Running head: HOW PEOPLE INTERPRET SITTING QUESTIONNAIRES

**How do people interpret and respond to self-report sitting time questionnaires?**

**A think-aloud study**

Benjamin Gardner \*

Ioanna Louca

Danai Mourouzis

Alessandra Calabrese

Aeysha Fida

&

Lee Smith

Author Note

Benjamin Gardner, Ioanna Louca, Danai Mourouzis, Alessandra Calabrese, Aeysha Fida: Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, De Crespigny Park, London, SE5 8AF, UK. Lee Smith: Cambridge Centre for Sport and Exercise Sciences, Anglia Ruskin University, Compass House, Cambridge, CB1 1PT, UK. Contact: [benjamin.gardner@kcl.ac.uk](mailto:benjamin.gardner@kcl.ac.uk); Phone: +44 207 848 6926.

\* Corresponding author.

Declaration of interest: All authors declare that they have no competing interests.

**How do people interpret and respond to self-report sitting time questionnaires?**

**A think-aloud study**

**Abstract**

Sedentary behaviour research to date has been predominantly based on self-reported sitting time, yet little attention has been paid to how respondents interpret sitting questionnaire items. 25 office workers participated in qualitative, ‘think-aloud’ interviews, describing their thoughts while completing 43 items derived from 9 existing questionnaires. Inductive Thematic Analysis identified four potential interpretation or response problems: misinterpretation and uncertainty; the mental calculation process involved in formulating responses; self-presentation concerns; and the affective and motivational impact of questionnaire completion. Results not only show that lay representations of sitting may diverge from those of researchers, but also highlight potential errors and biases encountered when generating sitting estimates. Additionally, reporting sitting may generate a desire to reduce sitting time. Findings suggest that domain-specific measures that estimate sitting across different settings may better correspond with participants’ perceptions. Future research should investigate the potential for sedentary behaviour questionnaire completion to change behaviour.

**Introduction**

Sitting time has been linked to mental and physical ill-health, and early death (1–3). Many studies have sought to document, understand or change sitting time (4,5), with especial attention paid to at-risk populations such as office workers, who typically sit for around two-thirds of their working and waking day (5,6).

Much sedentary behaviour research has been based on participant-reported sitting time. Many self-report sitting time questionnaires are available (7,8). These differ according to properties of the sitting episode (e.g., the purpose or domain [work, leisure], the physical or social environment, or time of day; (8,9)), properties of the recall period (e.g., typical versus actual period; single versus multiple days; weekdays versus weekends; (7)), and measurement dimensions (e.g. unit of time; duration versus bouts of sitting or breaks). For example, the International Physical Activity Questionnaire (IPAQ; 10) estimates sitting across a 7-day period by summing typical weekday sitting time across five days, and typical weekend sitting time across two days, via a direct recall item (e.g., ‘how much time did you spend sitting on a typical weekday?’). By comparison, the Multi-context Sitting Time Questionnaire (MSTQ; 11) assesses time spent in each of five typically-seated activities (e.g., watching TV), on workdays and non-workdays respectively.

Self-report questionnaires assume that participants share researchers’ interpretations of items, have insight into their sitting, and can accurately recall sitting when prompted. Yet, little attention has been paid to how people comprehend and respond to sitting questionnaires (12). Instead, the utility of sitting time questionnaires is typically established only against observational data (13), with measures that correlate with objective monitor assessments of sitting assumed to be valid and reliable. While many measures demonstrate validity in this respect, convergence is often imperfect (14); for example, studies have questioned the correspondence between the IPAQ, one of the most commonly used sitting time measures, and monitor assessments (15). Additionally, correlational evidence reveals only whether increases or decreases in monitor-assessed sitting time are mirrored by increases or decreases in self-reported sitting, not the accuracy of estimates. Indeed, the IPAQ has been shown to underestimate true sitting time despite adequate convergent validity (16).

‘Think aloud’ methods, whereby participants vocalise their thoughts, provide insight into respondents’ perceptions as they interpret and formulate responses to questionnaire items (17-20). Ericsson and Simon (18) distinguish between three ‘levels’ of think-aloud response. ‘Level 1’ responses are vocalisations of verbal thoughts encoded as part of a participant’s ‘inner speech’, and ‘Level 2’ responses are a more effortfully generated description of thoughts that would not otherwise have been verbally encoded. While responses corresponding to Levels 1 and 2 describe thoughts, ‘Level 3’ responses are reflections on thoughts or behaviours, such as explaining or assigning meaning or motive to them. While playing golf, for example, verbalisations at Levels 1 and 2 might reflect descriptions of task features (e.g. ‘it is mostly uphill’) or self-instruction (‘concentrate on this’), whereas Level 3 verbalisations might reflect explanations for performance, such as justifying club selection (21,22). Verbalisations corresponding to Levels 1 and 2 capture information stored momentarily in short-term memory that would otherwise be lost or subject to recall errors and biases if assessed retrospectively (23). While often used within sport and exercise psychology to describe thoughts during sports performance, or assess relationships between thoughts and performance (19-20, 22, 24-26), think-aloud methods as applied to questionnaire evaluation are designed to highlight comprehension and response problems that can threaten the validity of data (27). For example, participants’ interpretations may deviate from intended meanings (17), which can lead researchers to misinterpret responses. Studies in which ‘think aloud’ methods have been used to obtain Level 1 and Level 2 verbalisations can offer valuable information on how questionnaire items might be improved, as well as revealing how respondents represent and experience the focal phenomena.

To our knowledge, only one study to date has used ‘think aloud’ methods to probe sitting time questionnaires (12). Older adults experienced difficulty in understanding which activities to account for in their responses, and in estimating sitting across days and contexts, in response to the IPAQ and the Physical Activity Scale for the Elderly (12,28). While this study attested to the usefulness of ‘think aloud’ methods for documenting questionnaire comprehension, it focused on only two of the many available sitting time questionnaires.

The present study used ‘think aloud’ methods to investigate office workers’ interpretations and responses to a set of sitting items systematically sampled from multiple questionnaires to capture a range of question types (7-9). The study explored problems in item comprehension and response formulation. Office workers were chosen because they tend to be highly sedentary (5,6), so are often sampled within sitting time research.

**Materials and Method**

*Researcher positionality*

The two senior authors (BG, LS) have been active in the sedentary behaviour field since the mid 2010s, conducting research to variously observe, predict, explain and modify sitting. BG is a social psychologist with qualitative and quantitative research experience, and LS an epidemiologist with a background in sports science. All other authors were undergraduate psychology students with minimal extracurricular research experience, and no prior involvement in sedentary research. Authors IL and DM collected data at the telecommunications company, under supervision of BG, as coursework. Authors AC and AF collected data from the university sample, and contributed to analysis, as part of an internship supervised by BG. All authors regarded themselves as relatively sedentary, spending at least 6 hours per day sitting.

All authors approached the study from a post-positivist realist philosophical position (29). BG conceived of the study in response to anecdotal experience of participants expressing confusion about questionnaire measures of sedentary habits (30), having previously conducted a ‘think aloud’ study that revealed problems underlying self-reports of habitual behaviour (31). The ultimate goal of the study was to help researchers and practitioners to identify and develop optimal self-report measures of sedentary behaviour.

*Selection of questionnaire items*

Items were identified from an existing review of 60 sedentary behaviour questionnaires (8). We removed two non-English-language questionnaires, 29 questionnaires not designed for administration to, or validated among, adults from the general public (e.g. children, overweight adults), and 3 questionnaires not designed for self-administration. Six questionnaires identified as duplicates or derivatives (e.g. IPAQ short-form) were excluded. Four questionnaires not freely available were removed and, to minimise participant burden, so were two questionnaires each containing ≥20 sitting-related items. A list of excluded questionnaires is available as Supplemental Material.

Items from the remaining 16 questionnaires were inspected in full. Items relating only to physical activity were removed, as were identical or highly similar items across questionnaires. The final set comprised 43 items, extracted from nine questionnaires: IPAQ short-form (10); Workplace Sitting Breaks Questionnaire (32); Clemes Single Item Sitting Time Questionnaire (33); Survey Question on Television Viewing (34); Marshall Domain-Specific Sitting Time Questionnaire (35); Occupational Sitting and Physical Activity Questionnaire (36); Aadahl and Jorgensen Physical Activity Scale (37); MSTQ (11); and Active-Q (38). Together, these items capture direct recall (e.g. total sitting time) and domain-specific recall (e.g. sitting at work), across a range of periods (e.g., day, week; typical, actual), assessed with single or multiple items (see Table). The full set of questionnaire items is available as Supplementary Material.

INSERT TABLE HERE

*Participants and procedure*

Participants were office workers recruited from two UK-based organisations: a large multinational telecommunications company (N = 17; 13 female, 4 male), and a university (N = 8; 7 female, one male), both situated in inner-city locations in southeast England. These organisations were chosen because they both encompass a diverse range of office environments and employ workers from across the socioeconomic spectrum. Recruitment at the telecommunications company occurred over a four-week period in January-February 2019, and recruitment at the university, undertaken to boost sample size, over a two-week period in October 2019. Employees of the telecommunications company received an email advertising the study, circulated by management, and those from the university were recruited in person by knocking on office doors at the host site. Self-declared eligibility criteria were: fluent in English; part- or full-time employed in a ‘desk-based job’; no physical impairments necessitating prolonged sitting; and no psychology qualification. Due to researcher error, age was not recorded for employees from the multinational company. Six of the 7 university-based participants reported their age (range 21-69y; M = 33y, SD = 16.5y).

Data collection took place at local work sites. First, the researcher asked the participant to ‘*speak every and any thought that comes to mind as you answer each question’* and, to facilitate transcription and analysis, to *‘read the [item] you are referring to out loud first and then say what you think about it’.* To obtain descriptive think-aloud responses (i.e. ‘Level 2’ verbalisations) rather than reflections or rationalisations (i.e. ‘Level 3’ [18]), which can invoke reactivity (39), we asked that participants ‘*don’t plan out what you are saying or explain what you mean – instead just freely voice your thoughts as they come to you’.* To accustom participants to thinking aloud, the researcher provided a spoken example descriptive ‘think aloud’ response to a non-sitting-related item pertaining to typical breakfast consumption duration (‘*I would say between 10 and 15 minutes, usually about 15 minutes, so I will choose 11 to 20 minutes*’). Participants also completed a practice task, ‘thinking aloud’ in response to a sample item assessing coffee consumption. The researcher then left the room and began recording. While we were unable to verify in-situ that participants were ‘thinking aloud’ as instructed, inspection of recordings indicated that all participants generated codable descriptive data. The questionnaire was hosted online, with block-randomised item presentation to control for order effects, and completed using a study laptop. Procedures were verified as feasible among an independent sample of four desk-based employees and approved by the KCL Research Ethics Committee (MRS-18/19-8981).

*Analysis*

Inductive Thematic Analysis (40) of verbatim interview transcripts was undertaken by a team of three coders (BG, AC, AF). All coders independently performed data familiarisation, coding (i.e., assigning conceptual labels to pertinent ‘events’ in the data relating to comprehension and response formulation), and theme extraction procedures on three transcripts. Coders met in person to inductively develop a preliminary thematic framework using a ‘critical friends’ approach to ensure rigour and trustworthiness, whereby all coders offered critical feedback on each other’s interpretations to allow consideration of and reflection on multiple perspectives (41). Two further rounds were undertaken of application (by AC and AF) and critical reflection on the framework (by BG) in light of insights from coding and theme extraction of all remaining transcripts. Theme review and naming, and analysis write-up were undertaken by BG. All authors critically reviewed the final thematic structure and verified it as a valid representation of the data.

**Results**

Four themes were extracted, relating to: item misinterpretation and uncertainty; a mental calculation process involved in generating responses; self-presentation concerns; and the affective and motivational impact of questionnaire completion on participants.

*Misinterpretation and uncertainty*

Some participants misunderstood items. For example, one interpreted items relating to sitting over the course of a day to refer only to daylight hours:

Participant 25 (P25): *“Watching TV, day hours during the week [I would say is] zero … I only watch in the evening during the week.”*

There was uncertainty around whether a ‘workday’ referred to the entirety of waking days on which participants go to work, to time spent in the workplace, or more specifically to work-related activities in the workplace. Some items thus appeared nonsensical to some:

P8: *“[Indicate the number of hours and minutes per non-workday you spend … sitting while working.] Obviously I don’t work at the weekends. That’s weird phrasing.”*

Domain-specific measures were problematic, because some domains were seen as lacking modern-day relevance:

P3: *“I don’t use a computer anymore, I just my phone or tablet for everything. If you asked me how much time I spent looking at my phone, it would be a horrifically high number. But that hasn’t been asked.”*

Other domains were not experienced as mutually exclusive. For example, video streaming activity cut across ‘watching TV or movies’, ‘during transportation’, and ‘using a computer’:

P18: *“[These] distinction[s are] weird. Netflix doesn’t really work like that, does it?”*

Items requiring reports in hours and minutes were unclear to some. Some interpreted them to refer to the time of day *(“[for time spent in sleep or rest,] do I add the time? […] I thought they were asking what time I go to bed*”; P21), and others queried whether they were being asked “*to convert how many hours [I spent sitting] into minutes*” (P5).

While all participants provided quantitative responses to all items, they lacked confidence in the accuracy or meaningfulness of responses to items deemed unclear:

P13: *“‘[Do my answers] reflect my normal activity levels?’ What do you mean? I don’t know. Let’s go with no.”*

*Mental calculation*

Participants struggled to directly recall their sitting time. Some deferred to objective device-based data (“*my Fitbit tells me that I’m averaging about 7 hours 30 minutes [of sleep per night]”;* P9), while others inferred workplace sitting from time spent at work (*“[I work from] 8 until 5.30, so … that’s 9 and a half hours [of sitting]”*; P11). Many participants generated sitting estimates by mentally rehearsing relevant activities and routines:

P13: *“Time spent sitting in leisure time, not including television … it’s very hard to determine. What do I do? Monday [I go to] Virgin [the gym], Tuesday volleyball, Wednesday, Thursday gym, Friday. Hmm. I don’t really sit.”*

To aid activity-specific recall, some imbued items with additional context, thereby excluding other relevant actions or settings (“*there are probably other activities that fall into this … I’ll just think about reading, so that’s going to be [the basis for] my answer*”; P15). Some struggled to provide reliable estimates due to the variability of their routines:

P9: *“That’s a hard question because it completely depends. If I’ve got three back to back meeting, then I will be sat down for the three hours straight. If I don’t have back to back meetings, then I will be drinking a lot of water, so I will get up every half an hour for a toilet break or to fill up my water or go for a walk.”*

While some people appeared satisfied to offer imprecise global estimates based on specific activities (“*reading I guess could be an hour? No, I’ll put thirty minutes*”; P11), others engaged in a systematic mental calculation process, converting time spent in seated activities into item response metrics. This was burdensome (“*[I’m] not good at maths*”; P3), and some arithmetic errors were evident:

P4: *“[I sit in this domain for] six hours a day, so that’s what, about, so it’s like 15, no, sorry, that’s like 67 hours [in total]. Oh, I’ve done this completely wrong!”*

Calculated estimates were typically reviewed for accuracy and adjusted based either on whether they ‘felt’ right (“*texting … thirty minutes I should say, or fifteen minutes, I think half an hour is too long to text*”; P23) or, for items that seek to summarise all activities across a 24h period, whether responses totalled 24h:

P15: “*I’m going to say 6 [hours] for [standing, washing dishes or cooking, driving a car or truck] and then [for a previous item] put up the number of hours I spend working at a computer … to 7 so that it adds up to 24*”

*Self-presentation concerns*

Some participants worried that their responses reflected badly on them (“*God, this makes me sound awful!*”; P11). Many sought to assuage such concerns by justifying their sitting, highlighting a lack of autonomy (“*I don’t like sitting down for too long, but it is a desk-based job*”; P3), or emphasising compensatory activities:

P9: *“I’m sat down 93% of the time, but … that’s not including [the gym] … I repeat, that does not include my gym time that I have for an hour at lunch!”*

One participant appeared to modify her estimates for self-presentation reasons:

P16: *“That’s like, 9.5 hours [of sitting], oh God that’s really bad. Ok, I’m going to put 9 hours.”*

Participants sought consistency in their responses, such that earlier responses provided templates for later, similar items (“*as I said earlier, about maybe 70% of my time [is spent sitting]*”; P15). Similar items sometimes prompted a review and calibration of responses:

P8: “*This is the same [domain] as earlier, so I’m going to make that 1.5 [hours] and change my earlier answer to be 1.5 [hours].*”

*Affective and motivational impact of questionnaire completion*

Questionnaire completion prompted negative affect for some people. Some such responses arose in exasperation at the cognitive burden of completing the questionnaire (“*my God, this is way too much thinking*”; P14) or deciphering unclear items (“*this is a ridiculous question!*”; P8). Others arose because the questionnaire raised awareness of excessive sitting time (“*oh my God, do I really sit for like 80% of the day?*”; P12), a realisation that in turn motivated some to reduce their sitting time (“*oh God, I need to get up and move*!”; P18):

P16: “*That’s so bad. So that means that … 82% [of the time] is just [spent] sat down at my desk at work. God, I should really go to the gym more*.”

**Discussion**

This study demonstrated the thought processes underpinning interpretations of and responses to self-report items about sitting time. Our findings revealed the potential for participants to fail to understand items as intended by researchers, and for social desirability concerns to influence disclosure of sitting time. More fundamentally, we showed that participants experienced difficulty estimating sitting, because they appeared not to conceive of time spent in seated activities as sitting time. Perhaps consequently, generating sitting time estimates appeared to raise awareness, such that some participants were motivated to reduce what they saw as excessive sitting. These findings speak to the importance of using measures that prompt minimal reactivity from participants and assess sitting in a way that corresponds most closely to perceptions of seated activity.

The accuracy of self-report depends on researchers and participants sharing interpretations of questionnaire items. Yet, we documented important misunderstandings and perceived ambiguities in item wording and showed that responses to such items may lack validity. Perhaps unsurprisingly, where participants were unsure as to the meaning of an item, they were uncertain as to how to respond. Importantly, this could lead participants with identical sitting patterns to respond differently to the same item. Analogously, a previous ‘think aloud’ study showed that, in response to Likert scales assessing levels of agreement with a statement, some participants who did not understand the statement selected ‘strongly disagree’, whereas others chose to ‘neither agree nor disagree’ (34). Where participants within a sample differ in their perceptions of the meaning of items, their responses will lack validity and will be unsuitable for pooling with others’ data. Researchers must ensure that instructions and definitions are made clear and should consider adopting ‘think aloud’ methods to identify interpretation problems prior to questionnaire administration.

We also documented participants’ concerns about portraying themselves negatively by disclosing lengthy periods of sitting and showed that this could lead to under-reporting. Such self-presentation concerns may at least partly explain why self-report measures typically underestimate true sitting time (42,43), and indeed the broader phenomenon whereby people under-report engagement in health-risk behaviours and over-report engagement in health-promoting actions (44). This is particularly problematic given the prevalence of self-report measures for population-level assessment of sedentary behaviour (45-47), which shapes policy and practice (48). While there is evidence to suggest that some individuals consistently misreport their behaviour (44), questionnaire developers might mitigate this problem by emphasising the anonymity of responses (49).

A more fundamental problem highlighted by our data was that participants appeared not to mentally represent time spent sitting as ‘sitting time’. This echoes research demonstrating that people tend to encode sitting into memories of more meaningful acts performed while seated (50). Estimating sitting time required participants to engage in a burdensome and error-prone process of converting seated activities into global sitting time. This problem likely compounds other sources of bias associated with autobiographical recollection, such as memory decay, whereby people fail to remember relevant action episodes (51), and longer intervals between action and recall produce more erroneous responses (52). Actions that are personally meaningful are more likely to be recalled, and are recalled more quickly, than those that lack meaning (53,54), because they are more central elements within the cognitive structures that people use to organise and process information (55). By extension, self-report items may elicit more accurate information where they focus on activities as people perceive them. Activity-specific sitting measures might therefore alleviate burden and reduce error. Measures that estimate total sitting by summing across items assessing multiple contexts (e.g. the MSTQ: “indicate the number of hours and minutes that you spend…”; “sitting while watching TV or movies”, “sitting using a computer of video game”; [11]) may yield more valid data than direct recall items (e.g. IPAQ: ‘how much time did you usually spend sitting on a weekday?’ [10]).

The mental calculation process required to estimate sitting time raised participants’ awareness of excessive sitting time, inadvertently motivating them to change their behaviour. Intervention studies have demonstrated that raising awareness of sitting, albeit via feedback received from others, can prompt people to reduce their sedentary behaviour (56). Ours is the first study to demonstrate that feedback can effectively be self-administered: while participants were able to recall time in seated activities, combining discrete events into an unfamiliar metric corresponding to total sitting time provided novel and potentially motivational information. Such ‘mere measurement effects’, whereby reporting thoughts or actions leads participants to modify their behaviour, have been documented in other domains (57), and are thought to arise because self-reporting makes salient the mismatch between current and ideal behaviour (58). This poses particular problems for research contexts in which sitting is measured over multiple timepoints: attempts to observe natural fluctuations in sitting over time, or to detect effects of purposive interventions, would be contaminated if measurement alone encouraged sitting reduction. Conversely however, future work might investigate the potential for self-reporting sitting time to be used for intervention purposes.

Limitations of our study must be acknowledged. The act of ‘thinking aloud’ can modify attention and motivation such that questionnaire responses provided while thinking aloud may not resemble those that would otherwise be generated (18,59). Participants may have attended more closely than normal to items, potentially prompting more considered, deliberative responses. The knowledge that responses were being recorded by sedentary behaviour researchers may have prompted socially desirable responding, such as displaying heightened dismay at excessive sitting time. However, we sought to minimise such reactivity by instructing participants to provide descriptive rather than self-evaluative responses, conducting a practice task in the presence of the researcher to verify adherence to and understanding of instructions, and having participants complete the main ‘think aloud’ procedure alone, steps which can minimise reactive responding (18). We recruited a small sample of mostly female office workers, whose experiences and reflections may not generalise to other populations. Our aim, however, was not to provide an exhaustive or generalisable account, but rather to illustrate the thought patterns that may underpin questionnaire responses, and there is no reason to expect our participants’ thought processes to differ from those among others.

Additional limitations arise from our focus on sitting time, rather than sedentary behaviour more broadly. Critical commentators have highlighted the ableist assumption that sitting is not only synonymous with sedentary behaviour, but is also uniformly undesirable, as this excludes people for whom sitting is unavoidable due to disability, health or wellbeing needs (60). While often treated as synonymous with sedentary behaviour, sitting is a sub-type of sedentary behaviour, as defined as activity requiring ≤1.5 metabolic equivalents undertaken while seated or lying down (61). Not all sedentary behaviour involves sitting, and many seated activities (e.g., playing active computer games) are not sedentary, incurring energy expenditure above 1.5 metabolic equivalents (62). Physical activity guidelines increasingly focus on sedentary behaviour, rather than sitting per se (48), and it is likely that sedentary behaviour questionnaires will also evolve in response to these concerns, rendering sitting-focused questionnaires obsolete. However, our findings likely transcend sitting and reflect broader problems when self-reporting time spent in sedentary activity more broadly. For example, like sitting, people tend not to recognise ‘sedentary behaviour’ as a meaningful standalone activity, so are likely to undertake a mental calculation process when generating estimates of sedentary behaviour (63). In retrospect, a more fundamental limitation of our study is that we excluded people with disabilities that necessitating prolonged sitting, a population which could contribute most insight into the inclusivity of sedentary behaviour questionnaires (60). Developing a comprehensive understanding of responses to sedentary behaviour questionnaires will depend on adopting inclusive recruitment methods and procedures.

Our study points to potential difficulties affecting self-reported sitting time, most notably due to problems encountered when attempting to translate time in seated activities into total sitting time. Researchers should anticipate these problems when developing or administering questionnaires to measure sedentary behaviour.

**References**

1. Biddle SJH, Bennie JA, Bauman AE, Chau JY, Dunstan D, Owen N, et al: Too much sitting and all-cause mortality: is there a causal link? *BMC Public Health.* 2016, 26:635.

2. de Rezende LFM, Lopes MR, Rey-López JP, Matsudo VKR, Luiz O d. C: Sedentary behavior and health outcomes: An overview of systematic reviews. *PLoS ONE.* 2014, 9:e105620.

3. Smith L, Hamer M, Gardner B: Sedentary Behavior and Mental Health. In: Stubbs B, Rosenbaum S, editors. *Exercise-Based Interventions for Mental Illness.* Academic Press; 2018. p. 107–19.

4. Gardner B, Smith L, Lorencatto F, Hamer M, Biddle SJH: How to reduce sitting time? A review of behaviour change strategies used in sedentary behaviour reduction interventions among adults. *Health Psychol Rev*. 2016, 10: 89–112.

5. Smith L, Hamer M, Ucci M, Marmot A, Gardner B, Sawyer A, et al: Weekday and weekend patterns of objectively measured sitting, standing, and stepping in a sample of office-based workers: the active buildings study. *BMC Public Health*. 2015, 15:9.

6. Ryde GC, Brown HE, Gilson ND, Brown WJ: Are we chained to our desks? Describing desk-based sitting using a novel measure of occupational sitting. *J Phys Act Health.* 2014, 11: 1318–23.

7. Dall PM, Coulter EH, Fitzsimons CF, Skelton DA, Chastin SFM: TAxonomy of Self-reported Sedentary behaviour Tools (TASST) framework for development, comparison and evaluation of self-report tools: content analysis and systematic review. *BMJ Open.* 2017, 7: e013844.

8. Rivière F, Aubert S, Omorou AY, Ainsworth BE, Vuillemin A: Taxonomy-based content analysis of sedentary behavior questionnaires: A systematic review. *PLoS One.* 2018, 13: e0193812.

9. Chastin SFM, Schwarz U, Skelton DA: Development of a Consensus Taxonomy of Sedentary Behaviors (SIT): Report of Delphi Round 1. *PLoS One.* 2013, 8: e82313.

10. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al: International Physical Activity Questionnaire: 12-Country Reliability and Validity. *Med Sci Sports Exerc.* 2003, 35: 1381–95.

11. Whitfield GP, Gabriel KKP, Kohl HW: Assessing sitting across contexts: development of the multicontext sitting time questionnaire. *Res Q Exerc Sport.* 2013, 84: 323–8.

12. van Uffelen JG, Heesch KC, Hill RL, Brown WJ: A qualitative study of older adults’ responses to sitting-time questions: do we get the information we want? *BMC Public Health*. 2011, 11: 458.

13. Clark BK, Thorp AA, A. H. Winkler E, Gardiner PA, Healy GN, Owen N, et al: Validity of self-reported measures of workplace sitting time and breaks in sitting time. *Med Sci Sports Exerc.* 2011, 43: 1907–1912.

14. Gupta N, Christiansen CS, Hanisch C, Bay H, Burr H, Holtermann A. Is questionnaire-based sitting time inaccurate and can it be improved? A cross-sectional investigation using accelerometer-based sitting time. *BMJ Open.* 2017, 7: e013251.

15. Lee PH, Macfarlane DJ, Lam TH, Stewart SM: Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): A systematic review. *Int J Behav Nutr Phys Act.* 2011, 8: 115.

16. Rääsk T, Mäestu J, Lätt E, Jürimäe J, Jürimäe T, Vainik U, et al: Comparison of IPAQ-SF and two other physical activity questionnaires with accelerometer in adolescent boys. *PLoS One.* 2017, 12: e0169527.

17. French DP, Cooke R, McLean N, Williams M, Sutton S: What do people think about when they answer theory of planned behaviour questionnaires? A ‘think aloud’ study. *J Health Psychol*. 2007, 12: 672–87.

18. Ericsson KA, Simon HA. *Protocol analysis: Verbal reports as data*. Cambridge: MIT Press; 1993.

19. Eccles, D. W. Verbal reports of cognitive processes. In G. Tenenbaum, R. C. Eklund, & A. Kamata (Eds.), *Measurement in sport and exercise psychology.* Champaign, IL: Human Kinetics; 2012. pp. 103-117.

20. Eccles, D. W., & Arsal, G. The think aloud method: What is it and how do I use it? *Qual Res Sport Exerc Health*. 2017. 9:514-531.

21. Calmeiro L, Tenenbaum G. Concurrent verbal protocol analysis in sport: Illustration of thought processes during a golf-putting task. *J Clin Sport Psychol.* 2011. 5:223-36.

22. Whitehead AE, Taylor JA, Polman RC. Examination of the suitability of collecting in event cognitive processes using Think Aloud protocol in golf. *Front Psychol.* 2015. 6:1083.

23. Stone AA, Schwartz JE, Neale JM, Shiffman S, Marco CA, Hickcox M, Paty J, Porter LS, Cruise LJ. A comparison of coping assessed by ecological momentary assessment and retrospective recall. *J Pers Soc Psychol.* 1998. 74:1670-80.

24. Samson A, Simpson D, Kamphoff D, Langlier A. Think aloud: An examination of distance runners’ thought processes. *Int J Sport Exerc Psychol.* 2017. 15:176-89.

25. Whitehead AE, Jones HS, Williams EL, Rowley C, Quayle L, Marchant D, Polman RC. Investigating the relationship between cognitions, pacing strategies and performance in 16.1km cycling time trials using a think aloud protocol. *Psychol Sport Exerc.* 2018. 34:95-109.

26. Whitehead AE, Taylor JA, Polman RC. Evidence for skill level differences in the thought processes of golfers during high and low pressure situations. *Front Psychol.* 2016. 6:1974.

27. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *J Adv Nurs.* 2003. 42:57-63.

28. Washburn RA, Smith KW, Jette AM, Janney CA: The physical activity scale for the elderly (PASE): Development and evaluation. *J Clin Epidemiol*. 1993, 46: 153–62.

29. Guba EG, Lincoln YS. Competing paradigms in qualitative research. In Denzin NK, Lincoln YS, editors. *Handbook of qualitative research*. Thousand Oaks, CA: Sage; 1994. pp. 105-117.

30. Matei R, Thuné-Boyle I, Hamer M, Iliffe S, Fox KR, Jefferis BJ, Gardner B. Acceptability of a theory-based sedentary behaviour reduction intervention for older adults ('On Your Feet to Earn Your Seat'). *BMC Public Health.* 2015. 15:606.

31. Gardner B, Tang V. Reflecting on non-reflective action: An exploratory think-aloud study of self-report habit measures. *Brit J Health Psychol.* 2014. 19:258-73.

32. Pedisic Z, Bennie JA, Timperio AF, Crawford DA, Dunstan DW, Bauman AE, et al: Workplace Sitting Breaks Questionnaire (SITBRQ): an assessment of concurrent validity and test-retest reliability. *BMC Public Health.* 2014, 14: 1249.

33. Clemes SA, David BM, Zhao Y, Han X, Brown W. Validity of two self-report measures of sitting time. *J Phys Act Health.* 2012, 9: 533–9.

34. Pettee KK, Ham SA, Macera CA, Ainsworth BE: The reliability of a survey question on television viewing and associations with health risk factors in US adults. *Obes*. 2009, 17: 487–93.

35. Marshall AL, Miller YD, Burton NW, Brown WJ: Measuring total and domain-specific sitting: a study of reliability and validity. *Med Sci Sports Exerc*. 2010, 42: 1094–102.

36. Jancey J, Tye M, McGann S, Blackford K, Lee AH. Application of the Occupational Sitting and Physical Activity Questionnaire (OSPAQ) to office based workers. *BMC Public Health.* 2014, 14: 762.

37. Aadahl M, Jørgensen T. Validation of a New Self-Report Instrument for Measuring Physical Activity. *Med Sci Sports Exerc.* 2003, 35: 1196–1202.

38. Bonn SE, Lagerros YT, Christensen SE, Möller E, Wright A, Sjölander A, et al: Active-Q: Validation of the web-based physical activity questionnaire using doubly labeled water. *J Med Internet Res*. 2012, 14: e29.

39. Fox MC, Ericsson KA, Best R. Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychol Bull.* 2011. 137:316–44.

40. Braun V, Clarke V: Using thematic analysis in psychology. *Qual Res Psychol.* 2006, 3: 77–101.

41. Smith BR, McGannon KR: Developing rigor in qualitative research: problems and opportunities within sport and exercise psychology. *Int Rev Sport Exerc Psychol.* 2018. 11:101-21.

42. Affuso O, Stevens J, Catellier D, McMurray RG, Ward DS, Lytle L, Sothern MS, Young DR. Validity of self-reported leisure-time sedentary behavior in adolescents. *J Negat Results BioMed*. 2011. 10:2.

43. Gennuso KP, Matthews CE, Colbert LH. Reliability and validity of 2 self-report measures to assess sedentary behavior in older adults. *J Phys Act Health.* 2015. 12:727-32.

44. Black AE, Cole TJ. Biased over- or under-reporting is characteristic of individuals whether over time or by different assessment methods. *J Am Diet Assoc.* 2001. 101:70-80.

45. Bennie JA, [Chau JY](https://www.ncbi.nlm.nih.gov/pubmed/?term=Chau%20JY%5BAuthor%5D&cauthor=true&cauthor_uid=24020702), [van der Ploeg HP](https://www.ncbi.nlm.nih.gov/pubmed/?term=van%20der%20Ploeg%20HP%5BAuthor%5D&cauthor=true&cauthor_uid=24020702), [Stamatakis E](https://www.ncbi.nlm.nih.gov/pubmed/?term=Stamatakis%20E%5BAuthor%5D&cauthor=true&cauthor_uid=24020702), [Do A](https://www.ncbi.nlm.nih.gov/pubmed/?term=Do%20A%5BAuthor%5D&cauthor=true&cauthor_uid=24020702), [Bauman A](https://www.ncbi.nlm.nih.gov/pubmed/?term=Bauman%20A%5BAuthor%5D&cauthor=true&cauthor_uid=24020702). The prevalence and correlates of sitting in European adults - a comparison of 32 Eurobarometer-participating countries. *Int J Behav Nutr Phys Act.* 2013. 10:107.

46. Loyen A, van der Ploeg HP, Bauman A, Brug J, Lakerveld J. European Sitting Championship: Prevalence and correlates of self-reported sitting time in the 28 European Union member states. *PLoS ONE.* 2016. 11:e0149320.

47. Milton K, Gale J, Stamatakis E, Bauman A. Trends in prolonged sitting time among European adults: 27 country analysis. *Prev Med.* 2015. 77:11-16.

48. UK Chief Medical Officers. UK Chief Medical Officers' Physical Activity Guidelines. London: Department of Health; 2019.

49. Nederhof AJ: Methods of coping with social desirability bias: A review. *Euro J Soc Psychol*. 1985, 15: 263–80.

50. Gardner B, Flint S, Rebar AL, Dewitt S, Quail SK, Whall H, et al: Is sitting invisible? Exploring how people mentally represent sitting. *Int J Behav Nutr Phys Act.* 2019, 16: 85.

51. Smith ER: Mental representation and memory. In Gilbert DT, Fiske ST, Lindzey G, editors. *Handbook of Social Psychology Vol 1* (pp391-445). New York: Oxford University Press; 1998.

52. Durante R, Ainsworth BE. The recall of physical activity: using a cognitive model of the question-answering process. *Med Sci Sport Exerc.* 1996. 28:1282-91.

53. Ellis HC, Parente FJ, Shumate EC: Meaningfulness, perceptual grouping, and organization in recognition memory. *J Exp Psychol.* 1974.102: 308–13.

54. Markus H, Hamill R, Sentis KP. Thinking fat: Self-schemas for body weight and the processing of weight relevant information. *J Appl Soc Psychol.* 1987. 17:50-71.

55. Ghosh VE, Gilboa A. What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia.* 2014. 53:104-14.

56. Dewitt S, Hall J, Smith L, Buckley JP, Biddle SJH, Mansfield L, et al. Office workers’ experiences of attempts to reduce sitting-time: an exploratory, mixed-methods uncontrolled intervention pilot study. *BMC Public Health.* 2019, 19: 819.

57. Godin G, Sheeran P, Conner M, Germain M. Asking questions changes behavior: mere measurement effects on frequency of blood donation. *Health Psychol*. 2008, 27:179–84.

58. Silvia PJ, Duval TS. Objective self-awareness theory: recent progress and enduring problems. *Pers Soc Psychol Rev.* 2001. 5:230–41.

59. Double KS, Birney DP. Reactivity to measures of metacognition. *Front Psychol.* 2019. 10:2755*.*

60. Todd KR, Martin Ginis KA. Physical activity and spinal cord injury: Lessons learned at the lowest end of the physical activity spectrum. *Kinesiol Rev.* 2019. 8: 54-62.

61. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, et al. Sedentary Behaviour Research Network (SBRN) – Terminology Consensus Project process and outcome. *Int J Behav Nutr Phys Act.* 2017. 14:75.

62. Mansoubi M, Pearson N, Clemes SA, Biddle SJH, Bodicoat DH, Tolfrey K, et al. Energy expenditure during common sitting and standing tasks: examining the 1.5 MET definition of sedentary behaviour. *BMC Public Health.* 2015. 15:516.

63. Hadgraft NT, Brakenridge CL, Dunstan DW, Owen N, Healy GN, Lawler SP: Perceptions of the acceptability and feasibility of reducing occupational sitting: review and thematic synthesis. *Int J Behav Nutr Phys Act.* 2018. 15:90.

**Table**. Features of questionnaire items included in study questionnaire, coded according to relevant characteristics in Rivière et al’s taxonomy (8)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Questionnaire** | **Recall period** | **Number of sitting-related items** | **Dimensions** | | **Indicators** |
|  |  |  | **Time** | **Breaks** |  |
| International Physical Activity Questionnaire (IPAQ; short-form) (10) | 7 days | 2 | h-m/day | - | Duration |
| Workplace Sitting Breaks Questionnaire (32) | Typical workday | 2 | - | # breaks/h, total time of break during the day at work | # breaks |
| Clemes Single Item Sitting Time Questionnaire (33) | Today | 1 | h-m/day | - | Duration |
| Survey Question on Television Viewing (34) | Typical week | 1 | h-m/day | - | Duration |
| Marshall Domain-Specific Sitting Time Questionnaire (35) | Typical weekday and typical weekend day | 10 | h-m/day | - | Duration |
| Occupational Sitting and Physical Activity Questionnaire (36) | Typical workday in last 7 days | 3 | %, h-m/day | - | Duration |
| Aadahl and Jorgensen Physical Activity Scale (37) | Average weekday | 4 | h-m/day | - | Duration |
| Multi-context Sitting Time Questionnaire (11) | Typical workday and typical non-work day | 14 | h-m/day | - | Duration |
| Active-Q (38) | Typical day | 6 | h-m/day | - | Duration |