A new single item question for assessment of daily sitting time

- Criterion validity and test-retest reliability, with ActivPal as the criterion measure

Kristina Larsson
Abstract

Aim and research questions
The aim of this study was to investigate the criterion validity and test-retest reliability of the new single item question SED-GIH, with ActivPal as the criterion measure. Research questions: How strong is the level of agreement between sitting time measured subjectively with SED-GIH and objectively with ActivPal? How strong is the reliability of SED-GIH measured with test-retest?

Method
The validity section was a part of the large research project “Physical Activity and Healthy Brain Functions”. Participants answered a web questionnaire, completed a test-session, wore activity monitors and filled in a sleep and activity diary for one week. The criterion validity was assessed by the agreement between SED-GIH (a part of the web questionnaire) and the activity monitor ActivPal. Spearman’s correlation, Wilcoxon analysis and Bland Altman plot was conducted. A one-way independent ANOVA assessed differences between the categorical groups of SED-GIH and the objective values of sitting time from ActivPal. The reliability section was a part of the “Health Project” at GIH. The participants answered SED-GIH two times with a few weeks interval. The reliability was assessed with ICC, Wilcoxon analysis and a Bland Altman plot was conducted.

Results
The correlation between SED-GIH and ActivPal was moderate with Spearman’s rho = 0.31. Wilcoxon analysis showed significant differences between SED-GIH and ActivPal with p < 0.01. Bland Altman plot indicated that participants both under and overestimated their sitting time. One-way independent ANOVA between SED-GIH and ActivPal resulted in a significant p < 0.01. Tukey Post Hoc test presented that significant difference existed between some of the categorical answering options. ICC for the test-retest reliability of the SED-GIH was excellent with ICC = 0.86. Wilcoxon analysis showed non-significant differences with p = 0.12. The Bland Altman plot indicates that participants were consistent in the way they answered SED-GIH at test-session one and test-session two.

Conclusions
SED-GIH both over and underestimated sitting time, however a general slightly underestimation occurred. The validity and reliability of SED-GIH are in line, or even better, than other questions that are common to use when assessing sitting time. This make SED-GIH useful in many areas, like in the healthcare system and in national public health surveys.
Sammanfattning

Syfte och frågeställningar

Metod

Resultat

Slutsats
SED-GIH både under- och överskattade sittande tid, dock med en generell lätt underskattning. Validiteten och reliabiliteten för SED-GIH är i linje, eller till och med bättre, än andra frågor som är vanliga att använda för att ta reda på sittande tid. Detta gör SED-GIH användbar inom många områden, exempelvis i hälsosjukvården eller i nationella folkhälsounsundersökningar.
## Table of content

1 Background ........................................................................................................ 1  
   1.1 Sedentary behaviour ..................................................................................... 1  
      1.1.1 Our sedentary lifestyle ......................................................................... 1  
      1.1.2 Definitions ......................................................................................... 1  
      1.1.3 Health risks ....................................................................................... 2  
      1.1.4 Prevalence of sedentary and sitting time ............................................ 3  
1.2 Measurement methods of sedentary behaviour ............................................ 3  
   1.2.1 Objective measurement methods ............................................................. 3  
   1.2.2 Subjective measurement methods ............................................................ 5  
   1.2.3 Questionnaires validated with ActivPal ................................................... 6  
3 Aim ...................................................................................................................... 7  
4 Method .................................................................................................................. 7  
   4.1 Study design .................................................................................................. 7  
   4.2 Participants and sampling ............................................................................ 7  
      4.2.1 Ethical considerations .......................................................................... 8  
   4.3 Test procedures ............................................................................................ 8  
   4.4 Outcome measures ....................................................................................... 10  
      4.4.1 SED-GIH .............................................................................................. 10  
      4.4.2 ActivPal ............................................................................................... 10  
      4.4.3 Sleep and activity diaries ..................................................................... 10  
   4.5 Data analysis .................................................................................................. 11  
   4.6 Statistical analysis ......................................................................................... 11  
      4.6.1 Criterion validity .................................................................................. 12  
      4.6.2 Test-retest reliability ............................................................................ 12  
5 Results .................................................................................................................. 12  
   5.1 Participants characteristics .......................................................................... 12  
      5.1.1 "Physical Activity and Healthy Brain Functions" project ....................... 12  
      5.1.2 "Health Project" at GIH ....................................................................... 14  
   5.2 Criterion validity ........................................................................................... 14  
      5.2.1 Level of agreement between ActivPal and SED-GIH ......................... 14  
      5.2.2 Difference between the categories of SED-GIH ................................... 17  
   5.3 Test-retest reliability ..................................................................................... 18
6 Discussion ........................................................................................................................................... 20
6.1 Level of agreement between sitting time measured subjectively with SED-GIH and objectively with ActivPal ........................................................................................................................................ 20
6.2 How the objective sitting time measured with ActivPal varies depending on the subjective estimations measured with SED-GIH .................................................................................................................................. 21
6.3 Reliability of SED-GIH measured with test-retest ........................................................................ 22
6.4 Method discussion .......................................................................................................................... 22
6.5 Implications ..................................................................................................................................... 23
6.6 Conclusion ........................................................................................................................................ 24
References ............................................................................................................................................... 25

Appendix 1 – Literature search
Appendix 2 – Sleep and activity diary

Table and figure list

Figure 1 – Flow-chart of the test procedure ......................................................................................... 9
Figure 2 – Flow-chart of the number of participants ........................................................................... 13
Figure 3 – Correlation between SED-GIH and ActivPal .................................................................... 15
Figure 4 – Bland Altman Plot for the agreement between SED-GIH and ActivPal ......................... 16
Figure 5 – Correlation between test-session one and test-session two ............................................. 18
Figure 6 – Bland Altman Plot for test-retest of SED-GIH ................................................................. 19

Table 1 – Codes for the different categorical answering options of SED-GIH ........................... 11
Table 2 – Characteristics of participants who were included in the “Physical Activity and Healthy Brain Function” project ........................................................................................................... 13
Table 3 – Characteristics of the sample from the “Health Project” .................................................. 14
Table 4 – The answer options that had significant differences in hours of sitting measured with ActivPal between the groups ........................................................................................................... 17
Table 5 – The disposition of participants answers of SED-GIH and their mean hours of sitting per day measured with ActivPal ........................................................................................................... 17
1 Background

1.1 Sedentary behaviour

1.1.1 Our sedentary lifestyle

Since the middle of the last century the modern human lifestyle has changed rapidly. New habits regarding transportation, communications, workplace and domestic-entertainment technologies are reducing our daily movement, which results in more sedentary time. (Owen et al., 2010) To achieve a better understanding concerning the determinants of sedentary behaviour Owen and colleagues created an ecological model. The core components of the model are based on the sedentary behaviour domains: leisure time, household, occupation and transport. Sitting occurs in all these domains and is controlled by different factors such as attitudes, motivation and norms. There are strong existing norms that expect us to sit in classes, in meetings and at home while resting. These norms are implemented both by environmental and social contexts, such as questioning why someone is standing at a meeting, or providing employees with the opportunity to stand and work at the office. (2011)

1.1.2 Definitions

Behaviours defined as sedentary include both sitting and low intensity activities. The unit metabolic equivalent, MET, assess the levels of energy expenditure for different activities, where 1 MET is equal to sitting quietly. Sedentary behaviour is defined as ≤ 1.5 METs. (Ainsworth et al., 1993) Thus, there are differences between the definitions sedentary behaviour, sedentary time and sitting time. In this paper all terms will be used depending on the context.

It is important to distinguish between sedentary behaviour and physical inactivity. Physical inactivity can be defined as non-compliance to the guidelines for physical activity, while sedentary behaviour can be described as absence of muscular contractions. (World Health Organization, 2010) A person can be sufficiently active according to the physical activity guidelines but still spend too much time in prolonged sitting, also called The Active Couch Potato phenomenon (Owen et al., 2011; Owen et al., 2010). No generally agreed recommendations for sedentary behaviour exist, since we still lack sufficient evidence regarding how harmful specific amounts of sedentary time is for specified health outcomes.
1.1.3 Health risks

Sedentary behaviour is increasingly being recognized as a health risk, even if the recommendations for physical activity are fulfilled. Therefore, it is of big importance to understand the associations of sedentary behaviour and health risks, independent on physical activity level. To measure sedentary behaviour it is common to use both subjective and objective methods. The most commonly used objective methods are accelerometers, inclinometers, heart rate monitoring, and combined multi-unit monitors. The subjective ones are mainly self- and proxy-report questionnaires, behaviour logs, short-term recalls and diaries. (Atkin et al., 2012; Healy et al., 2011)

In a meta-analysis Biswas and colleagues concluded that adults’ sedentary time was associated with a greater risk for all-cause mortality, cancer incidence or mortality, cardiovascular disease incidence or mortality and type 2 diabetes. The result was the same whether participant’s physical activity level was high or low. Those who participated in higher levels of physical activity did however show lower deleterious outcome effects associated with sedentary time. All included studies except one used self-reported methods. (2015) Another meta-analysis, based on studies using subjective or objective measurement methods, found that each additional hour of daily sitting increased the risk for all-cause mortality by 2 %, after adjusting for levels of physical activity. Noticeably, for individuals sitting more than seven hours per day, the risk for all-cause mortality increased by 5 % for each additional hour of daily sitting. For individuals sitting 10 hours per day, the risk for all-cause mortality increased with as much as 34 % in total. (Chau et al., 2013)

A significant risk for cardiovascular disease was observed in individuals who had very high levels of sedentary time (> 10 hours per day) after adjustment for physical activity and other cardiovascular disease risk factors. For individuals with intermediate levels of sedentary time no such association was found. (Pandey et al., 2016) Thus, for promoting metabolic health, both increasing time spent in physical activity and decreasing sedentary time seems of importance. Based on accelerometry, significant associations have been found between sedentary time and waist circumference and clustered metabolic risk, independent of time spent in moderate to vigorous intense of physical activity. (Healy et al., 2008b) Bankoski and colleagues assessed sedentary behaviour with accelerometry. They found that participants with lower HDL cholesterol, larger waist circumference, higher triglycerides and the metabolic syndrome had higher percentage of sedentary time and lower number of breaks in prolonged sitting (2011). Breaking up prolonged periods of sitting could have positive effects
on metabolic outcomes. Measured with accelerometry, the dose-response effect of the number of breaks in sedentary behaviour still remains unknown, just as the effects of prolonged versus interrupted sitting time. (Chastin et al., 2015; Benatti & Ried-Larsen, 2015) Increasing the number of breaks in sedentary time is positively associated with waist circumference, BMI, triglycerides and 2-h plasma glucose (Healy et al., 2008a).

1.1.4 Prevalence of sedentary and sitting time

Based on subjective measures, European adults spend on average five hours per day sitting, though large between country differences exists. Sweden represents one of the countries with the highest amount of reported sitting time, where almost 30 % of the population sit more than 7 hours per day. In Portugal only 11 % of the population sit more than 7 hours per day. (Milton et al., 2015) Based on accelerometry, a sample of Swedes residing in Gothenburg, aged from 50 to 64 years, spent on average 8.7 hours sedentary (60.5 %) based on a daily median wear time of 14.3 hours per day (Ekblom-Bak et al., 2015). A review including studies with both subjective and objective measurement methods from ten different countries, present that older adults are sedentary about 9.4 hours per day, corresponding 65-80 % of their waking day (Harvey et al., 2015).

1.2 Measurement methods of sedentary behaviour

Since the prevalence of sedentary and sitting time is high and sedentary behaviour increasingly is being recognized as a health risk, there is of big importance to have tools for assessment of sitting time. At present both objective and subjective measurement methods are common to use.

1.2.1 Objective measurement methods

1.2.1.1 Direct observation

When assessing sedentary behaviour it is important to use measurement methods that are validated and have high accuracy. The most accurate method and gold standard for physical activity and sedentary behaviour is energy expenditure including direct or indirect calorimetry and doubly labelled water. Since this method is limited for use in free-living settings, an alternative method could be direct observation. Therefore, Lyden and colleagues compared direct observation estimations of time in activity intensity categories and total MET-hours, to a criterion measure of indirect calorimeter. The study provides evidence that in free-living
conditions direct observation (by trained observers) can be used as a criterion measure of sedentary behaviour and physical activity. (Lyden et al., 2014)

1.2.1.2 ActiGraph

Since direct observation can be difficult to use when observing a big group of people in their free-living conditions, other methodological devices have been developed. An accelerometer used for measuring physical activity and sedentary behaviour is a hip- or wrist-worn monitor, which detects body movement by using the unit counts per minute (cpm). A count is calculated from the amplitude and frequency of the acceleration. The accelerations are being measured in three axis, X-, Y-, and Z-axis, in three different plans, vertical, horizontal and lateral. One common accelerometer is ActiGraph GT3X+, which consists of both an accelerometer and an inclinometer in the same unit. (Peterson et al., 2015) ActiGraph has provided valid and reliable measurement when assessing physical activity (Trost et al., 1998). When assessing sedentary time with ActiGraph it is common to use the cut point < 100 cpm. Since ActiGraph is worn on the hip it is unable to detect postures, such as sitting and standing. The combination of the hip placement, and the cut point < 100 cpm often results in inaccurate assessment of standing low intensity activities (below 100 cpm), such as folding laundry and washing dishes, which will be inaccurate detected as sedentary time. Therefore, there is still much unknown about the ability of the ActiGraph to distinguish between light intensity physical activity and sedentary time. (Kozey et al., 2010; Kozey-Keadle et al., 2011)

1.2.1.3 ActivPal

ActivPal is also an accelerometer with an inclinometer. Compared to ActiGraph it is worn on the midline of the anterior aspect of the right thigh, with adhesive tape. It has a triaxial accelerometer with a high sampling frequency and identifies episodes of walking, sitting and standing. Unique for ActivPal is the ability to detect the body posture (sitting and standing) and postural transition, due to its position on the thigh. (Klenk et al., 2016; Grant et al., 2006) The validity of ActivPal is high compared with direct observation. An overall agreement between observation with digital camera and ActivPal was 95.9 %, when testing in a laboratory setting for 15-20 minutes on each session (Grant et al., 2006). Kim and colleagues used the same method and their result showed good validity and error from the absolute value was 4.11 %. (2015) When ActivPal was validated against direct observation with personal digital assistant as appliance, the result showed almost perfect correlations for sedentary time, absolute number of breaks and the break-rate (Lyden et al., 2012; Kozey-Keadle et al., 2011).
The correlation between ActivPal, when compared with actual METs from VO\textsubscript{2} as the criterion measure, was excellent with $r = 0.93$. When classifying sitting, standing and slow walking the levels of agreement between ActivPal and direct observation (recorded activity categories from a single observer) was 99.1 %. (Dowd et al., 2012) Since ActivPal has good validity it is often used as a criterion measurement when validating other objective and subjective measurement methods for sedentary behaviour (Van Nassau et al., 2015). When assessing the reliability for ActivPal Grant and colleagues allowed participants to wear three monitors simultaneously. The intraclass correlation for the interdevice reliability ranged from 0.79-0.99. (2006)

1.2.2 Subjective measurement methods

The most common subjective measurement methods for assessing sedentary behaviour are self- and proxy-report questionnaires, behaviour logs, short-term recall and diaries. Generally, these methods have low costs, are easy to implement on a large scale and have low burden both for the participants and the researcher. However, there are some general limitations considering recall and reporting bias, random and systematic reporting errors and low validity. (Atkin et al., 2012; Healy et al., 2011) Questionnaires for measuring sedentary behaviour can vary depending on which main focus they have. One of the most common is IPAQ (International Physical Activity Questionnaire), which assesses both physical activity and sedentary behaviour. The sitting question in IPAQ has moderate reliability and the validity is low to moderate, compared to accelerometer data (Craig et al., 2003). IPAQ tends to overestimate physical activity and underestimate sedentary time, with accelerometer as criterion measure (Van Dyck et al., 2015). Different fixed answer options (almost non of the time, one fourth of the time, half of the time, three fourths of the time and almost all of the time) is another way of asking about sitting time (Katzmarzyk et al., 2009). Sedentary behaviour both in prolonged time periods and on a regular basis, such as watching TV or sitting at work, in general have stronger reliability, compared to less regular sitting behaviour such as travel or other kind of sitting (Healy et al., 2011). Marshall and colleagues used log data and accelerometer as a reference method, to investigate sitting time in different domains (traveling, at work, watching TV, using a computer at home and leisure time excluding watching TV), both in weekdays and weekend days. The reliability was high in weekdays for time at work, watching TV and using a computer at home, but low in all domains for the weekend days. The validity was highest for the domains time at work and using a computer at home, for weekdays. In general the validity was low for the weekend. (2010) TV-viewing
time is a common domain for assessing sedentary behaviour, since it often occurs on regular basis in prolonged time periods. Therefore, TV-viewing time has good reliability, though the validity in different studies report large differences. (Clark et al., 2009)

1.2.2.1 A new single item question for assessment of daily sitting time

The Swedish government agency The National Board of Health and Welfare (Socialstyrelsen) is continuously working to develop National guidelines for disease preventing methods. One part of the guidelines means that healthcare professionals should offer consultation to everyone with insufficient levels of physical activity. It is important for The National Board of Health and Welfare to evaluate these guidelines. Evaluation requires validated indication questions regarding physical activity. Another important area of usage for this kind of question is to investigate in what proportion the Swedish population fulfill the recommendations for physical activity. The National Board of Health and Welfare have developed two questions regarding physical activity level. Both questions can be answered in three different ways, categorical, open or table options. Olsson and colleagues validated these different answering options with accelerometer as the criterion measure. The results suggest that questions regarding physical activity have better validity when categorical answers are used. (2016)

Since sedentary behavior is of big importance in this context, the new single item question called SED-GIH, regarding sitting time has been developed. This question has categorical answering options and is based on other sedentary questions with high validity. The aim was to develop these existing questions into one new that is easy to use in the healthcare to detect those who spend too much time sitting. Kallings and colleagues validated SED-GIH and three other sedentary questions with accelerometer as the criterion measure. SED-GIH had the strongest validity, although the differences between the questions were small. (Kallings et al., 2014) Since 2016 SED-GIH is included in the Swedish national public health survey, Health on Equal Terms (Folkhälsomyndigheten, 2016).

1.2.3 Questionnaires validated with ActivPal

The usage of ActivPal as criterion measurement when assessing validity for a questionnaire is common. When Chastin and colleagues validated sitting time by IPAQ with ActivPal as criterion measure, they found low correlations between the two methods where IPAQ underestimated sitting time with 2.2 hours per day for a total week and 4.6 hours per day during weekend days, when including transportation (2014). PAST (Past-day Adults
Sedentary Time) and PAST-U (modified version of PAST) are questionnaires in which participants are asked to report their time spent sitting or lying in different domains during the previous day. The validity for PAST and PAST-U with ActivPal as criterion measure was moderate to strong and test-retest reliability was fair to good. (Clark et al., 2013; Clark et al., 2016) Busschaert and co-workers tested validity and reliability in three different questionnaires measuring context-specific sedentary behaviour, which was summarized as total self-reported time. Test-retest reliability and the validity were acceptable. (2015)

Since subjective measurement methods are based on self-reported estimations it is imperative to know how well these measures correspond to the actual true value. Questionnaires measuring sedentary behaviour or sitting time often underestimate sitting time (Healy et al., 2011). Therefore, it is important to test the validity and reliability for SED-GIH. The validity has been tested previously, though with accelerometer as criterion measure (Kallings et al., 2014). Since ActivPal is more accurate for assess sitting time and can detect body postures (sitting, standing and reclining), it is of big interest to examine the validity for SED-GIH with ActivPal as the criterion measure. To the authors knowledge, the reliability for SED-GIH has not yet been examined.

3 Aim

The aim of this study was to investigate the criterion validity and test-retest reliability of the SED-GIH question, with ActivPal as the criterion measure.

Research questions:

- How strong is the level of agreement between sitting time measured subjectively with SED-GIH and objectively with ActivPal?
- How strong is the reliability of SED-GIH measured with test-retest?

4 Method

4.1 Study design

The data in this study was conducted from two larger projects (described below) with cross-sectional and cohort design respectively. The data was used to test the validity and reliability of SED-GIH.

4.2 Participants and sampling

The validity section of this study was a part of the large research project called “Physical
Activity and Healthy Brain Functions” at The Swedish school of Sport and Health science, GIH. The core question of that research project was “How can different components within physical activity patterns promote healthy brain functions?” The project had three sub projects were this study was a part of the first sub project, which was a cross-sectional study. Participants were recruited through convenience sampling at two companies in Stockholm and Gothenburg. The inclusion criteria were employees with an office-based work at these companies. As a start-up, information meetings were held at these companies. After these meetings, a total number of 1971 employees were invited to participate via mail. All measurements were conducted during working hours at an in-house test session.

The reliability section of this study was provided with data from a Health Project at GIH. The “Health Project” is a collaboration between GIH and the municipalities Solna and Lidingö. The municipalities informed about the project, which participants voluntary signed up for. The inclusion criterion was elderly individuals aged ≥ 65 years old. Students at GIH carried out the “Health Project”, as a part of their education.

4.2.1 Ethical considerations

The Stockholm regional ethics committee approved the sub project for “Physical Activity and Healthy Brain Functions”. Before the test session all participants signed a written informed consent form. The form included information considering the aim of the study, what benefits and risks participation could imply and that the participants were allowed to interrupt participation at any time without having to state why. All participation was anonymous, confidential and all contact was held by the researchers to ensure that the participants were truly participating on voluntary basis.

The Stockholm regional ethics committee approved the “Health Project” at GIH and all participants signed a written informed consent form. All participation was voluntary and all data was handled anonymous and confidential.

The risk of participating in both of these projects were low in relation to the potential benefits of finding a question aiming to assess sitting time, that might be used in health promotion.

4.3 Test procedures

In the sub project for the “Physical Activity and Healthy Brain Functions” study, participants responded to a web questionnaire before the in-house test session. The questionnaire included four areas; background, stress, work situation and health habits. The SED-GIH question was a part of this web questionnaire. In advance, participants also received information about the
test session and some preparations. The test session was completed by one of the three different test leaders. It took about one and a half hour and included cognition tests and a submaximal cardiovascular fitness test on a bicycle. Before leaving the test session, participants were equipped with two activity monitors, ActiGraph and ActivPal, and received instructions for a sleep and activity diary. The diary was filled out and the activity monitors were worn for a period of seven consecutive days. See the flow-chart of the test procedure in figure 1.

In the “Health Project” participants went through a test-session at GIH. First, they responded to a questionnaire with the topics perceived health, physical activity habits, diet, alcohol and tobacco use. SED-GIH was a part of this questionnaire. The test-session continued with anthropometric measures and different tests to assess level of physical fitness. About one week after the first test-session participants returned and completed the whole test-session one more time. Thereafter, they participated in an exercise program for a period of eight weeks. Finally, they returned to GIH and completed the whole test-session one last time.

Figure 1 – Flow-chart of the test procedure in “Physical Activity and Healthy Brain Functions” project
4.4 Outcome measures

From all variables collected by the “Physical Activity and Healthy Brain Functions” project and the “Health Project”, this study uses SED-GIH, ActivPal and the sleep and activity diaries.

4.4.1 SED-GIH

The question and its answering options reads:

How much time do you sit during a normal day, excluding sleep?
- Virtually all day
- 13-15 hours
- 10-12 hours
- 7-9 hours
- 4-6 hours
- 1-3 hours
- Never

4.4.2 ActivPal

The criterion measure for this study was ActivPal. To waterproof the ActivPal, it was placed in a small condom with transparent film around (Tegaderm Roll, 3M), which also was used to attach the ActivPal on the thigh. ActivPal continuously recorded the postures sitting/lying, standing and walking. The devices were initialized and downloaded using the activPAL software version 7.2.32.

4.4.3 Sleep and activity diaries

During the seven consecutive days participants wore the activity monitors. They were also logging sleep and activity in a diary. One side of the diary was filled in every night before bedtime. It contained questions regarding physical activity, working hours, when the activity monitor was not worn, estimations of sleepiness/alertness during the day and if something unusual had happened. The other side was filled in every morning directly after awakening. It contained questions regarding time for going to bed the previous day, how long time it took before falling asleep, time for awakening, sleep quality for the night, if anything special
happened that affected the sleep and if any sleeping drugs were taken. See the dairy in appendix 2.

4.5 Data analysis

The original data from ActivPal consisted of one large file for all seven consecutive days, one file for each participant. Excel HSC PAL analysis software V2 19s was used to merge the file with time parameters from the sleep and activity diaries to cut out the correct data. The time parameters used were awakening time, bedtime and working hours. For missing values in the diaries standardised times were used (awakening time 6 am, bedtime 11 pm, arrive to work 8 am and leave work 5 pm). If bedtime was later than 12 pm a separate analysis was conducted on the following date for the time passed 12 pm. These hours were added to the correct date afterwards. The new data was categorised into total wear time, sitting time, standing time and walking time. The data were presented in seconds that were converted into hours. The mean for all days, weekdays and weekend days were calculated. Inclusion criterion was at least four total consecutive days, with day one excluded since it was not a presentable 24-hour day. Inclusion criteria for weekend days were two consecutive days. The data that was presented in hours were converted into the same categorical answering options as used in SED-GIH. The categorical answering options and their codes (which will be used in the results) are presented in table 1. For the statistical analysis the ActivPal data was used both in hours (ratio data) and in the categorical answering options (ordinal data).

<table>
<thead>
<tr>
<th>Answering option</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtually all day</td>
<td>7</td>
</tr>
<tr>
<td>13-15 hours</td>
<td>6</td>
</tr>
<tr>
<td>10-12 hours</td>
<td>5</td>
</tr>
<tr>
<td>7-9 hours</td>
<td>4</td>
</tr>
<tr>
<td>4-6 hours</td>
<td>3</td>
</tr>
<tr>
<td>1-3 hours</td>
<td>2</td>
</tr>
<tr>
<td>Never</td>
<td>1</td>
</tr>
</tbody>
</table>

4.6 Statistical analysis

All analysis was conducted using IBM SPSS Statistics version 24. Statistical significance level were set at p < 0,05.
4.6.1 Criterion validity

4.6.1.1 Level of agreement between ActivPal and SED-GIH

Systematic errors between ActivPal and SED-GIH were assessed using Wilcoxon signed rank test, and random errors were assessed using Spearman’s rho. The associations was interpreted as weak (Spearman’s rho < 0.10), modest (Spearman’s rho 0.1-0.3), moderate (Spearman’s rho 0.3-0.5), strong (Spearman’s rho 0.5-0.8) or very strong (Spearman’s rho 0.8-1.0) (Muijs, 2011). To visually see the level of agreement a Bland Altman plot was conducted. The mean values of both methods were plotted against the difference. The limits of agreement were calculated based on values from a t-test of the difference between the both methods.

4.6.1.2 Difference between the categories of SED-GIH

A one-way independent ANOVA was conducted to assess if any differences existed between the categorical answering options from SED-GIH and the objective values of sitting time from ActivPal.

4.6.2 Test-retest reliability

Intraclass correlation (ICC) was calculated to assess the reliability of the test-retest of SED-GIH. The values from the ICC were interpreted as poor (ICC < 0.40), fair (ICC 0.40-0.59), good (ICC 0.60-0.74) or excellent (ICC 0.75-1.00) (Cicchetti, 1994). A Wilcoxon analysis was conducted to assess if any differences existed between the two occasions. To visually see the level of agreement a Bland Altman plot was conducted. The mean values of the both occasions were plotted against the difference. The limits of agreement were calculated based on values from a t-test of the difference between the two occasions.

5 Results

5.1 Participants characteristics

5.1.1 ”Physical Activity and Healthy Brain Functions” project

A total number of 284 participants were included in the validity part of this study. The number of participants from each company and number of participant who fulfilled both the test-session and the web questionnaire are presented in figure 2. The characteristics of the sample are presented in Table 2.
Figure 2 – Flow-chart of the number of participants in “Physical Activity and Healthy Brain Functions” project

Table 2 – Characteristics of participants who were included in the “Physical Activity and Healthy Brain Function” project

<table>
<thead>
<tr>
<th></th>
<th>Missing, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>n total</td>
<td>284</td>
</tr>
<tr>
<td>Gender, %</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>33.5</td>
</tr>
<tr>
<td>Female</td>
<td>66.2</td>
</tr>
<tr>
<td>Mean age, years (SD)</td>
<td>41.91 (8.9)</td>
</tr>
<tr>
<td>Education, %</td>
<td></td>
</tr>
<tr>
<td>Elementary school</td>
<td>2.1</td>
</tr>
<tr>
<td>High school</td>
<td>38.7</td>
</tr>
<tr>
<td>College</td>
<td>54.2</td>
</tr>
<tr>
<td>Higher academic education</td>
<td>4.2</td>
</tr>
</tbody>
</table>
5.1.2 ”Health Project” at GIH

A total number of 95 participants fulfilled both the fist and the second test-session, and were included in the reliability part of this study. The characteristics of the sample are presented in Table 3.

Table 3 – Characteristics of the sample from the “Health Project”

<table>
<thead>
<tr>
<th>n total</th>
<th>95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, %</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>30.5</td>
</tr>
<tr>
<td>Female</td>
<td>69.5</td>
</tr>
<tr>
<td>Mean age, years (SD)</td>
<td>70.38 (4.95)</td>
</tr>
<tr>
<td>Mean diff. days T1-T2</td>
<td>5.24</td>
</tr>
</tbody>
</table>

5.2 Criterion validity

5.2.1 Level of agreement between ActivPal and SED-GIH

The correlation for the mean value for all days between SED-GIH and ActivPal was moderate with Spearman’s rho = 0.31 (figure 3). The correlation for weekdays was also moderate with Spearman’s rho = 0.31 and the correlation for weekend days was modest with Spearman’s rho = 0.18. From now on only mean value for all days will be used. A Wilcoxon analysis showed significant differences between SED-GIH and ActivPal with p < 0.01. This indicates a systematic error, where participants slightly underestimated their sitting time. A Bland Altman analysis between the mean values of ActivPal and SED-GIH was plotted against the difference (figure 4). Upper limits of agreement were 2.01, lower limits of agreement were -2.57 and the mean of the difference were -0.28. The plot indicates that participants who overestimated their sitting time with SED-GIH, actually sits more than the average. Participants who underestimates their sitting time with SED-GIH, actually sits less than the average.
Figure 3 – Correlation between sitting time measured with SED-GIH (in categorical data) and ActivPal (in hours). Shading behind presents several participants.
Figure 4 – Bland Altman Plot for agreement between SED-GIH and ActivPal. Shading behind presents several participants.
5.2.2 Difference between the categories of SED-GIH

A one-way independent ANOVA resulted in a significant \( p < 0.01 \), which indicates that differences between the categorical answering options of SED-GIH exist regarding sitting time measured objectively with ActivPal. Tukey Post Hoc test resulted in significant difference between the categorical answers options “Virtually all day” and “13-15 hours”, “7-9 hours”, “4-6 hours” and “1-3 hours”. Significant differences also existed between “10-12 hours” and “4-6 hours” and “1-3 hours” (table 4).

The disposition of participants answers of SED-GIH and their mean hours of sitting per day measured with ActivPal, are presented in table 5.

<table>
<thead>
<tr>
<th>Answer options</th>
<th>n</th>
<th>% of total n</th>
<th>ActivPal, mean hours of sitting time per day</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td>Virtually all day</td>
<td>11</td>
<td>3.9</td>
<td>11.3</td>
<td>10.4</td>
</tr>
<tr>
<td>13-15 hours</td>
<td>24</td>
<td>8.5</td>
<td>9.8</td>
<td>9.2</td>
</tr>
<tr>
<td>10-12 hours</td>
<td>84</td>
<td>29.6</td>
<td>10.1</td>
<td>9.8</td>
</tr>
<tr>
<td>7-9 hours</td>
<td>102</td>
<td>35.9</td>
<td>9.6</td>
<td>9.3</td>
</tr>
<tr>
<td>4-6 hours</td>
<td>45</td>
<td>15.8</td>
<td>9.2</td>
<td>8.8</td>
</tr>
<tr>
<td>1-3 hours</td>
<td>16</td>
<td>5.6</td>
<td>8.5</td>
<td>7.9</td>
</tr>
<tr>
<td>Never</td>
<td>2</td>
<td>0.7</td>
<td>9.6</td>
<td>6.8</td>
</tr>
<tr>
<td>Total</td>
<td>284</td>
<td>100</td>
<td>9.7</td>
<td>9.6</td>
</tr>
</tbody>
</table>
5.3 Test-retest reliability

Intraclass correlation coefficient for the test-retest reliability of the SED-GIH was excellent with ICC = 0.86. A Wilcoxon analysis showed non-significant differences between the two occasions with p = 0.12. A Bland Altman analysis between the mean values of test-session one and test-session two was plotted against the difference. Upper limits of agreement were 0.93, lower limits of agreement were -1.1 and the mean of the difference were 0.08. The plot indicates that participants were consistent in the way they answered SED-GIH at test-session one and test-session two.

Figure 5 – Correlation between test-session one and test-session two. Shading behind presents several participants.
Figure 6 – Bland Altman Plot for test-retest of SED-GIH. Shading behind presents several participants.
6 Discussion

The aim of the current study was to investigate the criterion validity and test-retest reliability of the SED-GIH question, with ActivPal as the criterion measure. The main findings were a moderate correlation between SED-GIH and ActivPal, were SED-GIH slightly underestimated sitting time. The reliability of SED-GIH was excellent.

6.1 Level of agreement between sitting time measured subjectively with SED-GIH and objectively with ActivPal

The current study showed a moderate correlation between sitting time measured objectively with ActivPal and sitting time measured subjectively with SED-GIH. Even though the results only indicates a moderate correlation, the results are in line with other questions and questionnaires that assess sitting time or sedentary behaviour. When Chastin and colleagues validated sitting time by IPAQ with ActivPal as criterion measure, they found low and non-significant correlations between the two methods were IPAQ underestimated sitting time (ICC = 0.11-0.28, r = 0.01-0.16) (2014). The validity for the questionnaires PAST and PAST-U with ActivPal as criterion measure, was moderate to strong (ICC = 0.64, r = 0.57) (Clark et al., 2013; Clark et al., 2016). When Busschaert and co-workers tested validity in three different questionnaires measuring context-specific sedentary behaviour they found acceptable validity (r = 0.02-0.52) (2015).

The correlation for weekdays sitting time in the current study was, just as the correlation for the total sitting time, moderate and the correlation for weekend days was only modest. The results are in line with the findings from Chastin and colleagues when they validated IPAQ and found that participants reported less accurate sitting time on weekend day than weekdays (2014). These findings may be explained by more structured daily activities during weekdays than weekend day, which can make it easier to estimate sitting time and also apply the SED-GIH term “how much time we spend sitting during a normal day” on weekdays.

The results from the current study indicate systematic errors, since a Wilcoxon analysis showed significant differences between SED-GIH and ActivPal. These results were expected since it is hard to estimate exactly how much time you spend sitting on a normal day, which is common for questions and questionnaires assessing both sitting time, sedentary behaviour and physical activity (Atkin et al., 2012; Healy et al., 2011).

The Bland Altman plot in the current study indicates that participants both under and overestimated their sitting time, with a general slight underestimation. Limits of agreement
were wide (2.01 to -2.57 in categorical answering units) and the mean of the difference were -0.28. The results indicates the same as the Bland Altman plot between PAST and ActivPal, where the limits of agreement also were wide (-4.90 to 4.60 hours), a large distribution in the data existed with both under and overestimations but in general a slight underestimation of sitting time with PAST occurred (Clark et al., 2013). When Chastin and colleagues created a Bland Altman plot of the difference between IPAQ and ActivPal against their mean, they found that IPAQ underestimated sitting time (2014). However, both PAST-U and the three different questionnaires measuring context-specific sedentary behaviour overestimated sedentary time, with ActivPal as the criterion measure (Clark et al., 2016; Busschaert et al., 2015). Thus, both under and overestimations of sedentary behaviour or sitting time occurs in commonly used questions or questionnaires. This may be due to how the question is asked. IPAQ is asking for sitting time during the last seven days, which is similar to SED-GIH. PAST and PAST-U are asking for sitting time during the previous day, and the three context specific questions asks for sitting time in different context. To address these over and underestimations one solution could be to adjust for this in a model. If SED-GIH is used in a population study, an adjustment model could recalculate the results with the underestimation of 0.28 categorical units in mind.

6.2 How the objective sitting time measured with ActivPal varies depending on the subjective estimations measured with SED-GIH

The ANOVA with Tukey Post Hoc test showed that differences exist between the categorical answering options of SED-GIH (p < 0.01). However, the mean values of sitting time measured with ActivPal did not differ much between the categorical answering options of SED-GIH. The lowest mean value was 8.5 hours per day, the highest was 11.3 hours per day and the total mean value for all participants was 9.7 hours per day. Thus, in general all participants objectively measured average sitting time per day was close to each other, even though they have estimated their sitting time with SED-GIH in a wide range.

No significant differences existed between “13-15 hours” and any of the other groups (except “Virtually all day”), which were unexpected since the two groups nearby (“Virtually all day” and “10-12 hours”) showed significant differences. This can be described by large variations regarding the objective values of sitting time measured with ActivPal and the number of participants in these groups (see table 5).

The total number of participants for the categorical answering options “10-12 hours per day” and “7-9 hours per day” was 186, which is 65.5 % of the total. The mean values of sitting
time measured with ActivPal in these two groups were 10.1 and 9.6 hours per day. If we cluster these two groups into one (“7-12 hours”), all participants in this big group answered SED-GIH correctly based on their mean values from ActivPal (9.9 hours per day). This indicates that SED-GIH in advantage can be used on population levels. The number of participants in the categorical answering options “Virtually all day”, “13-15 hours”, “1-3 hours” and “Never” is low, which is a weakness to the study.

6.3 Reliability of SED-GIH measured with test-retest

Test-retest reliability of the SED-GIH was excellent (ICC = 0.86), which is better than other reliability tested questions. When Clark and colleagues tested the reliability for PAST with test-retest it resulted in fair to good values (ICC = 0.50) (2013). Busschaert and co-workers tested reliability in three different questionnaires measuring context-specific sedentary behaviour, were test-retest reliability was acceptable (ICC = 0.37-0.80) (2015).

The Bland Altman plot between the mean values of test-session one and test-session two, against their difference had narrow limits of agreement (0.93 and -1.1), and almost all values were placed in this span. The plot indicates that participants were consistent in the way they answered SED-GIH at test-session one and test-session two. The reliability of SED-GIH was strong and the Bland Altman plot indicates a consistency in the way participants choose their answers, which is a strong argument for its usage in healthcare.

6.4 Method discussion

The current study is part of two larger projects which entails several strengths like well and detailed planned methods and a large number of participants which results in great power to the study. All participation in the two projects was on voluntary basis, which may have a positive effect on the motivation to fulfilment. Limitations to the current study have been observed considering the methods and the processing of the data. During the test-session it is possible that the test leaders may have affected the participants habits regarding behaviour of sitting, by informing regarding how periods of prolonged sitting may cause health risks. It is also possible that the participants may have been conscious about their unhealthy habits regarding sitting time when they answered the web questionnaire some weeks before the test-session, which can have caused a mismatch between SED-GIH and ActivPal. The fact that it was three different test leaders who conducted the test-sessions can have affected the information the participants received. Another impact on the internal validity to the current study is the accuracy of the participants dedication to fill in the diaries correctly. Since the
diaries are required to obtain the ActivPal data, incorrect dairies will affect the whole dataset. Some careless mistakes and not correct fulfilled diaries were detected during the analysis, which cause a doubt to the accuracy. If the timespan of wearing ActivPal is incorrect it can result in too much or too little data that will be categorized as sitting time, which is unsatisfying in the current study where the aim is to assess sitting time. The analysis to obtain the ActivPal data is monotonously and it is possible to make careless mistakes when adding all the time variables from the diaries, which also can affect the internal validity of the data. In the reliability part of the current study all participants were elderly. This may have an effect on the results since elderly can have reduced memory function.

6.5 Implications

Today we spend the majority of our awake time in some kind of sitting or sedentary activities and many circumstances affect our sitting behaviour, such as norms, attitudes and motivation. Norms are implemented both by environmental and social contexts, like the strong norm to sit in classes and meetings. (Owen et al., 2011; Owen et al., 2010) The fact that sitting, independent of the level of physical activity, is increasingly being recognized as a health risk for all-cause mortality, cancer, cardiovascular diseases, type 2 diabetes and metabolic risks, makes this an area that needs high priority (Biswas et al., 2015; Healy et al., 2008b). To assess this problematic area, a tool to easily measure sitting time is needed. The best measurement methods are the objective ones like direct observation, accelerometers or ActivPal. Since these are expensive a cheaper and user-friendly tool is necessary, like questionnaires or even better, a single item question. Kallings and colleagues developed SED-GIH with categorical answering options, since previous studies indicated that as the best alternative (2014). The validity for SED-GIH had only been evaluated with accelerometer as criterion measure. Since the accelerometer cannot separate the postures sitting and standing, ActivPal is a much better criterion measure. Therefore, the current study assessed the criterion validity for SED-GIH with ActivPal as the criterion measure. The agreement between the two methods was moderate, which is in line with other questions that assess sitting time. The test-retest reliability of SED-GIH was excellent. The results can be compared with the commonly used IPAQ, which actually provided low and non-significant correlations when the agreement was assessed with ActivPal as the criterion measure, were IPAQ underestimated sitting time (Chastin et al., 2014). To use TV-viewing time as a domain for assessing sedentary behaviour is common, since it often occurs on regular basis in prolonged time periods. TV-viewing time has good reliability, though the validity in different studies report large differences. (Clark et
In this context SED-GIH is a better alternative to use since it does not measure any specific context or behaviours on a regular basis like TV-viewing time, but it still has excellent reliability.

Questionnaires and other subjective measurement methods in general have low costs, are easy to implement on a large scale and have low burden both for the participants and for the ones who summarize the results (Atkin et al., 2012; Healy et al., 2011). Since SED-GIH has strong arguments for acceptable validity and excellent reliability it can likely be useful in healthcare systems. SED-GIH can be used as a tool in the primary care when defining determinants for unhealthy habits like too much sitting, but also in national public health surveys. SED-GIH is since 2016 included in the Swedish national public health survey, Health on Equal Terms (Folkhälsomyndigheten, 2016). If sitting time will be measured more frequently with an easy measurement method as SED-GIH, it can help us to become more aware of how much we actually sit. If we also add the important information regarding the health risks with sitting and how important it is to break up periods of prolonged sitting, we can provide great health benefits on a population level. This in turn, can have positive effects on the sitting norm that exist in our society today and hopefully reverse the sitting trend.

The current study is the first to assess reliability for SED-GIH and second to validate SED-GIH, but the first one to use ActivPal as criterion measure. Further research should be completed on other groups, in different ages and with different kind of settings, not only in office settings.

6.6 Conclusion

SED-GIH both over and underestimated sitting time, however a general slightly underestimation occurred. The validity and reliability of SED-GIH are in line, or even better, than other questions that are common to use when assessing sitting time. This make SED-GIH useful in many areas, like in the healthcare system and in national public health surveys.
References


Folkhälsomyndigheten (2016). *Hälsa på lika villkor?* Stockholm


Appendix 1

Literature search

Aim and research questions:

The aim of this study was to investigate the criterion validity and test-retest reliability of the SED-GIH question, with ActivPal as the criterion measure.

Research questions:

• How strong is the level of agreement between sitting time measured subjectively with SED-GIH and objectively with ActivPal?
• How strong is the reliability of SED-GIH measured with test-retest?

Searches that gave relevant results:

<table>
<thead>
<tr>
<th>Database</th>
<th>Search string</th>
<th>Number of hits</th>
<th>Number of relevant hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>PubMed</td>
<td>&quot;Sedentary Lifestyle&quot;[Mesh] AND measurement English Adult: 19+ years</td>
<td>165</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>(&quot;Sedentary Lifestyle&quot;[Mesh]) AND &quot;Health Status Indicators&quot;[Mesh] English Adult: 19+ years</td>
<td>87</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sedentary AND &quot;health risk&quot; English Adult: 19+ years</td>
<td>104</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>&quot;Prevalence&quot;[Mesh] AND &quot;sitting time&quot; English Adult: 19+ years</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>&quot;Sedentary Lifestyle&quot;[Mesh] AND validity English Adult: 19+ years</td>
<td>94</td>
<td>32</td>
</tr>
</tbody>
</table>
Appendix 2

Sleep and activity diary

<table>
<thead>
<tr>
<th>Dag</th>
<th>Har du utfört någon fysisk aktivitet/träning under dagen (klockslag, typ)?</th>
<th>Har du tagit av aktivitetsmätaren under träningen?</th>
<th>Arbetstid (klockslag): När arbetade du idag (på kontor, på väg, hemma, etc.)?</th>
<th>Anteckna när aktivitetsmätaren har varit borttagen och anledningen (utöver träning)</th>
<th>Om du tänker på hur sömnig/pigg du varit under största delen av dagen, vilket påstående (1-9) passar in på dig?</th>
<th>Om du tagit en tupplur under dagen, ange tidperioden.</th>
<th>Har det inträffat något speciellt under dagen som var utöver det vanliga, t ex att du var sjuk?</th>
</tr>
</thead>
</table>
| 1   | "_dag_ (kväll)"
| 2   | "_dag_ (kväll)"
| 3   | "_dag_ (kväll)"
| 4   | "_dag_ (kväll)"
| 5   | "_dag_ (kväll)"
| 6   | "_dag_ (kväll)"
| 7   | "_dag_ (kväll)"

Exempel: tex. Måndag Cyklade till jobbet (7.30-7.45) Simträning 19-20) X 8.15-12.00, 13.00-16.00 21.00-23.00 6.45, 15min, dusch 20.10-20.20 dusch 2 16.20-16.40 nej

KOD:

STARTDATUM:

STARTTID:
Svara på dessa frågor **varje morgon** direkt när du har vaknat

<table>
<thead>
<tr>
<th>Dag:</th>
<th>Igår kväll släckte jag lampan kl. ...</th>
<th>Efter jag släckte lampan somnade jag på ... (min, ca)</th>
<th>I morse vaknade jag kl. ... (slutgiltigt)</th>
<th>Hur skulle du skatta din sömnkvalité?</th>
<th>Om det har hänt någonting speciellt utöver det vanliga under natten, beskriv det kort här</th>
<th>Har du tagit sömnmedel för att kunna sova? Anteckna preparatnamn och dos.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exempel</td>
<td>tex. Måndag</td>
<td>22.35</td>
<td>5</td>
<td>6.13</td>
<td>4</td>
<td>Mitt barn sjukt och vaknade varje timme.</td>
</tr>
<tr>
<td>Dagen du får aktivitetsmätarna</td>
<td>___dag (morgon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>___dag (morgon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>___dag (morgon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>___dag (morgon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>___dag (morgon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>___dag (morgon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>___dag (morgon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>___dag (morgon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tack för att du fylldes in sömndagboken. Du kan lämna in kuvertet med aktivitetsmätarna och sömndagboken i den vita brevlådan utanför massagerummet. Vi är tacksamma om du kan lämna in det snarast.**