self-reported measures, the IPAQ is only a proxy for actual activity, and associations with objective data (usually derived from accelerometers) are often very modest. Craig et al. (2003), for example, report mean correlations across 16 samples for the short form of just 0.30, which they regard as acceptable. Similarly, Lee et al. (2011) report, from their review of validation studies, correlations for total physical activity ranging from just .09 to 0.39, but which average around 0.28. However, and in contrast to Craig et al. (2003), they argue that this is a level that does not reach minimal acceptable standards. Lee et al. (2011) also point to evidence of high levels of over-estimation, with five out of six studies finding estimates of physical activity (when converted into metabolic equivalent minutes) from IPAQ that are substantially greater than the MET equivalent derived from accelerometry – between 36% and 173% greater.⁴

That said, the ability of these measures to correctly classify people as being sufficiently active (i.e., meeting some recommended activity threshold) is generally very good, with around 80% of all individuals covered by the sample analysed by Craig et al. (2003) being similarly classified by both the IPAQ forms and the accelerometer data.

It can also be argued that comparisons of self-reported time data with time estimates derived from accelerometry are not strictly valid. First, the usefulness of count data from accelerometers for measuring physical activity other than that which mainly involves lower-extremity or trunk acceleration, such as walking, jogging, running and climbing stairs, is questionable (Hendelman et al. 2000). Second, the time data derived from self-reports (and diaries) concern elapsed time and thus will include periods of inactivity (though accelerometry studies can make allowance for short interruptions). It is thus not surprising that estimated time spent in physical activity will be larger in self-reported measures than in

³ Information about the IPAQ, including scoring protocols and downloadable versions of the instruments, is available at: <https://sites.google.com/site/theipaq/>

⁴ See also Boon et al. (2010).

measures derived from accelerometer counts; some of the difference simply reflects differences in the underlying construct being measured.

Ultimately, with use of accelerometers considered neither practical nor feasible for use in the HILDA Survey, the only choices we faced was whether the survey would benefit from the inclusion of a more detailed self-reported measure of physical activity than currently collected, and if yes, which measure to include.

Implementing the IPAQ in the HILDA Survey

With the long version involving 27 questions, compared with the 7 questions for the short version, it was determined that the former was not practical for administration within the HILDA Survey, especially given the competition for interview time and questionnaire space. It was also determined that the repetitive nature of the long format would not be well received by some participants, a conclusion supported by observations made by Craig et al. (2003, p. 1388). Such considerations are very important in longitudinal surveys where the survey experience in one wave can influence sample member participation at the next.

Space constraints in wave 13 also dictated the use of the interviewer-administered version rather than the self-administered version. While the former is designated as being designed specifically for telephone administration, we could see no reason why the same script could not also be used for administration directly to respondents in their home by an interviewer via computer-assisted (CAPI) methods (as is the predominant practice in the HILDA Survey).

The question sequence ultimately administered in wave 13 of the HILDA Survey, however, was not a precise duplication of the IPAQ short format instrument, with a number of modifications made following the dress rehearsal. These were as follows:

1. The question on sitting was omitted.

While the IPAQ short format instrument is not long, the overall length of the HILDA Survey instrument that was trialled for wave 13 was still in excess of the limit specified in our sub-contract with Roy Morgan Research (the organisation sub-contracted to undertake the data collection). It was, therefore, decided to omit the question on sitting time. This was the question within the sequence expected to take most time. Unlike the questions on physical activities, all sample members would be expected to record a positive response and thus have to undertake a mental calculation. Further, that calculation was expected to be associated with considerable measurement error, both because of the wide range of sitting activities and because of the difficulty separating time spent sitting from time spent on one's feet. Finally, responses to the question on sitting do not affect the recommended summary measure of physical activity.

2. The introduction was amended to make it clear to respondents that they would first be asked about vigorous activity, then moderate activity and finally walking.
3. When asking about time spent on activity during a day, an explicit response option was provided for respondents who indicated that activity time varied from day to day. Persons selecting this option were then asked to estimate the total time spent on that activity during the entire 7-day period.

This is in line with the intent of the recommended IPAQ sequence. The difference is that we made the provision of a “don't know – varies from day to day” option explicit.

4. All persons who provided an answer on time spent that was in excess of 3 hours on any of the three activities were asked a further question clarifying that their answer only covered a single day (and not some longer period).

The inclusion of this question reflected concerns about the number of large outliers in the dress rehearsal sample. It was hypothesized that in some of these cases respondents may be mistakenly reporting time spent in activity over a week rather than over a single day. It was felt that the questions were conducive to such errors given respondents are first asked to recall the numbers of days during the last week spent on each activity group.

5. The list of examples of moderate physical activities was amended from “carrying light loads, bicycling at a regular pace, or doubles tennis” to “carrying light loads, gentle swimming, cycling at a modest pace, or social tennis”.
6. An additional question was included, which asked respondents to indicate whether their activity levels in the last 7 days were more, less or about the same as usual.

A paper representation of the question sequence administered in wave 13 is provided in Appendix A.

Comparisons with the Active Australia Survey instrument

Different survey instruments have been used to assess levels of physical activity within populations, with previous research demonstrating that the prevalence estimates derived from these different surveys can be far apart. Notably, an Australian study has found large differences in reported physical activity time across the same individuals when administered different instruments (Brown et al. 2004). In particular, it was found that prevalence rates based on IPAQ items were much higher than estimates based on the Active Australia (AA) Survey items.

As previously noted, the AA instrument has become the dominant method for measuring the prevalence of physical activity in Australia, as reflected in its use in the 2011-2012 Australian Health Survey. Nevertheless, we opted for inclusion of the IPAQ rather than the AA items. The reasons for this were at least twofold.

First, the AA instrument requires respondents to report on total time spent on each activity over the entire week. The experience from the wave 9 dress rehearsal was that this required mental calculations that some respondents found difficult, and which increased interview time. In contrast, the IPAQ focuses on just the time spent on one day during the week, which we felt would be simpler for respondents to answer (which is then multiplied by the number of days on which that activity type is undertaken).

Second, the AA instrument does not clearly define whether all physical activity is covered or only activity during leisure time. The ABS (2013) recognises this problem. It claims that “the purpose of the physical activity questions (regarding walking, moderate and vigorous activity) is meant to be focused on leisure time”. But if so, this is far from clear from the question wording; as the ABS notes, the questions do not specifically exclude work-based activity. Interviewers are apparently trained to prompt respondents to exclude activity during working time, but the ABS admits that this will not be clear to some respondents who will likely include activity during work. We took the view that the IPAQ approach of including all types of physical activity, regardless of where it occurred, was both more straightforward and preferable.

Survey Administration

As already noted, the IPAQ items were administered as part of the personal interview component of the HILDA Survey in wave 13. Given physical activity will typically be viewed as a socially desirable activity, we might expect a tendency for self-reports to cause activity levels to be overstated. Unfortunately, we cannot quantify the level of upward reporting bias that exists in the data. Somewhat differently, such upward biases might be expected to be greater in an instrument administered in person, rather than over the phone or where the instrument is self-administered. This we can test for, with evidence from simple regression models of physical activity time suggesting that responses do not vary systematically with response mode.⁵

Responses on physical activity undertaken during a week prior to interview might also be expected to be sensitive to both the amount of daylight and the weather during that week (see Tucker & Gilliland 2007), and hence to the time of the year (and the location) in which the survey is conducted. The interviews for wave 13 of the HILDA Survey were conducted between late July 2013 and the end of February 2014, but with most interviews (just over 80%) conducted in the two months August and September (i.e., late winter and early spring). Some evidence on the presence (or absence) of seasonality in the survey responses on physical activity levels is presented later in this paper.

Data Processing

The data collected were edited and processed (mostly) in line with rules recommended for the IPAQ Short-Form.⁶ Most importantly, this included the construction of derived variables measuring total activity time for each activity type measured in metabolic equivalent of task (or MET) minutes, and a categorical variable that sorts the population into three groups: low, moderate, and high activity levels.

Data Editing

Any cases where the sum of the reported daily time on walking, moderate activity and physical activity exceeded 16 hours were set to missing on the grounds that the values were unreasonably high. This affected 97 cases (or just 0.5% of the total responding sample).

The recommended processing rules also required that activity be recorded in minimum blocks of at least 10 minutes, and hence any responses to the daily time questions of 1 to 9 minutes were set to zero. No such adjustments, however, needed to be made to the wave 13 HILDA Survey data.

The guidelines also require any values of '15', '30', '45', '60' or '90' that were recorded as hours to be recoded as minutes, on the grounds that these are probably instances of miscoding. In the HILDA Survey CAPI script, however, answers in excess of 24 are not accepted. Further, all persons who provide an answer that is recorded as being in excess of three hours are given an additional question confirming that their answer is indeed for one day. Given this, we have opted not to edit the data in the way recommended. That said, we

⁵ We estimated simple linear regression models of measures of physical activity time (for men and women separately) that included controls for age, marital status, the number of dependent children, self-assessed health, labour force status and working time, month of interview, the presence of another adult during the interview, and survey mode. In none of these models was survey mode a significant predictor. Interestingly, the presence of another adult during the interview was associated with significantly higher reported activity times, consistent with social desirability arguments.

⁶ See *Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) – Short and Long Forms*, November 2005: at <http://www.ipaq.ki.se/scoring.pdf>.

accept that it is possible that some large round numbers (e.g., 10, 15 or 20) may have been reported in minutes but mistakenly coded as hours.

Derived Variables

We constructed, for each of the three activity types, a measure of daily time, with all responses converted to minutes. Additionally, for any respondent who provided an answer for an entire week (because their activity pattern was too irregular to provide an answer for a typical day), a daily equivalent was obtained by dividing by the reported number of days on which that activity occurred.⁷

In an attempt to obtain a more normal distribution, the IPAQ guidelines also specify that each of the three daily activity time variables are truncated so that any values exceeding 180 minutes are recoded to be equal to 180 minutes, meaning that the total time spent on any one activity in a week cannot exceed 21 hours. This affected a relatively large fraction of responses – 8.7, 9.6 and 6.4 per cent of all cases with respect to walking, moderate activity and vigorous activity time.

From this we then produced eight measures of weekly activity time:

- (i) Walking time per week = Daily walking minutes x Walking days;
- (ii) Moderate activity time per week = Daily moderate activity minutes x Moderate activity days;
- (iii) Vigorous activity time per week = Daily vigorous activity minutes x Vigorous activity days;
- (iv) Total physical activity time per week = Walking time + Moderate activity time + Vigorous activity time.
- (v) Walking MET-minutes per week = 3.3 x Walking time per week;
- (vi) Moderate activity MET-minutes per week = 4.0 x Moderate activity time per week;
- (vii) Vigorous activity MET-minutes per week = 8.0 x Vigorous activity time per week;
- and
- (viii) Total physical activity MET-minutes per week = Walking MET-minutes + Moderate MET-minutes + Vigorous MET-minutes.

The selected MET values of 3.3, 4.0 and 8.0 are as recommended by IPAQ, which, in turn, are based on the earlier work of Craig et al. (2003).⁸

Finally, a summary categorical variable is produced, with every valid responding case assigned to one of three categories, as follows:

1. High

Vigorous activity on at least 3 days, achieving a minimum total physical activity of at least 1500 MET-minutes per week,

OR

7 or more days of any combination of the three activity types, achieving a minimum total physical activity level of at least 3000 MET-minutes per week.

⁷ The IPAQ guidelines recommends dividing by seven, which makes no sense given the underlying variable is the time spent on an activity on a day when that activity is undertaken.

⁸ These are only estimates and hence other values could be used. The Australian Health Survey, for example, assigns values of 3.5 to walking for fitness, 5 to moderate intensity activity and 7.5 to vigorous intensity activity.

2. Moderate

3 or more days of vigorous activity of at least 20 minutes per day,

OR

5 or more days of moderate activity and / or walking of at least 30 minutes per day

OR

5 or more days of any combination of the three activity types, achieving a minimum total physical activity level of at least 600 MET-minutes per week.

3. Low

Any individual who does not meet any of the criteria recommended above.

Obviously the data can also be used to construct other measures. That said, what is not so easily assessed with these data is whether activity levels meet the National Physical Activity Guidelines.⁹ For 18 to 64 year olds, the guidelines specify that people should: (i) be active on most days of the week (which has generally been interpreted as meaning at least 5 days); (ii) accumulate a minimum of 150 minutes of moderate intensity exercise or 75 minutes of vigorous activity (or an equivalent combination); and (iii) muscle strengthening on at least two days a week. No data on the latter requirement is collected in the IPAQ. But more importantly, the walking time data collected in the IPAQ does not identify whether that activity meets the moderate intensity level. In the national guidelines this is defined to be any activity that “takes some effort, but you are still able to talk while doing them”, with “brisk walking” provided as an explicit example.

Population Estimates

Summary Statistics

We next present population weighted estimates of the prevalence of physical activity derived from wave 13 of the HILDA Survey.¹⁰ We begin, in Table 1, by presenting estimates of the distribution of time spent on each of the three types of activities for the entire sampled population (persons aged 15 years or older), as well as the median, mean and standard deviation for each activity type.

As can be seen, the distribution of time spent on each activity is highly skewed. This is especially the case with vigorous and moderate physical activity, where the proportions inactive are 53% and 41% respectively. The distribution of walking time, on the other hand, is far less skewed. And when time is summed across all three activities the distribution obtained is much closer to ‘normal’ looking.

As we might expect, both median and mean activity levels decline with the intensity of the activity. Median activity time per week declines from two hours and 20 minutes for walking, to one hour for moderate physical activity, to zero for vigorous activity. Average levels are much higher, reflecting the long tails in the time distributions, and range from over 4 hours per week for walking to a little over 2 hours for vigorous physical activity.

⁹ As recommended by the Australian Government Department of Health (and available at: <http://www.health.gov.au/internet/main/publishing.nsf/content/health-publth-strateg-phys-act-guidelines>).

¹⁰ The weighting factors used here come from those used in a beta-version of release 13. There may, therefore, be small differences between the estimates reported here and equivalent estimates derived from the final release 13 dataset.

Table 1: Distribution of estimated time spent on physical activity last week (persons aged 15+)

	<i>Vigorous activity</i>	<i>Moderate activity</i>	<i>Walking</i>	<i>Total</i>
% distribution				
Zero	53.4	40.8	15.9	6.7
< 2.5 hours	20.7	28.1	36.9	18.3
2.5 to < 5 hours	11.6	12.0	19.9	20.1
5 to < 10 hours	8.5	9.3	13.8	22.9
10 to < 20 hours	5.0	7.6	8.2	17.8
20 to < 30 hours	0.9	2.3	5.3	8.8
30 hours or more	0.0	0.0	0.0	5.4
Median minutes	0	60	140	360
Mean minutes	126.7	175.8	264.2	565.2
Standard deviation	233.3	286.1	333.2	603.9

Note: Estimates are population weighted.

Comparisons with the AHS

While the distributions presented in Table 1 seem reasonable, the mean levels are relatively high. As previously noted, overstatement by some (if not many) respondents is to be expected given the social desirability associated with being seen to be physically active. Nevertheless, the mean levels of physical activity reported in the HILDA Survey are also much higher than that reported in other self-report surveys, and more specifically the 2011-2012 Australian Health Survey (AHS).

A detailed comparison is provided in Table 2, which reports data from both the AHS and the HILDA Survey on mean time spent on each of the three activity types, by age group, for the population aged 18 years and over. As I hope has already been made clear, the two data sources are not directly comparable. The estimates of vigorous and moderate activity from the AHS reported in Table 2 are restricted to leisure-time activities, while the estimate of walking time is restricted to walking for fitness, recreation and sport and walking to get to and from places. We thus expect larger estimates in the HILDA Survey.¹¹ Nevertheless, are the differences we find too large? The estimated mean time spent on vigorous activity is about twice as large in the HILDA Survey as in the AHS; time on moderate physical activity is more than six times greater; and walking time about 80% higher. In terms of total activity time, the mean activity time in the HILDA Survey (551 minutes) is about 2.4 times higher than in the AHS.

A similar (indeed slightly larger) differential was reported by Brown et al. (2004) when comparing responses to the IPAQ with responses to the Active Australia (AA) Survey among a small sample of Australians interviewed by telephone using both sets of instruments. Brown et al. explained this differential as a function of two main factors. First, and as already noted, the IPAQ items cover, in theory, physical activity that occurs in all life domains, whereas the

¹¹ On the other hand, the timing of the two surveys might lead us to expect lower activity levels within the HILDA Survey sample. As noted earlier, the HILDA Survey interviews were concentrated in two months (August and September) that, in the southern states at least, are relatively cool. In contrast, interviewing for the two surveys that comprise the AHS was spread over all 12 months of the year.

AA Survey is mainly restricted to activity during leisure time.¹² Second, the differential may reflect differences in the way the questions are posed. Notably, the AA instrument (like the AHS) asks respondents to estimate activity time over the entire 7-day recall period, whereas the IPAQ only seeks an estimate for one day, with a weekly estimate obtained by multiplying that estimate by the number of days in the week the activity is undertaken. The IPAQ approach will produce higher estimates if respondents have a tendency to select the day during the week that they spent most time on that activity.¹³

Table 2: Average time (mins) spent on physical activity last week by age: HILDA Survey (wave 13) and Australian Health Survey (2001-12) compared (persons aged 18+)

<i>Age group</i>	<i>Vigorous activity</i>	<i>Moderate activity</i>	<i>Walking</i>	<i>Total</i>
HILDA, 2013				
18-24	184	173	298	653
25-34	158	182	293	630
35-44	130	173	266	565
45-54	121	173	265	558
55-64	102	176	234	512
65-74	70	179	218	466
75+	39	138	153	325
Total	123	173	256	551
AHS, 2011-2012				
18-24	114	29	134	276
25-34	75	22	140	236
35-44	59	19	142	219
45-54	60	23	157	239
55-64	32	37	158	226
65-74	21	41	153	214
75+	7	25	109	142
Total	58	27	144	228

Notes: All estimates are population weighted.

Walking time in the AHS is the sum of walking for fitness, recreation or sport and walking for transport.

Sources: The AHS data come from Australian Bureau of Statistics, *Australian Health Survey: Physical Activity, 2011-2012* (ABS cat. no. 4364.0.55.004), Data cubes, Table 1.1.

But if the differences in the way the questions are asked are a major contributor to this cross-survey differential, we might expect the differential to be similar across all three activity types. This is not the case. Indeed, the AA Survey does include an additional item about “vigorous gardening or heavy work around the yard” which was also administered in the AHS. If this estimate is added to the AHS estimate of vigorous physical activity we arrive at

¹² As noted below, the AA Survey also includes a question on “vigorous gardening or heavy work around the yard”. Further, the recommended item on walking covers not only “recreation” and “exercise”, but also getting “to or from places”.

¹³ They also argue that use of a ‘usual week’ reference period, rather than the last 7-days (as used in AA), in the version of the IPAQ they administered may have contributed to higher estimates. This, however, is not relevant for this comparison given that subsequent versions of IPAQ switched to the last 7-day recall method.

a revised mean time for vigorous physical activity of 116 minutes, only slightly less than the estimated mean of 123 minutes in the HILDA Survey.

These observations suggest that the large differences in moderate physical activity time, as well as in walking time, are being driven in large part by the differences in activity coverage, and especially the exclusion of activity that occurs during working time from the AA Survey (and hence from the AHS). The inclusion of working time would be expected to mainly affect moderate physical activity and walking. It would obviously include physical work by labourers (and other workers in blue-collar jobs), but many employees in the service industry (e.g., wait staff) will also spend considerable periods of time walking. It might be argued that some of these work-related activities do not involve activity that is continuous for at least 10 minutes, but I doubt that is how the average respondent thinks of physical activity. Activities such as golf and gym work-outs, for example, all involve frequent interruptions, yet these are all regarded as perfectly acceptable forms of physical activity.

That said, if the treatment of activity during working time was the principal source of differences in estimates between the two surveys, then surely the differential in estimates would be much smaller among the oldest members of the population (where paid employment is relatively uncommon)? Table 2 shows that the differentials in the estimates of both moderate activity and walking time are smallest for the people in the oldest age groups, but these differentials are still very large.

Overall, I am drawn to two conclusions. First, comparisons with the AHS probably tell us very little about the quality of the physical activity data being collected in wave 13 of the HILDA Survey. There are large differences, but then that is what we might expect given the differences in the range of activity covered by the two surveys. Second, there nevertheless remain concerns that the use of the IPAQ is associated with a significant overstatement of physical activity, especially of moderate physical activity and walking.

Participation in Sufficient Physical Activity

As previously noted, the data provided by the IPAQ do not enable the construction of a measure that identifies persons who are active at levels that perfectly align with the National Physical Activity Guidelines, and neither do the AA Survey items. The ABS (2013), however, constructs a measure that is very close, which they describe as “sufficient physical activity”. In this measure, an individual is sufficiently active during the reference week if they undertook at least five sessions of physical activity involving a total of at least 150 minutes of moderate physical activity or walking (for transport or fitness reasons) or 75 minutes of vigorous physical activity (or an equivalent combination of both). For comparative purposes we construct two alternative measures of ‘sufficient activity’ from the HILDA Survey. They are both defined as above (but bearing in mind the differences in activity scope) but with one measure including walking and the other excluding it. Summary results disaggregated by age are reported in Table 3.

As can be seen, the treatment of walking is critical. If walking is excluded then a little less than one in three Australians would meet the sufficient activity threshold. Once walking time is included this rises to 62%. By comparison, the estimate from the AHS falls between these two extremes – 43%. This is as expected given the AHS definition is restricted to walking for transport and fitness / recreation purposes, and other physical activity is restricted to activity during leisure time.

Table 3: Proportion of persons ‘sufficiently active’ last week by age: HILDA Survey (wave 13) and Australian Health Survey (2011-12) compared (persons aged 18+)

<i>Age group</i>	<i>Australian Health Survey</i>	<i>HILDA Survey, (excl. walking)</i>	<i>HILDA Survey, (incl. walking)</i>
18-24	53.4	43.7	71.8
25-34	47.0	39.0	70.0
35-44	44.0	33.3	61.8
45-54	44.0	32.1	61.4
55-64	38.9	28.6	58.8
65-74	37.7	24.4	56.4
75+	25.2	17.9	42.5
Total	43.0	32.7	62.1

Note: All estimates are population weighted.

Sources: The AHS data come from Australian Bureau of Statistics, *Australian Health Survey: Physical Activity, 2011-2012* (ABS cat. no. 4364.0.55.004), Data cubes, Table 1.1.

IPAQ categories

As previously noted, the IPAQ also recommends the construction of a categorical variable that divides the population into three broad groups based on the overall physical activity level of each individual. Table 4 presents a summary of this distribution cross-classified by sex and broad age group. According to this categorisation system, close to one in three Australians have low activity levels (27% of men and 36% of women), and a similar proportion overall (35%) have high activity levels (though prevalence varies markedly with sex: 42% of men but only 28% of women).

Table 4: Prevalence of IPAQ categories by age and sex (persons aged 18+)

<i>Sex / Age group</i>	<i>Low (%)</i>	<i>Moderate (%)</i>	<i>High (%)</i>	<i>Total (%)</i>
Male				
18-39	19.9	28.3	51.9	100.0
40-64	29.4	31.0	39.6	100.0
65+	38.2	36.0	25.7	100.0
Sub-total	27.1	30.8	42.2	100.0
Female				
18-39	28.5	36.1	35.4	100.0
40-64	37.0	36.2	26.8	100.0
65+	50.2	34.8	15.0	100.0
Sub-total	36.1	35.9	28.0	100.0

Note: Estimates are population weighted.

In the only other study that I am aware of that has administered the IPAQ data to a nationally representative Australian population sample and reported data by these categories, the reported rates of ‘high activity’ are actually much larger (66% of men, and 52% of women,

aged between 18 and 64 in June 2003) (Bauman et al. 2009).¹⁴ The high prevalence of high-activity individuals reported for Australia in that study is difficult to believe.¹⁵ While entirely speculative, it might be that the social desirability bias associated with self-reporting of physical activity could be greater in a survey where physical activity is the focal point. Selection effects might also be at work, with a survey on physical activity likely to attract greater interest from person who themselves are very active. The HILDA Survey is not immune to reporting biases, but there is little reason to think that participation in the HILDA Survey, which is a broad socio-economic survey, is strongly related to physical activity levels except within the extreme tail of the distribution where people are inactive because of severe illness and disability, which in turn will inhibit survey participation.^{16, 17}

Seasonality

As mentioned earlier, physical activity levels might be expected to vary with the time of the year in which the survey takes place. There is some weak evidence for this in Table 5. Across all persons, mean MET minutes of physical activity are lowest in July / August, when days are shortest and average temperatures are coolest, and highest in the November to February period when days are longer and temperatures warmer, though the magnitude of this differential is modest – 230 minutes, or a little less than 10% of total weekly activity averaged across the entire sample.¹⁸ The differences between the other periods (the months of September, October and the November to February period), however, are small and statistically insignificant. Further, the patterns are inconsistent across the different States and Territories of Australia, with the level of activity in July / August being only significantly lower than at other times of the year in South Australia and Queensland (which has been combined with Northern Territory given the relatively small number of cases in the latter).¹⁹

We also checked, and confirmed, that similar results were obtained using categorical measures; e.g., the proportion of low-activity persons.

At face value, therefore, it seems that seasonality in physical activity is not a big issue for the HILDA Survey data.

¹⁴ The equivalent proportions from the HILDA Survey are 45.6% of men and 31% of women aged between 18 and 64 years.

¹⁵ Similarly large estimates of “high activity” people are reported for samples from other developed countries, including Canada, New Zealand and the USA. Nevertheless, a key feature of the results presented by Bauman et al. (2009) is the large variation across their 20 country samples. The estimates reported for some other countries – notably Norway, Spain and Sweden – are very similar to those reported here from the HILDA Survey.

¹⁶ Attrition within the HILDA Survey has been shown to be significantly related to the presence of serious long-term health conditions (Watson and Wooden 2009). Nevertheless, such affects should be adjusted for through the use of the population weights that are provided in the data set. What will not be well adjusted for is non-response due to illness and disability at the initial survey wave.

¹⁷ Differences in the timing of the surveys might also provide a small part of the explanation, with the Australian sample in Bauman et al. (2009) surveyed in April 2003, a month which included three public holidays, whereas as already noted, the HILDA Survey interviews were mainly conducted in August or September (of 2013).

¹⁸ Further, in a simple regression model which controlled for individual characteristics, physical activity time was not significantly less among persons interviewed in July / August than in most other months.

¹⁹ Simple independent samples t-tests were conducted using weights which adjust proportions but leave the total sample size unchanged.

Table 5: Mean MET minutes by month of interview and State

<i>State</i>	<i>Month of interview</i>				<i>Total</i>
	<i>July/ August</i>	<i>September</i>	<i>October</i>	<i>November- February</i>	
NSW / ACT	2340	2376	2488	2444	2375
Victoria	2377	2332	2095	2305	2316
SA	2126	2832	2932	2780	2628
Tasmania	2490	2357	2136	2267	2420
WA	2692	2419	3235	3126	2666
Qld / NT	2719	3084	3054	3132	2942
Australia	2444	2561	2593	2674	2528

Note: Estimates are population weighted.

Construct Validity

As previously discussed, there is some debate about how well the IPAQ performs as a measure of physical activity. The HILDA Survey data is unable to directly inform this debate. We can, however, examine associations between the physical activity measures and other variables collected in the HILDA Survey that might be expected to be associated with physical activity levels. Most obvious here are measures of physical health, including the physical functioning scale within the SF-36, body mass index (BMI) and girth (waist to height ratio). I therefore report, in Table 5, rank-order correlations between four summary measures of physical activity time from the IPAQ (as well as the single-item activity measure included every year in the SCQ) and a series of indicators of health-related outcomes and behaviours.

There at least three key features of the results presented in this table worth noting.

First, with one exception, all correlations are in the expected direction; higher physical activity levels are associated with better health outcomes and behaviours. The exception here is smoking, with small positive correlations between a current smoker vs non-smoker dummy and each of the four IPAQ measures found. This surprising finding might reflect associations with age, with smoking rates (while declining over time) still highest among younger members of the population (Wilkins 2013). However, the within age-group correlations still show positive correlations for persons in the 25 to 54 year age range.

Second, the magnitudes of these correlations are arguably not very large. The size of the correlation between total MET minutes per week and the physical functioning sub-scale of the SF36 (which is intended to measure the extent to which poor health limits the ability to undertake a range of common daily activities), for example, is just .26. Nevertheless, relatively small correlations should not be surprising given the relatively modest correlations between self-reported physical activity and objective measures of activity discussed earlier.

Table 5: Spearman rank correlation coefficients

	<i>IPAQ measures</i>				<i>SCQ</i>
	<i>Walking (mins / week)</i>	<i>Moderate PA (mins / week)</i>	<i>Vigorous PA (mins / week)</i>	<i>Total PA (MET mins per week)</i>	<i>Moderate / intensive PA frequency</i>
SF36 health status scales					
Physical functioning	.133**	.116**	.330**	.263**	.285**
Role – physical	.135**	.110**	.220**	.216**	.210**
Bodily pain	.096**	.062**	.182**	.148**	.175**
General health	.127**	.121**	.248**	.232**	.304**
Vitality	.114**	.130**	.191**	.205**	.289**
Social functioning	.104**	.108**	.163**	.177**	.206**
Role – emotional	.079**	.095**	.114**	.140**	.158**
Mental health	.061**	.088**	.082**	.111**	.177**
Presence of LT health condition / disability					
	-.108**	-.091**	-.204**	-.182**	-.166**
BMI (kg/m ²)	-.073**	-.031**	-.121**	-.099**	-.146**
Waist-height ratio	-.104**	-.066**	-.219**	-.179**	-.207**
Obese (BMI > 30)	-.070**	-.045**	-.125**	-.106**	-.147**
Smoker	.021*	.029**	.010	.052**	-.042**

Notes: ** and * denote $p < .001$ and $p < .01$, respectively, in a 2-tailed test.

Third, correlations with the single-item measure included every year in the SCQ (and reported in the final column of Table 4) are no smaller.²⁰ Indeed, if anything, associations with this variable are mostly larger than those observed for any of the IPAQ variables. This would seem to suggest that the single-item SCQ measure performs just as well as, if not better than, the various IPAQ measures. One potential criticism of this conclusion is that the comparison made here is not strictly valid given the SCQ item is a measure of frequency while the summary IPAQ variables are all measures of time. Nevertheless, comparison with the measures of frequency that are collected in the IPAQ continue to suggest that as a measure of the frequency of participation in moderate and vigorous physical activity, the SCQ item performs at least as equally well as the IPAQ. Indeed, a measure of the frequency of moderate and vigorous physical activity derived from the IPAQ tends to produce slightly smaller correlations.²¹ One possible explanation for the superior performance of the SCQ variable is that it is a measure of usual activity, rather than activity in the last 7-days, and of course it is usual behaviour that is most relevant for health outcomes.

²⁰ As noted earlier, this variable measures the frequency of participation in moderate or vigorous physical activity in a usual week. Responses have been re-coded to provide a continuous variable. It can take the values 0, 0.5, 1.5, 3, 5 and 7.

²¹ Correlations are generally strongest for the IPAQ measure of weekly frequency of vigorous activity.

Summary and Conclusions

Wave 13 of the HILDA Survey saw the administration of the short-form of the IPAQ. The data collected suggest that, on average, Australians in (the latter half of) 2013 spent about two hours a week engaged in vigorous physical activity, almost three hours in moderate activity, and over four hours a week walking. Nevertheless, there are sizeable proportions of Australians who reported undertaking either no activity or very little activity. Indeed, based on categories recommended by IPAQ, the data suggest that about 27% of men and 36% of women have low activity levels.

These estimates of activity levels are, however, much higher than reported in the Australian Health Survey. In part, the large differences simply reflect the difference in the coverage of different measures. Nevertheless, it is difficult to avoid drawing the conclusion that levels of physical activity, and especially moderate intensity activity and walking, are overstated in the HILDA Survey. This might occur, for example, if respondents with irregular activity patterns, when asked to think of a day last week when they were physically active, have a tendency to select the day on which they were active for the longest.

On the other hand, comparisons with another survey of an Australian sample using the IPAQ instrument revealed much lower estimates of physical activity in the HILDA Survey. If the IPAQ has a tendency to cause activity levels to be overstated, the extent of that overstatement is much less in the HILDA Survey.

We also suspect that much of the difference in estimated mean times is being driven by the cases at the upper end of the distribution. It thus may be that throughout much of the distribution, differences are far less pronounced.²²

Finally, evidence was also presented suggesting that the single-item measure of physical activity that has been included in every wave of the HILDA Survey to date, while not providing an estimate of the actual time spent on physical activity, performs just as well as the IPAQ measures when it comes to predicting health outcomes and behaviours. This is good news for those HILDA Survey data users who wish to analyse relationships with physical exercise over the duration of the panel.²³

References

Australian Bureau of Statistics (ABS). 2013. *Australian Health Survey: Users' Guide, 2011-2013* (ABS cat. no. 4363.0.55.001). ABS: Canberra.

Australian Institute of Health and Welfare (AIHW). 2003. *The Active Australia Survey: A Guide and Manual for Implementation, Analysis and Reporting*. AIHW: Canberra.

Bauman A, Bull F, Chey T, Craig CL, Ainsworth BE, Sallis JF, Bowles HR, Hagströmer M, Sjöström M, Pratt M, and the IPS Group. 2009. The International Prevalence Study on Physical Activity: Results from 20 Countries. *International Journal of Behavioral Nutrition and Physical Activity* 6: Article 21.

Boon RM, Hamlin MJ, Steel GD and Ross JJ. 2010. Validation of the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF) with accelerometry. *British Journal of Sports Medicine* 44(10): 741-746.

²² Testing this directly, however, requires accessing the unit record data from the AHS.

²³ Examples of such uses include Brown and Roberts (2011) and Perales, del Pozo-Cruz & del Pozo-Cruz (2014a, 2014b).

- Brown W, Bauman A, Chey T, Trost S and Mummery K. 2004. Comparison of surveys used to measure physical activity. *Australian and New Zealand Journal of Public Health* 28(2): 128-134.
- Brown H and Roberts J. 2011. Exercising choice: the economic determinants of physical activity behaviour of an employed population. *Social Science and Medicine* 73(3): 383-390.
- Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF and Oja P. 2003. International Physical Activity Questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise* 35(8): 1381-1395.
- Hendelman D, Miller K, Baggett C, Debold E & Freedson P. 2000. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine & Science in Sports & Exercise* 32(9 Suppl): S442-449.
- Lee PH, Macfarlane DJ, Lam TH and Stewart SM. 2011. Validity of the international physical activity questionnaire short form (IPAQ-SF): a systematic review. *International Journal of Behavioural Nutrition and Physical Activity* 8: Article 115.
- Perales F, del Pozo-Cruz J and del Pojo-Cruz B. 2014a. On the associations between physical activity and quality of life: findings from an Australian nationally representative panel survey. *Quality of Life Research* 23(7):1921-1933.
- Perales F, del Pozo-Cruz J and del Pojo-Cruz B. 2014b. Impact of physical activity on psychological distress: prospective analysis of an Australian national sample. *American Journal of Public Health* 104(12): e91–e97.
- Sjöström M, Oja P, Hagströmer M, Smith BJ and Bauman A. 2006. Health-enhancing physical activity across European Union countries: the Eurobarometer study. *Journal of Public Health* 14(5): 291-300.
- Tucker P and Gilliland J. 2007. The effect of season and weather on physical activity: a systematic review. *Public Health* 121(12): 909-922.
- Watson N and Wooden M. 2009. Identifying factors affecting longitudinal survey response. In P Lynn (ed.), *Methodology of Longitudinal Surveys* (pp. 157-182). Wiley: Chichester (UK).
- Wilkins R. 2013. Who stopped smoking between 2001 and 2010? In R Wilkins (ed.), *Families, Incomes and Jobs, Volume 8* (pp. 102-105). Melbourne: Melbourne Institute of Applied Economic and Social Research.

Appendix A: Paper representation of physical activity question sequence used in wave 13, HILDA Survey

Physical Activity

K55. I am now going to ask you about the time you spent being physically active in the last 7 days. Please answer each question even if you do not consider yourself to be an active person. Think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

I will first ask about vigorous activities, then moderate activities, and finally walking.

So think now about all the vigorous activities which take hard physical effort that you did in the last 7 days. Vigorous activities make you breathe much harder than normal. Examples may include heavy lifting, digging, jogging, aerobics, or fast cycling. Think only about those physical activities that you did for at least 10 minutes at a time.

During the last 7 days, on how many days did you do vigorous physical activities?

(IF NECESSARY, REPEAT: Think only about those physical activities that you did for at least 10 minutes at a time.)

- | | | |
|--------------------------|--------------------------------|-------|
| <input type="checkbox"/> | Record number of days per week | [1] |
| <input type="radio"/> | Refused | [998] |
| <input type="radio"/> | Don't know | [999] |

IF K55.number >0 go to K56
ELSE go to K57

K56a. How much time did you usually spend doing vigorous physical activities on one of those days?

(IF NECESSARY, REPEAT: Think only about those physical activities that you did for at least 10 minutes at a time.)

- | | | |
|--------------------------|-------------------------------------|-------|
| <input type="checkbox"/> | Record number of hours | [1] |
| <input type="checkbox"/> | Record number of minutes | [2] |
| <input type="radio"/> | Refused | [998] |
| <input type="radio"/> | Don't know – varies from day to day | [997] |
| <input type="radio"/> | Don't know | [999] |

IF K56a = 997 go to K56b
IF K56a > 3 hrs / 180 mins go to K56c
ELSE go to K57

K56b. How much time in total did you spend over the last 7 days doing vigorous physical activities?

- | | | |
|--------------------------|--------------------------|-------|
| <input type="checkbox"/> | Record number of hours | [1] |
| <input type="checkbox"/> | Record number of minutes | [2] |
| <input type="radio"/> | Refused | [998] |
| <input type="radio"/> | Don't know | [999] |

Go to K57

**K56c. CONFIRM ANSWER IS FOR A SINGLE DAY. IF NECESSARY, ASK:
And can I confirm that your answer is the time spent on just one day?**

- Yes – one day [1]
- No – 7 days [2]
- No – other (specify number of days) [3]

K57. Now think about activities which take moderate physical effort that you did in the last 7 days. Moderate physical activities make you breathe somewhat harder than normal. Examples may include carrying light loads, gentle swimming, cycling at a modest pace, or social tennis. Do not include walking. Again, think about only those physical activities that you did for at least 10 minutes at a time.

During the last 7 days, on how many days did you do moderate physical activities? Do not include walking.

(IF NECESSARY, REPEAT: Think only about those physical activities that you did for at least 10 minutes at a time.)

- Record number of days per week [1]
- Refused [998]
- Don't Know [999]

IF K57.number >0 go to K58
ELSE go to K59

K58a. How much time did you usually spend doing moderate physical activities on one of those days?

(IF NECESSARY, REPEAT: Think only about those physical activities that you did for at least 10 minutes at a time.)

- Record number of hours [1]
- Record number of minutes [2]
- Refused [998]
- Don't know – varies from day to day [997]
- Don't know [999]

IF K58a = 997 go to K58b
IF K58a > 3 hrs / 180 mins go to K58c
ELSE go to K59

K58b. How much time in total did you spend over the last 7 days doing moderate physical activities?

- Record number of hours [1]
- Record number of minutes [2]
- Refused [998]
- Don't know [999]

Go to K59

K58c. CONFIRM ANSWER IS FOR A SINGLE DAY. IF NECESSARY, ASK:
And can I confirm that your answer is the time spent on just one day?

- Yes – one day [1]
- No – 7 days [2]
- No – other (specify number of days) [3]

K59. Now think about the time you spent walking in the last 7 days. This includes at work and at home, walking from place to place, and any other walking that you might do solely for recreation, sport or exercise.

During the last 7 days, on how many days did you walk for at least 10 minutes at a time?

(IF NECESSARY, REPEAT: Think only about those physical activities that you did for at least 10 minutes at a time.)

- Record number of days per week [1]
- Refused [998]
- Don't Know [999]

IF K59.number >0 go to K60
ELSE go to K61

K60a. How much time did you usually spend walking on one of those days?

(IF NECESSARY, REPEAT: Think only about walking that you did for at least 10 minutes at a time.)

- Record number of hours [1]
- Record number of minutes [2]
- Refused [998]
- Don't know – varies from day to day [997]
- Don't know [999]

IF K60a = 997 go to K60b
IF K60a > 3 hrs / 180 mins go to K60c
ELSE go to K61

K60b. How much time in total did you spend walking over the last 7 days?

- Record number of hours [1]
- Record number of minutes [2]
- Refused [998]
- Don't know [999]

Go to K61

K60c. CONFIRM ANSWER IS FOR A SINGLE DAY. IF NECESSARY, ASK:

And can I confirm that your answer is the time spent on just one day?

- Yes – one day [1]
- No – 7 days [2]
- No – other (specify number of days) [3]

K61. Now think about all your physical activity during the last 7 days. Would this be more than usual, less than usual, or about the same as you would usually do in a 7-day period?

- More than usual [1]
- About the same as usual [2]
- Less than usual [3]
- Refused [8]
- Don't Know [9]