How do older adults respond to active australia physical activity questions? lessons from cognitive interviews

Heesch, Kristiann C.; Van Uffelen, Jannique; Brown, Wendy J.

Published in:
Journal of Aging and Physical Activity

DOI:
10.1123/JAPA.2012-0175

Licence:
Other

How Do Older Adults Respond to Active Australia Physical Activity Questions? Lessons from Cognitive Interviews

Abstract

The aim of this study was to examine older adults' understanding and interpretation of a validated questionnaire for physical activity surveillance, the Active Australia Survey (AAS). To address this aim, cognitive interviewing techniques were used during face-to-face semi-structured interviews with 44 adults aged 65-89 years. Qualitative data analysis revealed that participants were confused with questionnaire phrasing, misunderstood the scope of activities to include in answers, and misunderstood the time frame of activities to report. They also struggled to accurately estimate the frequency and duration of their activities. Our findings suggest that AAS questions may be interpreted differently by older adults than intended by survey developers. Findings also suggest that older adults use a range of methods for calculating PA frequency and duration. The issues revealed in this study may be useful for adapting AAS for use in older community-dwelling adults.

Keywords: physical activity, exercise, qualitative methods, interviews, questionnaire, survey, aged, elderly

Evidence on the beneficial effects of physical activity (PA) in older adults (Chodzko-Zajko et al., 2009; Nelson et al., 2007) is largely based on epidemiological studies, in which questionnaires have, in light of their ease of administration and relatively low cost, been the preferred method of assessment (Bauman, Phongsavan, Schoeppe, & Owen, 2006; Lagerros & Lagiou, 2007). In Australia, the Active Australia Survey ([AAS]; Australian Institute of Health and Welfare, 2003) is widely used for federal (Armstrong, Bauman, & Davies, 2000) and state-specific physical activity surveillance (e.g., Chau, Smith, Chey, Merom, & Bauman, 2007; Tomlin, Joyce, & Patterson, 2012), as well as in population-wide intervention research (Marshall et al., 2004; Merom et al., 2004) . Because of its relatively strong measurement properties, including good validity and reliability (Brown, Burton, Marshall, & Miller, 2008; Brown, Trost, Bauman, Mummery, & Owen, 2004; Timperio, Salmon, & Crawford, 2003) and its responsiveness to change (Reeves et al., 2010), AAS is increasingly being used in other countries. In line with a growing focus on the importance of PA for older people, AAS is also increasingly being used to assess PA levels in this population group. However, little is known about how older people cope with the recall and formulation aspects of responding to questions about a behaviour as complex as PA.

Our aim was therefore to examine older people's understanding and interpretation of the questions asked in AAS, in order to uncover cognitive problems that they may have with responding to PA questions. We used cognitive interviewing methods, as use of these methods has uncovered cognitive problems during the development of questionnaires in other fields (Collins, 2003; Drennan, 2003) and during the refinement and development of other PA measures (Fisher, Falkner, Trevisan, & McCauley, 2000; McKenna, Foster, & Page, 2004; Nigg et al., 2005; Topolski et al., 2006).

Method

Participants

As reported previously [redacted], we recruited 55 community-dwelling residents of Brisbane, Australia, aged ≥65 years. We purposively sampled to include men and women with varying education and PA levels given that these factors influence understanding of PA questionnaires (Durante & Ainsworth, 1996). To recruit older adults with different levels of PA and education levels, flyers were displayed at a variety of voluntary organisations with large numbers of older adult members and located in different suburbs of the city (e.g., bridge clubs, senior centres, and exercise centres for older adults), and emails and notices in e-newsletters were sent to university staff to encourage them to pass along the study details to eligible friends and relatives. Potential participants were initially screened by telephone. Those who reported an inability to walk >100 meters without aid, were unable to speak and understand English, or were unable to cognitively respond to our questioning, were ineligible. The study protocol was approved by the [redacted] Ethics Committee.

Cognitive Interviewing

The theoretical framework underpinning the interview process was the Conrad and Blair question-and-answer model of survey response (Conrad & Blair, 1996), a revised version of the Tourangeau question-and-answer model (Tourangeau, 1984). In accordance with this model, we sought to understand participants' progression through three stages of responding to a question: 1) comprehension (understanding the meaning of the words used and the information being requested); 2) response formulation (information retrieval and mental calculations required to develop responses); and 3) response mapping (fitting answers into pre-specified response

formats). To document participants' progress through these stages, two cognitive interviewing techniques were used. The first was concurrent think-aloud, and the second was scripted and unscripted probing (Beatty & Willis, 2007; Drennan, 2003; Jobe & Mingay, 1989). The interview protocol was pre-tested in three practice interviews and refined after each of these interviews.

Interview Protocol

Each participant provided informed consent and completed a questionnaire on socio-demographic and health-related characteristics before the start of a face-to-face interview with two interviewers. One interviewer [redacted] led the interview while another [redacted] (or a research assistant) audio-recorded, noted non-verbal communication and asked probing questions.

During the interview, the participant completed four short PA surveillance questionnaires as part of a larger study on measurement of PA in older adults. In addition to AAS, the participant responded to the Behavioral Risk Factor Surveillance System (Centers for Disease Control and Prevention, 2003), the PA Scale for the Elderly (Washburn, Smith, Jette, & Janney, 1993), and the International PA Questionnaire (Craig et al., 2003). These questionnaires were included in the larger study because they are the most frequently used PA surveillance questionnaires and they have been used with older adults. To decrease possible bias from questionnaire ordering, the order of questionnaire administration was determined by a computer-based random order generator, and the participant was instructed before the start of each new questionnaire after the first one, to respond as if the person had not already completed other questionnaires. Interviews concluded with the participant receiving a $20 gift voucher.

For the current analysis, only AAS questions were examined, and thus the specific

interview protocol used for gathering data about AAS is described next. Each participant was asked to practice the cognitive interviewing think-aloud technique and then to use the technique in responding to AAS questions (See Appendix for AAS questions). The participant was asked to read each AAS question aloud, verbalize the process of formulating an answer, and write the answer on a printed copy of the questionnaire. Thus, although interviews were held in a face-to-face setting, the process of completing AAS replicated a survey context as much as possible, in that each participant completed a pen and paper version of AAS as part of the interview process. If a participant did not adequately verbalize the process used for formulating an answer, the interviewers asked probing questions to learn whether the participant understood the intent of the question (e.g., *What activities are you including in your answer?*) and how the person had formulated a response (e.g., *How did you come up with this answer?*). The interviewers also noted any difficulties in recording responses on the printed copy of the questionnaire.

Active Australia Physical Activity Survey

AAS is an 8-item measure developed for PA surveillance of adults aged 18 to 75 years (Australian Institute of Health and Welfare, 2003): see Appendix. It assesses four PA domains in the following order: walking (*for recreation, exercise or to get to and from places*), vigorous gardening and heavy work around the yard (*which made you breather harder or puff and pant*), vigorous PA (*that makes you breathe harder or puff and pant, [e.g., jogging, cycling, aerobics, competitive tennis]*), and moderate PA (*e.g., gentle swimming, social tennis, golf*). For each domain, participants reported the number of times the activity was performed in the previous week, as a measure of frequency, and the total time (hours and minutes) spent in the activity in the last week, as a measure of duration. The vigorous gardening and yardwork questions are included after the walking questions to allow participants to report such activities but to dissuade

them from reporting these activities in questions about vigorous PA that follow. Reponses to the gardening and yardwork questions are not subsequently used for computing PA scores, and thus were not analysed in the present study.

Data Analysis

As described elsewhere ([redacted]), each interview for the larger study, which included the use of cognitive interviewing techniques with the four PA questionnaires, was transcribed into a word processing file, and field notes about the interview were added to the file. Data within the file were labeled to reflect the specific questionnaire and PA domain to which they pertained (e.g., AAS walking questions). Each file was then imported into NVivo 8 qualitative analysis software (QSR International, Melbourne, Australia) to organize, manage and code the data. The importation process automatically created a category in NVivo for each label, to allow for the coding of data separately by domain within each questionnaire (e.g., walking, vigorous PA, and moderate PA domains in AAS). Data pertaining to AAS were used for the analysis described in this paper.

Using the Conrad and Blair model (Conrad & Blair, 1996), a participant's AAS responses were further coded into the comprehension, response formulation, and response mapping stages of questionnaire response. This model suggests that within each stage, participants can encounter computational, lexical, logical, omission, or temporal problems (Conrad & Blair, 1996), and thus, in coding AAS data, we specifically sought out evidence of these problems in completing AAS within each stage. In addition, we created a separate 'Activities' theme to describe the variety of activities that participants reported that they had included in their response to AAS questions.

Two researchers ([redacted]) jointly developed the initial themes within each stage in the Conrad and Blair model. They began by first analyzing AAS data from participants who completed AAS as the first of the four questionnaires they completed for the larger study described above. Next, the researchers used these themes to jointly code transcripts from participants who completed AAS as their second or third questionnaire, and they noted any themes that they had not noted in transcripts already coded. The first author then reviewed all themes, merged those that overlapped, and summarized the findings. To assess the credibility of the findings, she reviewed the transcripts of participants who completed the AAS as the first questionnaire in their interview, to ensure that the summary reflected only themes present in these transcripts and thereby to exclude the possibility that themes could have been influenced by the completion of PA questionnaires prior to completion of AAS. The second author then reviewed the summary to confirm that the major themes were captured and the data were well represented.

Results

Characteristics of the analysis sample (n = 44) are presented in Table 1. Participants' ages ranged from 65 to 89 years. Most participants were born in Australia and were retired although a few still worked, at least part-time. Half did not have a university degree, and some reported difficulty managing on their income. All participants reported at least good health although a few were limited in their ability to walk 1 km.

The activities included in participant responses to questions about walking, vigorous PA, and moderate PA are shown in Table 2. Participants reported these activities during the think-aloud process or, when not verbalized during that process, during probing of the activities they

were including in their responses. Activities are categorized into transport-related activities, housework, exercise/sport, and other activities.

There was variation in the way that participants understood AAS instructions and wording and in the way they formulated and mapped responses. This led to both under-reporting and over-reporting of frequency and duration of activities. Specific problems with comprehension, response formulation and response mapping are summarized in Table 3 and described in detail below.

Comprehension of the Requested Information

Misunderstanding the scope of activities for inclusion in each domain. Most participants understood that they were to include walking for exercise and for transport in responses to walking questions. A few participants, however, when probed, indicated that they had not included walking for transport and thus under-reported their walking. Some included start-and-stop walking in the house, walking around the yard, or walking while doing housework, and some included walking as part of playing a sport (e.g., golf, lawn bowls), resulting in over-reporting. One man asked, “What if I also did some running?...I’m going to include running as well as when I’m walking.”

With the examples of vigorous and moderate PA provided in AAS limited to sports and exercise, most participants limited the activities they reported as vigorous or moderate to such activities and may have under-reported their activities as a result. For example, one woman said, “I don’t play competitive sport anymore,” so reported no moderate PA. A few participants considered the examples to be the only activities to report, and reported no activity if they had not done any of the example activities.

Some participants said they did no activities that made them puff and pant and therefore reported no vigorous PA, as would be appropriate based on the AAS definition of vigorous PA. One male, who focused on the puff and pant criterion, limited his response to the portion of his cycling that made him puff and pant (i.e., the riding uphill). Most participants, however, said they did not puff or pant but they did vigorous PA that made them breathe heavier, "work harder," or "get warmer" and included such activities in their responses, which may have resulted in over-reporting of vigorous PA. One man explained the inclusion of his golfing as vigorous PA, "Golf is probably the most vigorous that I did,... and it is hard work getting up some of the hills pulling a buggy;...I might breathe a little bit harder, but I don't get to the puffing or panting stage."

Most participants interpreted moderate PA to only include exercise and sports. Some however interpreted the term more broadly and likely over-reported their activities. One man interpreted moderate PA to be "physical activities that were not working, not sort of pottering around the house." Another man, who questioned what types of activities were left to report as moderate after having reported walking and vigorous PA, concluded that he should record walking that he had not already reported, including "walking that wasn't for 10 minutes…ordinary walking around the house, household chores". Others also included their housework as the previous questions (about vigorous PA) had specifically instructed to not include housework. For example, one woman said, "Okay, now I get to talk about the other kinds [of activity];…cleaning house would count, so let's see…mopping, sweeping, vacuuming, spraying the carpet, and straightening up and doing other things - moving things around, re-arranging."

Differing interpretation of 'times'. The word *times* as a measure of activity frequency was interpreted in two different ways. In reporting walking or cycling for transport, some participants counted each portion of the journey as a separate time (e.g., walking "into town and out of town" as two times), and others counted the whole journey as one time. AAS is not clear on which definition of *time* is appropriate and thus it is not clear if there was under- or over-reporting by some participants. Almost all participants correctly counted each session of continuous exercise as one time (e.g., one "time" for one session at the gym that included multiple activities) although one woman over-reported by counting each session at the gym as two times, one for the time she spent on a stationary bike and the second time for the time she spent on a rowing machine.

Misunderstanding 'at least 10 min' of continuous activity. All participants seemed to understand that they were to include walking lasting at least 10 min, but some were confused by the *walking continuously for at least 10 min* wording. A few participants interpreted the phrasing to mean that they were only to consider each 10+ min walk to have lasted only 10 min, no matter the total duration of the walking. For example, one woman reported that the total time of her four walks (each up to 1 hour in length) took 40 min (each walk allowed to be counted as 10 min x 4 walks). Thus she under-reported her walking duration. One man also thought that each of his walks could only be counted as lasting 10 min in duration even though his walks lasted longer than 10 min. He decided to report a greater frequency of walking than actually done, to compensate for the fact that he was counting each walk as a 10-min walk (i.e., 10 walking sessions lasting more than 10 min, each reported as 15 walking sessions each lasting 10 min). As a result, he over-estimated his walking frequency in order to more accurately capture duration. One woman understood that she was to report the number of days, instead of the number of

sessions, in the previous week that she had walked continuously for 10 min, and therefore she recorded 7 times of walking, and as a result, under-reported her walking frequency.

Misunderstanding the duration of activity. Some participants over-reported the duration of activities by reporting the entire duration of an activity rather than the portion that pertained to the question being asked. For example, some participants reported the total duration of their shopping within their responses to the walking duration question although the shopping inevitably involved standing as well. Golfers typically described their entire time on the golf course as time spent walking or doing moderate or vigorous PA even though they were not moving the entire time. However, a few participants correctly reported as vigorous PA only that part of an activity that made them breathe heavier, such as the walking uphill or the most strenuous part of a bicycle ride.

Problems with Formulating Responses

Miscalculating frequency. Participants used a variety of strategies to calculate activity frequency that resulted in under-reporting for some and over-reporting for others (listed in Table 4). A common strategy was to mentally go back through the week and count the number of sessions. However, this process was difficult for recalling walking. Participants would often start with counting the number of walking sessions but then needed to do some guesswork because they could not remember every walking session. A woman said,

> I'm just trying to remember what we did on the weekend, it's a little hazy. I remember we did shopping and various things. I can't remember for certain if we did some exercise [walking], but I think we did some...so, yeah, I just added another day for the weekend because I can't remember exactly what.

A few participants reported that they were physically active every day, and therefore recorded seven times to represent 7 days of activity. For some participants the seven times represented the minimum amount of times. For example, one man explained that he only reported the early morning walks he did "religiously every day" but that he had not included other walking.

Most participants reported activity of the past 7 days; however, some did not. Using words like "on a regular basis," "on average", "weekly" or using the present tense (e.g., "I walk at least once or twice a week"), these participants reported their general impressions about the amount of activity they did as part of their daily lives and such reporting may have inflated their estimates of PA frequency. One woman reported, "I walk to [the] bridge and then I walk back home, so that's twice a week, that's four times. And then I have...walking to school on the volunteer day and 3 days that I just walk in the morning." One man explained that it was better for him to remember the number of times he "regularly" did the activity given the previous week was "erratic" for which he reported "difficulty answering your questions."

The next strategy was to multiply the number of sessions "normally" done per day by 7 days (e.g., "Really probably 3 times a day, x 7 days, 21 times."). This strategy was used most frequently for calculating the number of walks, given difficulties in remembering all walking sessions. A related strategy was to guess what was normally done during the week. One woman said, "I'm only guessing because I can't possibly remember how many times I go [walking], but I do one every morning, that's 10 minutes, and I go to the shops maybe...it depends what I'm doing so I can only guess". A unique strategy used by one woman was to subtract the number of days that she did not cycle from 7 days (as 'it was raining"), to arrive at the number of days she did moderate PA in the previous week. Her rationale was that "I do normally get out on the

bike", so if she could recall specific instances of not cycling, she would arrive at the correct number of cycling sessions.

Miscalculation of duration. Participants calculated the total time spent in activities of the previous week or of a "regular" week using a range of strategies that resulted in under-reporting for some and over-reporting for others (Table 4). Calculations included multiplication (of the minutes of an activity repeated during the week) and/or summing (adding up duration of different activities). For some individual activities, such as incidental walking or "doing things with the grandkids", a guesstimate of an activity's duration was made before calculations were done to arrive at the total weekly duration. For other activities, particularly exercise and sporting activities, participants knew the duration of their activities as they had timed the activities or the activities were a set length, like an aqua aerobics class or an exercise walk done regularly. Typically then, participants summed times spent in activities with known durations with time spent in activities for which they had to guess the durations.

In making calculations, some participants clearly over-reported duration as they rounded up the total time they were reporting. One man admitted, "I've overestimated a little bit." Sometimes rounding up was to account for incidental walking that was not included in the response to the walking frequency question. For example, estimated duration of walks included "another couple of hours for messing around down in the yard" or "whatever 10 minutes I did around the house". In contrast, a few participants indicated that they had reported the minimum duration of their activities. For example, one man reported, "I think 6 hours would probably be underestimating [my walking], but 6 hours minimum anyhow."

A number of participants miscalculated the total time spent in an activity domain. Some errors were simple math miscalculations: "2 hours a day, 7 days in a week: that's 140 hours."

The main difficulties, however, were with calculating the total weekly duration for activities of varying lengths, trying to account for incidental activities, and multiplying the minutes of an activity performed on multiple days and then adding this to the minutes of other activities. One woman reported, “Well, 6 days at 40 minutes, that’s 2 hours and 40 minutes, and the one day is 60 minutes. Now have I done that sum rightly? So together its 3 hours and 40 minutes?”

Some strategies for calculating duration were only used in calculating the duration of walking activities. One such strategy was estimating the total time spent walking on a “normal” day and multiplying that by 7 days to arrive at a total weekly duration of walking. The participants who understood that the duration of each walk could only count as 10 min multiplied the number of walks, or guessed the number of walks, by 10 min, no matter the actual length of the walks.

Reporting the same activity in multiple domains. Some participants reported the same activity in the walking domain and either the vigorous or moderate PA domain and thus over-reported the frequency and duration of their PA. For example, walking for exercise was often included in both the walking and vigorous PA domains. Some participants included sports and exercise, like golf and aqua aerobics, as both walking and moderate PA. Activities for which the intensity varied over the course of the session (e.g., exercise classes) were typically reported as vigorous PA only.

Problems with Mapping Response Options

Once participants had formulated their responses to questions about frequency of activities (number of times), they had no difficulties mapping responses to the response boxes. However, they did encounter some difficulty with mapping responses to questions about duration (total time in the previous week) to the response options. The response options included two boxes to

record hours followed by two boxes to record minutes. Therefore, some participants divided to convert minutes into hours, in order to fit answers into the boxes. One participant calculated, "20 times x 10 [minutes] is 200 minutes," but to then put her response into the response boxes, she miscalculated: "it would be 4 hours, wouldn't it?" Such miscalculations lead to both under-reporting and over-reporting of duration of activities.

Discussion

This study documents the cognitive processes used by a sample of older Australian adults to respond to a self-administered version of AAS, a PA surveillance questionnaire that uses open-ended response options for questions about the frequency of sessions of walking, moderate PA and vigorous PA over the past week and the duration (total time) of these activities for that time period. The data from cognitive interviews suggest variation in the way that older adults *understood* the questions, and *formulated* and *mapped* their responses. This led to both under-reporting and over-reporting. Below we summarise our findings and make suggestions for possible ways to improve the accuracy of the data collected from older adults and possibly other adults as well, acknowledging that further research is needed to clarify if the same cognitive processes and problems occur in younger age groups, and that any changes made to AAS would need to be tested in both older and younger adults, before being used for population level surveillance.

Understanding the Questions

Walking questions were the first to be asked, and participants wanted to include all their walking, including activities that were more appropriate to include later as either vigorous PA or moderate PA (e.g., golf) and that often got included in these other domains as well. Such 'double-dipping' has been reported for other PA questionnaires, in samples of middle-aged

adults (Altschuler et al., 2009) and, by us, of older adults ([redacted]. To decrease such misclassification and over-reporting, we do not recommend changing the ordering in which domains are asked. In adapting a PA questionnaire for older adults, researchers in Sweden found that participants were uncomfortable with first being presented with questions about activities done rarely or seldom (vigorous and moderate PA), and therefore, questions about activities done frequently (walking) were arranged to be first (Hurtig-Wennlof, Hagstromer, & Olsson, 2010). We do suggest that to decrease the tendency to report the same activity more than once, in introductory instructions, participants be instructed that AAS will ask separately about walking, vigorous PA and moderate PA and that activities reported in one domain of activity are not to be reported in other domains.

Another issue was that some participants did not understand that *continuous* meant non-stop walking. We have found this to be a problem for the short 7-day form of the International PA Questionnaire (IPAQ) too ([redacted]). Adding that start-and-stop walking, such as walking around the house or in shops, is not to be included or simply rephrasing *continuous* to *non-stop* walking may help address this problem. The wording *at least 10 min* in the walking frequency question was also problematic. Other researchers have found that participants do not consider the 10-min rule to be an absolute cut-off for walking questions (Rzewnicki, Van den Auweele, & De Bourdeaudhuij, 2003). In the current study, some participants did not understand that *at least 10 min* was to be applied to the walking duration question as well. Clarifying that participants are only to include walking reported in the frequency question (e.g., re-stating that only walking lasting at least 10 min is to be reported for the duration question or revising the term *walking this way* that is used in the duration question) could address the issue.

Another issue most relevant to walking questions was misunderstanding about the number of sessions to report for transport-related activity. The simplest suggestion for change would be to instruct how to count *each* session of active travel. However, for calculating the prevalence of people meeting PA guidelines that call for doing PA at least 5 days of the week, or most days of the week (Australian Government Department of Health and Aged Care, 1999), it would be more useful to inquire about the number of days, instead of number of *times* for calculating PA frequency. Doing so would also reduce the cognitive processing required to answer the question and decrease other problems we uncovered in reporting frequency of activity.

Questions about vigorous PA were the easiest to understand as these activities tend to be planned, structured activities, and we recommend that no changes be made to these questions. Most participants correctly relied on the description of vigorous PA rather than on the examples. Similar to our previous findings for IPAQ ([redacted]), only a few participants considered the examples to be choice-limiting checklists, although most participants reported an entire session of activity rather than the component of the session that was vigorous-intensity. However, in comparison to IPAQ, in completing AAS, participants tended to struggle much less with deciding what activities to include as vigorous PA. This may reflect a number of factors. First, IPAQ provides examples of vigorous PA but no description, and participants perceived that the IPAQ examples were activities that could be performed at different intensities. In the present analysis, participants did not express such confusion with the examples, and no participant questioned the meaning of the vigorous PA description (*breathe harder* or *puff and pant*). Some participants did include moderate PA or walking as vigorous PA because their breathing increased when they did those activities. In contrast to IPAQ and most other PA questionnaires, AAS asks participants to report vigorous gardening and yard activities immediately before the

vigorous PA question to discourage participants from reporting these activities as vigorous PA. Moreover, AAS instructs participants to exclude household chores, gardening and yardwork from responses to vigorous PA questions, decreasing the range of activities participants have to consider in their responses.

Participants were, however, uncertain about the type of PA to include for *other more moderate* PA, given no definition of moderate was provided and most of the walking and exercise had already been reported. Thus, a definition and examples are suggested. Our previous research also found that older adults were uncertain about the types of PA to report as moderate PA for IPAQ ([redacted]), and confusion with the intensity of activity to include as moderate PA has been reported for other PA questionnaires in younger populations (Altschuler et al., 2009; Craig et al., 2003; Ham, Macera, Jones, Ainsworth, & Turczyn, 2004; Rzewnicki et al., 2003). In focus groups with older Australian women (Brown, Fuller, Lee, Cockburn, & Adamson, 1999), participants agreed that many activities (e.g., walking, exercise classes, golf, cycling) could be performed at different intensities, which makes decisions about categorising PA as moderate difficult without definitions of intensity provided. Such findings may explain why questions measuring moderate PA have lower test-retest reliability estimates than walking and vigorous PA questions (Brown, Trost, Bauman, Mummery, & Owen, 2004).

Formulating and Mapping a Response

As shown for other PA questionnaires ([redacted]) and for answering questions in surveys more generally (Tourangeau, Rips, & Rasinski, 2000), participants used a number of strategies and often used a mix of strategies for calculating frequency and duration of their activities. Walking in particular required a mix of strategies. Asking participants to recall incidental behaviour like walking is known to be

problematic, and so some degree of individual error is likely. Providing further clarification about the type of walking to include and exclude (e.g., include non-stop walking; exclude walking in shops) could decrease the number of sessions that participants have to recall, and as a result may somewhat improve computations, and therefore estimations.

However, it should be acknowledged that for habitual activities like walking, it was very difficult for participants to recall with accuracy each session and the duration of each session in the last week. Thus, few participants relied exclusively on episodic enumeration (recalling and counting) for calculating frequency or summing total weekly min of walking for calculating total duration. Most needed to at least supplement calculations with estimations based on their general impressions of frequency and duration. Turning such impressions into numerical responses can be a challenge and may result in flawed estimations (Tourangeau et al., 2000). Indeed, for some participants, the cognitive effort required to recall each session and its duration was too great and therefore they guessed. This finding is in line with those from previous studies of how people respond to surveys (Tourangeau et al., 2000).

This difficulty with retrieval of PA sessions may be explained by the role of memory in responding to questionnaires. Durante and Ainsworth (1996) discuss theories of memory relevant to the recall of PA behaviour. In short, these theories contend that general, abstract memories are distinct from memories of specific events, but that general memories “intrude” when trying to retrieve the memory of a specific event, potentially decreasing the accuracy of the recall of the event. Also, as the time since the event increases, the memory of the specific event become less clear and blends with similar memories, leading to further inaccuracies in the recall of the event. Durante and Ainsworth (1996) explain that these issues are particularly problematic for recalling specific events that are highly habituated. In the current analysis, the theories of memory indicate

that retrieval of specific sessions of PA, particularly a highly habituated activity like walking, may be influenced by general recollections of "usual" PA, leading to an inaccurate assessment of the frequency and duration of specific sessions.

For the measurement of PA, there is debate about whether activities of a usual week or a specific week should be assessed. Our findings indicate that for at least some older adults, reporting a usual week may be preferred, a finding supported by our analysis of older adults' responses to IPAQ ([redacted]) and by others (Durante & Ainsworth, 1996). Problems with short-term memory may be a particular problem for some older adults (Jonker, Geerlings, & Schmand, 2000) for whom being asked to recall specific events could be stressful.

For calculating frequency, we have already suggested that days be used instead of *times*. Such a change would decrease the number of possible events (i.e., days from times) to recall, making the episode enumeration process easier. If days were used, in conjunction with the existing 'last week' timeframe, participants could be instructed to consider whether the activity had been performed the previous day and then work backwards to count the total number of days. This process is used with 7-day recall instruments and is suggested for recall of PA in questionnaires with short referent periods like AAS (Durante & Ainsworth, 1996). This in turn may help participants focus on the events of the last week.

To decrease errors in converting minutes to hours or vice versa, sufficient numbers of response boxes (with appropriate instruction) should be provided to allow participants to write responses in minutes or hours only.

Limitations

Responses of participants who completed AAS as the second or third questionnaire may have been influenced by questions asked in earlier questionnaires included in the larger study.

However, we confirmed that all the main issues we describe were evident in the transcripts from participants who completed AAS first. Also of note is that the design of this research did not allow for comparisons of cognitive errors between older and younger adults. Therefore, it remains unknown whether cognitive errors are greater in older populations.

There are also limitations to our use of cognitive interviewing techniques. First, as with cognitive interviewing more generally (Willis, 2005), participants were self-selected and were thus not representative of all Australian older adults. Our sample was generally well educated and cognitively healthy. More problems with completing AAS could be expected in people with cognitive decline, in less well educated populations, and in culturally and linguistically diverse populations, as these groups may have different interpretations of the concepts in the questionnaire. Second, the cognitive interview setting creates an artificial testing environment (Willis, 2005). Participants may have responded differently if they had completed AAS on their own or if they had been administered AAS using a different mode of delivery (e.g., interviewer-administered). To decrease the influence of the testing environment, participants were provided no guidance in how to respond, and probing only occurred once participants had completed questions of a PA domain. Particular concerns with the think-aloud process are that some participants do not perform this task well (Willis, 2005) and that it may interfere with the cognitive process of responding (Beatty & Wallis, 2007). In the current study, some participants did not engage in thinking aloud, or did it poorly. For those participants, probing questions were asked, to reveal as much about the thought process as possible. Evidence suggests that responses to probes like ours, asked immediately following questions, are likely to not be much different than think-aloud reports (Beatty & Willis, 2007). Third, analysis of cognitive interview data can result in faulty conclusions if problems are identified that do not exist in survey administration in

the field (Beatty & Willis, 2007), and thus testing of suggested changes to AAS is needed before a revised version is used for surveillance.

Future Research

We need to be cautious that making any changes, particularly adding instructions and more examples, does not make AAS more burdensome to complete and does not confuse people who would have had no problems completing the current written AAS version. Thus any changes made to AAS based on our findings would need testing, such as split ballot testing, to learn whether an altered version that addresses the problems we uncovered provides a more reliable and valid measure of PA than the current version. However first, learning whether the problems we uncovered are found across all adult age groups is needed, given AAS is used for population surveillance. We would suggest cognitive interviewing to do this, prior to comparing an altered version to the AAS version used in the current study.

Conclusions

Our findings suggest that AAS questions about walking, vigorous PA and moderate PA may not be interpreted by older adults as intended by the survey developers. They also suggest that there is variation in the way that older adults interpret AAS instructions and understand the meaning of the words used in AAS. Data collected from cognitive interviews revealed challenges in understanding the scope of activities to include, understanding particular wording used in the questions and instructions, and making the required calculations for formulating and recording responses. Participants had the most difficulty with the walking questions. Over-reporting errors were more common than under-reporting errors. Based on the findings, we have made recommendations that could be used for adapting the self-administered version of AAS for use in populations that include older community-dwelling adults, although additional research is

needed to test the impact of these changes on the accuracy of the data collected. We acknowledge that making any changes to PA questionnaires like AAS can disrupt the monitoring of PA over time, and therefore the advantages and disadvantages of using a revised version need to be carefully considered.

References

Altschuler, A., Picchi, T., Nelson, M., Rogers, J. D., Hart, J., & Sternfeld, B. (2009). Physical activity questionnaire comprehension: Lessons from cognitive interviews. *Medicine and Science in Sports and Exercise, 41*(2), 336-343. doi:10.1249/MSS.0b013e318186b1b1

Armstrong, T., Bauman, A., & Davies, J. (2000). *Physical activity patterns of Australian adults* (Cat. No. CVD 10). Canberra, ACT: Australian Institute of Health and Welfare.

Australian Government Department of Health and Aged Care. (1999). *An active way to better health: National physical activity guidelines for adults*. Canberra, ACT: Australian Government Publishing Service.

Australian Institute of Health and Welfare. (2003). *The Active Australia Survey: A guide and manual for implementation, analysis and reporting.* (Cat. No. CVD 22). Canberra, ACT: Australian Institute of Health and Welfare.

Bauman, A., Phongsavan, P., Schoeppe, S., & Owen, N. (2006). Physical activity measurement--A primer for health promotion. *Promotion & Education, 13*(2), 92-103. doi: 10.1177/10253823060130020103

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71*(2), 287-311.

Brown, W. J., Burton, N. W., Marshall, A. L., & Miller, Y. D. (2008). Reliability and validity of a modified self-administered version of the Active Australia physical activity survey in a sample of mid-age women. *Austrialian and New Zealand Journal of Public Health, 32*(6), 535-541.

Brown, W. J., Fuller, B., Lee, C., Cockburn, J., & Adamson, L. (1999). Never too late: Older peoples' perceptions of physical activity. *Health Promotion Journal of Australia, 9*(1), 55-63.

Brown, W. J., Trost, S. G., Bauman, A., Mummery, K., & Owen, N. (2004). Test-retest reliability of four physical activity measures used in population surveys. *Journal of Science and Medicine in Sport, 7*(2), 205-215.

Centers for Disease Control and Prevention. (2003). Behavioral Risk Factor Surveillance System Survey Questionnaire. Retrieved from http://www.cdc.gov/brfss/questionnaires/questionnaires.htm

Chau, J., Smith, B., Chey, T., Merom, D., & Bauman, A. (2007). *Trends in population levels of sufficient physical activity in NSW, 1998 to 2005: Full report* (Report No. CPAH06-001c). Sydney, NSW: NSW Centre for Physical Activity and Health.

Chodzko-Zajko, W. J., Proctor, D. N., Fiatarone Singh, M. A., Minson, C. T., Nigg, C. R., Salem, G. J., & Skinner, J. S. (2009). American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Medicine and Science in Sports and Exercise, 41*(7), 1510-1530. doi:10.1249/MSS.0b013e3181a0c95c

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research, 12*(3), 229-238.

Conrad, F., & Blair, J. (1996). From impressions to data: Increasing the objectivity of cognitive interviews. In *Proceedings of the Survey Research Methods Section of the American Statistical Association* (pp. 1-10). Alexandria, VA: American Statistical Association.

Craig, C. L., Marshall, A. L., Sjostrom, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., …Oja, P. (2003). International Physical Activity Questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise, 35*(8), 1381-1395.

Drennan, J. (2003). Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing, 42*(1), 57-63.

Durante, R., & Ainsworth, B. E. (1996). The recall of physical activity: Using a cognitive model of the question-answering process. *Medicine and Science in Sports and Exercise, 28*(10), 1282-1291.

Fisher, R. P., Falkner, K. L., Trevisan, M., & McCauley, M. R. (2000). Adapting the cognitive interview to enhance long-term (35 years) recall of physical activities. *Journal of Applied Psychology, 85*(2), 180-189.

Ham, S. A., Macera, C. A., Jones, D. A., Ainsworth, B. E., & Turczyn, K. M. (2004). Preliminary considerations for physical activity research: Variations on a theme. *Journal of Physical Activty and Health, 1*, 98-113.

[redacted]

Hurtig-Wennlof, A., Hagstromer, M., & Olsson, L. A. (2010). The International Physical Activity Questionnaire modified for the elderly: Aspects of validity and feasibility. *Public Health Nutrition, 13*(11), 1847-1854. doi: S1368980010000157

Jobe, J. B., & Mingay, D. J. (1989). Cognitive research improves questionnaires. *American Journal of Public Health, 79*(8), 1053-1055.

Jonker, C., Geerlings, M. I., & Schmand, B. (2000). Are memory complaints predictive for dementia? A review of clinical and population-based studies. *International Journal of Geriatric Psychiatry, 15*(11), 983-991. doi:10.1002/1099-1166(200011)15:11

Lagerros, Y. T., & Lagiou, P. (2007). Assessment of physical activity and energy expenditure in epidemiological research of chronic diseases. *European Journal of Epidemiology, 22*(6), 353-362.

Marshall, A.L., Bauman, A.E., Owen, N., Booth, M.L., Crawford, D., & Marcus, B.H. (2004). Reaching out to promote physical activity in Australia: A statewide randomized controlled trial of a stage-targeted intervention. *American Journal of Health Promotion, 18*(4), 283-287.

McKenna, J., Foster, L. J., & Page, A. (2004). Exploring recall of physical activity in young people using qualitative interviewing. *Pediatric Exercise Science, 16*(1), 5-14.

Merom, D., Rissel, C., Phongsavan, P., Smith, B.J.,Van Kemenade, C., Brown, W.J., & Bauman, A.E. (2004). Promoting walking with pedometers in the community. *American Journal of Preventive Medicine, 32*(4):290-297.

Nelson, M. E., Rejeski, W. J., Blair, S. N., Duncan, P. W., Judge, J. O., King, A. C., …Castaneda-Sceppa, C. (2007). Physical activity and public health in older adults: Recommendation from the American College of Sports Medicine and the American Heart Association. *Medicine and Science in Sports and Exercise, 39*(8), 1435-1445.

Nigg, C., Hellsten, L., Norman, G., Braun, L., Breger, R., Burbank, P., …Williams, G. (2005). Physical activity staging distribution: Establishing a heuristic using multiple studies. *Annals of Behavioral Medicine, 29*(Suppl), 35-45.

Reeves, M.M., Marshall, A.L., Owen, N., Winkler, E.A.H., & Eakin, E.G. (2010). Measuring physical activity change in broad-reach intervention trials. *Journal of Physical Activity & Health, 7*(2):194-202.

Rzewnicki, R., Van den Auweele Y., & De Bourdeaudhuij, I. (2003). Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutrition, 6*(3), 299-305.

Timperio, A., Salmon, J., & Crawford, D. (2003). Validity and reliability of a physical activity recall instrument among overweight and non-overweight men and women. *Journal of Science and Medicine in Sport, 6*(4), 477-491.

Tomlin, S., Joyce, S., & Patterson, C. (2012). *Health and wellbeing of adults in Western Australia 2011: Overview and trends*. Perth, WA: Government of Western Australia Department of Health.

Topolski, T. D., LoGerfo, J., Patrick, D. L., Williams, B., Walwick, J., & Patrick, M. B. (2006). The Rapid Assessment of Physical Activity (RAPA) among older adults. *Preventing Chronic Disease, 3*(4), A118.

Tourangeau, R. (1984). Cognitive science and survey methods: A cognitive perspective. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey design: Building a bridge between the disciplines* (pp. 73-100). Washington, DC: National Academy Press.

Tourangeau, R., Rips, L.J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, UK: Cambridge University Press.

[redacted]

[redacted]

[redacted]

Washburn, R. A., Smith, K. W., Jette, A. M., & Janney, C. A. (1993). The Physical Activity Scale for the Elderly (PASE): Development and evaluation. *Journal of Clinical Epidemiology, 46*(2), 153-162.

Willis, G.B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.