Test-retest reliability and agreement of the Satisfaction with the Assistive Technology Services (SATS) instrument in two Nordic countries

Sund, Terje; Iwarsson, Susanne; Anttila, Heidi; Helle, Tina; Brandt, Ase

Published in:
Physiotherapy Theory and Practice

DOI:
10.3109/09593985.2013.876478

2014

Document Version:
Peer reviewed version (aka post-print)

Physiotherapy Theory and Practice

TEST-RETEST RELIABILITY AND AGREEMENT OF THE SATISFACTION WITH THE ASSISTIVE TECHNOLOGY SERVICES (SATS) INSTRUMENT IN TWO NORDIC COUNTRIES

Terje Sund, PT, MSc[1,2], Susanne Iwarsson, Reg OT, PhD, Professor[2], Heidi Anttila, PT, PhD[3], Tina Helle, Reg OT, PhD[2,4] & Åse Brandt, Reg OT, PhD, Senior Researcher[5]

1 Department of Assistive Technology, The Norwegian Labour and Welfare Service, Oslo, Norway
2 Department of Health Sciences, Faculty of Medicine, Lund University, Lund, Sweden
3 Health and Social Services Development Unit, Service System Department, National Institute for Health and Welfare, Helsinki, Finland
4 Department of Occupational Therapy, University College Northern Denmark, Aalborg, Denmark
5 The National Board of Social Services, Odense, Denmark

Conflicts of interest: none declared

Corresponding author: Terje Sund, NAV Hjelpemidler og tilrettelegging, PO box 5 St. Olavs plass, 0130 Oslo, Norway. Telephone: + 47 21 06 86 66. E-mail: terje.sund@nav.no

Abstract

Purpose. The purpose of this study was to investigate test-retest reliability, agreement, internal consistency, and floor- and ceiling effects of the Danish and Finnish versions of the Satisfaction with Assistive Technology Services (SATS) instrument among adult users of powered wheelchairs (PWC) or powered scooters (scooters).

Method. Test-retest design, two telephone interviews 7-18 days apart of 40 informants, with mean age of 67.5 (SD 13.09) years in the Danish, and 54, with mean age of 55.6 (SD 12.09) years in the Finnish sample.

Results. The Intraclass Correlation Coefficient (ICC) varied between 0.57-0.93 for items in the Danish and between 0.41-0.93 in the Finnish sample. The percentage agreement varied between 54.2 and 79.5 for items in the Danish and between 69.2 and 81.1 in the Finnish sample, while the Cronbach's alpha values varied between 0.87-0.96 in the two samples. A ceiling effect was found in all items of both samples.

Conclusion. This study indicates that the SATS may be reliably administered for telephone interviews among adult PWC and scooter users, and give information about aspects of the service delivery process for quality development improvement purposes. Further psychometric testing of the SATS is required.

Keywords**:** assistive technology, instrument development, intra-observer reliability, powered wheelchairs, psychometric testing, rehabilitation, scooters, service delivery

Introduction

Provision of assistive devices is an important part of rehabilitation services (Scherer, 2002; Donabedian, 2003; Sackett et al, 2007; Brandt, Kreiner, Iwarsson, 2010; Lenker et al, 2010). Users' assessment of assistive technology services is vital in a client-based rehabilitation context, since such assessments can inform, guide and support decision-makers to identify optimal ways to organise the services (Lenker et al, 2005; Sund, Iwarsson, Andersen, Brandt, 2013), and they may be a prerequisite for positive outcomes (Auger et al, 2010). User assessments are often conducted in terms of user satisfaction, frequently defined as the user's critical evaluation of different aspects of assistive devices and related services, which is influenced by individual expectations, perceptions, attitudes and personal values related to services given (Demers, Weiss-Lambrou, Ska, 2000). This article, however, only concerns assessment of user satisfaction with the service delivery process (SDP) of assistive devices, and specifically it is about psychometric testing of the' Satisfaction with the Assistive Technology Services (SATS)' instrument, which is constructed to assess users' satisfaction with the SDP.

The SDP directly assists in the selection, acquisition and use of an assistive device. First of all, it is important to decide whether an assistive device is likely to solve the user's practical problems in daily life, in which case it is necessary to select a product that fits the user's needs. The SDP is important in achieving this (Cook and Polgar, 2008). At present, evidence-based knowledge concerning user satisfaction with the SDP is limited (Berndt, van der Pijl, De Witte, 2009; Lenker et al, 2010).

The importance of the assistive devices is increasing due to rapid increase of chronic diseases and ageing population living with chronic diseases, as older people tend to have a higher

prevalence of disability restricting participation (World Health Organization WHO and The World Bank, 2011). Powered wheelchairs (PWC) and scooters are some of the most common and expensive assistive devices for enhancing mobility-related participation (Cook and Polgar, 2008). Available Nordic statistics stated that in 2012 in all 4.46 euros per inhabitant were spent on PWCs and 1.26 on scooters in Norway (The Norwegian Labour and Welfare Service NAV, 2013). It is believed that this is similar in the other Nordic countries due to their social welfare systems providing assistive devices for citizens with functional limitations.

Existing instruments

Knowledge about user satisfaction with the SDP should be generated from data collected with valid and reliable instruments (Streiner and Norman, 2008). A recent literature review based on searches in the CINAHL, Pubmed and Cochrane databases revealed only one instrument constructed to assess satisfaction with the assistive technology, i.e. the device and the SDP, namely the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST 2.0) (Demers, Weiss-Lambrou, Ska, 1996). The SDP part of the QUEST 2.0 consists of four items and has some experienced-based limitations, e.g. that the four QUEST 2.0 service items require a different timing of when the users should be asked, and that a 'does not know' response alternative is not included (Wressle and Samuelsson, 2004; Brandt, 2005; Samuelsson and Wressle, 2008). Besides, some authors have criticized the QUEST 2.0 for providing too little detailed information to really support quality development of the services (Sund, 2004; Brandt, 2005). Therefore, there is a need for an instrument which can assess user satisfaction with specific aspects of the SDP.

The Satisfaction with the Assistive Technology Services (SATS) Instrument

Consequently, four experienced Norwegian therapists/researchers (including the first author) constructed the Satisfaction with the Assistive Technology Services (SATS) instrument, aiming at assessing user satisfaction with the SDP of assistive devices in a cost effective way. The instrument was to be usable in praxis contexts and therefore
it was decided that it should satisfy grade A according to Andresen's (Andresen, 2000) criteria for assessing tools of disability outcomes research. That is, an instrument should be easily scored, easy to understand and be completed within 15 minutes. First, the purpose of the new instrument was stated. As a second step, and being as inclusive and open-minded as possible, an item pool consisting of items relevant for assessing user satisfaction with the SDP of assistive devices was created. The next step was to critically appraise and discuss the items, and to select those that met the needs identified. Many of the selected items coincided with the those of the KWAZO, a Dutch instrument constructed to assess quality aspects of the SDP in terms of goal attainment (Dijcks, Wessels, De Vlieger, Post, 2006), why the KWAZO items were chosen for the SATS. The KWAZO is based on criteria for quality of care found in literature and in the Horizontal European Activities of Rehabilitation Technology (HEART) study (Heart-Line, 1994). After factor and internal consistency analyses, the final version of the KWAZO had seven items, which were used: 'accessibility to the professionals', 'information', 'coordination between the professionals', 'knowledge of the professionals', 'waiting time', 'participation', 'instruction and training' (Dijcks, Wessels, De Vlieger, Post, 2006). We were not, however, interested in the goal attainment outcome dimension, but rather in user satisfaction, which is why the items were rephrased as questions about satisfaction. In order to optimise content validity of the SATS, five representatives of the three Norwegian disability user organisations (umbrella organisations) were consulted. It was decided to add an

item on 'follow-up services', as this was considered to be an important aspect of the SDP. In addition, an 'overall satisfaction' item was included. Thus the SATS instrument tested in this study consisted of nine items. A five-point ordinal rating scale ranging from '1=very dissatisfied' to '5=very satisfied and a 'does not know' response option were offered. In addition, an option to give comments on the items was included. The instrument was constructed in a structured interview format and to be applicable to the provision of all kinds of assistive devices. For details, see table 1.

(Insert table 1 in about here)

Once the Norwegian version was established, the SATS was translated into Danish and Finnish. The Danish translation was carried out by a simple one way translation process as written Norwegian and Danish are very similar. Two Danish health professionals independently translated the SATS, one of whom lived in Norway and the other (the last author) had a PhD degree in occupational therapy. In addition, a person with a master degree in Danish language contributed to the translation. Based on discussions, some adjustments were made before a final Danish version was established. For the Finnish translation, a two-panel method was used (Swaine-Verdier et al, 2004), and the Finnish version was tested for face and linguistic validity and for two different administration modes with 19 persons with disabilities who recently had got a rollator or a wheelchair. The two administration modes seemed reliable and the users considered the SATS items acceptable, easy to answer (clear understanding of the wording of each item) and relevant for assessing user satisfaction with the SDP of assistive devices (Ahtola, Heinonen, Haikonen, Anttila, 2011). No further psychometric information about the Danish and Finnish versions of the SATS is available.

Psychometric properties

Reliability and agreement, internal consistency and floor or ceiling effects are important issues in the development of an instrument (Kottner et al, 2011). The test-retest reliability of an instrument is the ability to differentiate between cases and is defined as the ratio of variability between subjects to the total variability of all measurements in a sample (Streiner and Norman, 2008). Agreement is the degree to which repeated ratings or scores are identical (de Vet, Terwee, Knol, Bouter et al, 2006; Kottner et al, 2011). Internal consistency is defined as the degree of the interrelatedness among the items of an instrument (Mokkink et al, 2010) and provides information on whether several items that propose to measure the same general construct produce similar scores (Streiner, 2003). A floor or ceiling effect is dependent on the clustering of responses at the lowest or highest scoring options (Andresen, 2000). High proportions of such clustering reduce an instruments ability to differentiate between cases (Streiner and Norman, 2008), and most likely reduce its sensitivity to changes.

Therefore, the aim of this study was to investigate the test-retest reliability, agreement, internal consistency as well as floor- and ceiling effects of the Danish and Finnish versions of the SATS among adult users of PWCs and scooters.

Materials and methods

Sample

This study was based on data collected in Denmark and Finland. The informants were adults who had been provided a PWC or a scooter under the terms of the Social Service Act in Denmark and the Health Care Act in Finland. The inclusion criteria were: a) persons at least 18 years of age who had been provided a PWC or scooter within the latest 1-3 months, b) sufficient cognitive function to be able to reflect on the items and to express personal opinions

on the phenomena under study during a telephone interview (based on assessments by the case managers on basis on their prior knowledge about the informants and discussions with them) c) living in ordinary housing.

Based on textbook recommendations (Streiner and Norman, 2008; De Vet, Terwee, Mokkink, Knol, 2011) a sample of about 50 informants was aimed for from each country. The Danish informants were sampled by the interviewers in two steps among users who had been provided a PWC or scooter from July to October 2010. In the first step municipalities were selected. Out of a total of the 98 Danish municipalities, 46 were randomly selected and their assistive technology departments were contacted by telephone. Out of these, 29 were interested in participating and were sent written information about the study. Finally, 14 municipalities representing a national geographical variation accepted to participate. As a second step and based on the inclusion criteria, the informants were recruited consecutively by case managers in the 14 municipalities. A total of 65 informants in Denmark were invited, of whom 24 declined to participate, mostly because they did not want to. During the study, one informant dropped out due to illness, leaving us with N=40 (response rate 61.5%).

In Finland, data were collected as part of a larger Nordic study. All twenty assistive technology centres serving municipalities in twenty hospitals districts were sent written information about the study and invited to participate. Eleven centres were interested, of which 10 participated and recruited informants consecutively by case managers. In all, 68 informants consented to participate. Two dropped out prior to the SATS telephone interviews because they did not want to continue. Another 10 cases were excluded because the time interval between the interviews exceeded a limit of 18 days because of e.g. holiday time. One questionnaire was lost in the mail, and one was discarded because of a different retest

interviewer. Thus, the final Finnish SATS test-retest sample consisted of 54 informants (response rate 77.1%).

Most informants were retired, with the Danish informants nearly 12 years older than the Finnish on average. Nearly half of the informants lived in a city in both countries, and more of the Finnish informants were living alone compared to those of the Danish sample. In both samples, most informants had been provided a scooter. In fact, in the Danish sample only two informants had received a PWC. There was a significant difference in self-reported diagnoses between the two samples. Demographic and health data are presented in table 2.

(Insert table 2 about here)

Ethical considerations

According to current national legislation, a formal ethical approval was not necessary in Denmark. In Finland the study received research ethics approval from the Ethical Council of the Hospital District of Helsinki and Uusimaa (Record no: 417/13/03/00/09).
All principles of ethical guidelines for human research were followed. The informants were guaranteed anonymity and confidentiality and were informed that participation was voluntary, and that they at any time could withdraw from the study without any consequences for future services.

Procedures

The informants were informed about the study and contacted by telephone in Denmark. Those who agreed to participate received written information as well as a letter of consent, which the informants signed and returned prior to the interviews.

Telephone interviews were chosen for data collection as this method has been found to be reliable for investigating psychometric properties of similar instruments (Auger, Demers et al, 2010). Telephone interviews are less time consuming than face-to-face interviews (Holbrook, Green, Krosnick, 2003) and give access to broad geographical areas at a lower cost.

In Denmark, four occupational therapy students performed the interviews, while in Finland the interviews were conducted by 17 specifically trained assistive technology case managers. The interviewers were briefed about the study prior to the data collection by the fourth author in Denmark and the third author in Finland. The interviewers gained experience in using the SATS instrument by interviewing each other or colleagues. The same interviewer interviewed the same informants on two different occasions (T1 and T2), aiming at a 14 days interval (maximum 18 days). This interval was chosen because it was considered to be adequate to decrease the possibility of a memory effect, and still short enough to minimize the risk of unforeseen changes (Streiner and Norman, 2008). The mean number of days between the two interview occasions was 13.5 (range 10-15, SD 1.09) days in the Danish and 11.8 (range 7-18, SD 3.32) in the Finnish sample.

In Denmark, a couple of days before the interviews the informants received the SATS instrument form by mail, enabling them to reflect on the questions and complete the form before the telephone call. In Finland, the informants received the SATS instrument form about one month before the first interview. During the interviews and over the telephone, the interviewer registered the informants' answers to each item on the SATS form. In both countries, the SATS interviews lasted for 10-15 minutes.

Data analysis

Data entry control was performed by the interviewers in Denmark, and by the third author in Finland. Descriptive statistics, independent two-sample t-test or chi-squared test were used to compare the two samples concerning demographic and health data, type of assistive device (PWC or scooter), and other assistive devices (yes/no). Since reliability is context- and population-specific (Streiner and Norman, 2008; Mokkink et al, 2012), the Danish and Finnish samples were analysed separately.

The percentage of 'does not know' responses to the items varied between 0.0% and 12.5% of all cases in both samples, except for the item 'follow-up services', which varied between 25.9% and 35.0%. Since 'does not know' was not part of the ordinal satisfaction scale, the 'does not know' responses were excluded from further analysis. For details, see table 3.

(Insert table 3 about here)

Reliability was assessed for each item by means of the intraclass correlation coefficient (ICC) (2,1) (Weir, 2005). Item ICC was reported because clinically each item provides information about aspects of the SDP. A coefficient of 0 represents a totally unreliable measurement and 1 indicates perfect reliability; levels > 0.7 are recommended (Streiner and Norman, 2008). In addition, percentage agreement was calculated. Internal consistency was investigated separately for T1 and T2 using Cronbach's alpha, with values ranging from 0.70 to 0.95 considered as acceptable (Bland and Altman, 1997; Nunnaly and Bernstein, 1997). For items with inter-item correlation <0.20, an item to total correlation analysis was conducted to examine the impact of removing these items. Values >0.70 may indicate item redundancy (de Vet, Terwee, Mokkink, Knol, 2011). To study floor and ceiling effects, the proportion of

scores at each end of the scale were calculated, with a floor or ceiling effect defined if a clustering of more than 20% of the responses at the lowest or highest score occurred (in this study rating score at 1 or 5) (Andresen, 2000).

The level of statistical significance was set to $p \leq 0.05$. SPSS 19.0 was used for all analyses (SPSS Inc, 2009).

Results

The ICC at item level varied between 0.57 ('participation' item) and 0.93 ('information' item) (mean ICC=0.79) in the Danish and 0.41('instruction' item) and 0.93 ('waiting time' item) (mean ICC=0.74) in the Finnish sample. For all items in the Danish sample, except for 'possibility to participate', the ICC values were >0.70. In the Finnish sample the ICC values of six items were >0.70. The percentage agreement varied between 54.2% and 79.5% (mean=70.0%) in the Danish and between 69.2% and 81.1% (mean=75.4%) in the Finnish sample. For details, see table 4.

(Insert table 4 about here)

In the Danish sample, the Cronbach's alpha values were 0.96 at both T1 and T2 (n=25). None of the Danish inter-item correlations were <0.20. Regarding possible item redundancy and at T1, the inter-item correlation for 'overall satisfaction' varied between 0.71-0.94. The item total correlation was 0.95. Removing the item reduced the Cronbach's alpha from 0.96 to 0.94. At T2, the inter-item correlation for 'overall satisfaction' varied between 0.50-0.91. The item total correlation was 0.90. Removing 'overall satisfaction' from the SATS reduced the Cronbach's alpha from 0.96 to 0.95. Inter-item correlation for the other items were <0.70.

In the Finnish sample, the Cronbach's alpha values were 0.88 and 0.87 at T1 (n=33) and T2 (n=36), respectively. At T1, the inter-item correlation values between 'instruction and training' and five other items ('accessibility to the professionals', 'cooperation between the professionals', 'waiting time', 'participation', 'overall satisfaction') varied between -0.05-0.09. The item total correlation value was 0.16. Removing 'instruction and training' from the Finnish version of the instrument increased the Cronbach's alpha from 0.88 to 0.89. At T2, the inter-item correlation values between 'instruction and training' and the same items as above, except 'overall satisfaction', were between 0.05-0.19. The item total correlation value was 0.31, and deleting 'instruction and training' from the instrument increased the Cronbach's alpha from 0.87 to 0.88. In the Finnish sample and regarding possible item redundancy, the inter-item correlation for 'overall satisfaction' was 0.74 for 'information' at T1 and 0.81 at T2. The corresponding values for the other items were <0.70.

'Very satisfied' and 'satisfied' were the predominant scores in both national samples. A ceiling effect was identified for all items for both samples, since much more than 20% of the responses were clustered at the positive end of the scale (5='very satisfied'). At T1 more than 50% of the informants had given a 'very satisfied' response to four SATS items in the Danish and eight SATS items in the Finnish sample, while the corresponding figures at T2 were one item in the Danish and eight items in the Finnish sample. Very few informants had given a 'dissatisfied' or 'very dissatisfied' response. For details, see table 5.

(Insert table 5 about here)

Discussion

This study tested psychometric properties of two national versions of the SATS, in two Nordic countries. The main results were that the ICC values of all but one of the items in the Danish and all but three of the items in Finnish sample were above the recommended level of 0.7, which indicates that the SATS may be reliably administered for telephone interviews among adult PWC and scooter users (see table 4). However, it should be kept in mind that since only two informants in the Danish sample had received a PWC (see table 2), it cannot be concluded that the Danish version of the SATS is applicable for studies with this user group.

The fact that the Cronbach's alpha coefficients of the SATS were high (0.87-0.96) in both samples indicates high internal consistency. That is, the SATS items seem to capture the same phenomenon, namely user satisfaction with the SDP. However, Cronbach's alpha values above 0.95, and as some authors claim that alpha values above 0.90 most likely indicate unnecessary redundancy rather than a desirable level of internal consistency (Streiner, 2003), there is a need to reconsider whether some of the items of the SATS could be excluded. Also, the results in the Finnish sample that the inter-item correlation between 'instruction and training' and five and four other items at T1 and T2, respectively, were <0.20 and fairly low item total correlation, indicate that 'instruction and training' is not able to discriminate users on satisfaction with the SDP very well (de Vet, Terwee, Mokkink, Knol, 2011). In addition, a previous Finnish SATS validity study (n=19) suggested to exclude the 'follow-up' item from the instrument. The Finnish validity study also suggested excluding the 'overall satisfaction' item with the argument that sufficient information could be collected by means of the other items of the SATS (Ahtola, Heinonen, Haikonen, Anttila, 2011).

Further, the fact that the inter-item correlation for 'overall satisfaction' in the Danish sample were >0.70 may indicate item redundancy. For this reason, removing 'overall satisfaction' from the Danish version of the SATS should be considered, but before doing so, more psychometric testing is required. This applies to both national versions of the SATS. On the other hand, the reason for constructing the SATS was to provide an instrument with the ability to assess different aspects of the SDP for quality development purposes. In that perspective it is important to have an instrument with a fair amount of items concerning essential aspects of the SDP. Consequently, we recommend a review of important aspects of the SDP.

The fact that there were relatively high numbers of 'does not know' responses for some of the items supported the inclusion of this response alternative in the SATS (see table 3). Relatively high proportions of 'does not know' responses to the 'follow-up services' item in both samples may indicate a need for different timing for asking this item – the same problem as for the QUEST 2.0 (Demers, Weiss-Lambrou, Ska, 2000; Demers et al, 2001; Wressle and Samuelsson, 2004; Brandt, 2005; Samuelsson and Wressle, 2008). That is, the other items can be administered shortly after the delivery of the assistive devices, when the users still can remember the process, while concerning follow-up the devices need to be used for longer, that is until the users have experienced using the devices in everyday activities.

The fact that 20.4% (T1) and 11.1% (T2) of the informants in the Finnish sample responded 'does not know' to the 'coordination between the professionals' item might indicate that the coordination had taken place without the informants' knowledge. A relatively high non-response rate appeared for the equivalent item in a Dutch study using the KWAZO instrument, probably indicating that the informants did not feel competent to rate this aspect of the services (Dijcks, Wessels, De Vlieger, Post, 2006). High proportions of 'does not

know' responses represent a validity problem (Streiner and Norman, 2008; Mokkink et al, 2012) because the item might not be relevant for the phenomenon under investigation. Consequently, the 'does not know' responses should be carefully monitored in future studies.

Since considerably more than 20% of the informants gave a 'very satisfied' response to all items in both national samples (see table 4), a marked ceiling effect was identified (Streiner and Norman, 2008). Similar ceiling effects were found in a Danish/Norwegian study among adult scooter users (n=86) in which the SATS instrument was used, where 31-62% of the informants gave a 'very satisfied' response to a number of items (Sund, Iwarsson, Andersen, Brandt, 2013). Other assistive technology instruments such as the Wheelchair Outcome Measure (WhOM) (an individualized, goal oriented measure of outcome related to wheelchair intervention) (Auger et al, 2010), the Pscychosocial Impact of Assistive Devices Scale (PIADS) (designed to assess the effects of an assistive device on functional independence, well-being and quality of life) (Demers et al, 2002) and the QUEST 2.0 (Demers, Weiss-Lambrou, Ska, 2002) have the same problem. So have surveys within health services (Rao, Peters, Bandeen-Roche, 2006; Brazil et al, 2013). A way to solve the problem of ceiling effects could be to construct an unbalanced scale with more 'satisfied' than 'unsatisfied' response alternatives. This might result in median scores closer to the middle of the scale and effectively increase the variability among the informants and the instrument's sensitivity to change, and so improving the ICCs without necessarily decreasing percentage agreement (Streiner and Norman, 2008). The fact that a large proportion of our informants gave a 'very satisfied' response may question what they actually responded to. Our observations regarding ceiling effects emphasize the importance of integrating open-ended questions that allow the informants to give qualitative comments to the various items, and to interpret these comments very carefully (Brazil et al, 2013). Moreover, as most of the SATS items were replicated from

the KWAZO and were not tested further for validity, a qualitative study to investigate how informants interpret the meaning of the different SATS items could be valuable, for example by the application of cognitive interviewing techniques (Willis, Reeve, Barofsky, 2005). This could supplement the Finnish SATS face and linguistic validity study of the SATS (Ahtola, Heinonen, Haikonen, Anttila, 2011).

Differences in the number of interviewers in the two countries combined with different levels of experience from using standardised rating instruments could have influenced the variance error, as more interviewers could increase the risk of performing the interviews differently. In order to explore this, as recommended by Slaug et al. (2012) an analysis of the characteristics of the interviewers and the settings in which the interviews took place would have been be useful. Unfortunately, we did not collect any data for such an analysis. On the other hand, in telephone interviews using structured questions the risk of interviewer bias is small (Steeh, 2008). Furthermore, the interviewers in this study were all trained to perform the telephone interviews, and there was no information indicating that the interviews were performed in a sub-optimal way or differently in the two countries. A SATS manual might have given the interviewers more guidance and perhaps strengthened the reliability even further (Gerdes et al., 2012). Even if it for scientific reasons is preferred to have a small number of trained interviewers in order to maximize reliability, the fact that the present study had many interviewers reflects the complex clinical practice reality. That is, since the SATS will be used by numerous interviewers and not only in ideal research circumstances by trained researchers, the number of interviewers may be considered a strength (Iwarsson et al, 1996).

In this study ICC was used as a measure of reliability in spite of the fact that several authors recommend weighted kappa for analysis of ordinal data (de Vet, Terwee, Knol, Bouter, 2006;

Mokkink et al, 2010; Kottner et al, 2011). However, since the results of ICC and weighted kappa analyses are identical, Streiner and Norman (Streiner and Norman, 2008) recommend using the ICC for all analysis except for nominal datasets.

As each telephone interview lasted for 10-15 minutes in both countries, the SATS instrument proved to be time-efficient. Also, as the Finnish face and linguistic validity study of the SATS concluded that the instrument had reasonably high acceptability to the informants (Ahtola, Heinonen, Haikonen, Anttila, 2011), the SATS therefore seems to achieve grade A according to Andresen et al's (Andresen, 2000) criteria for assessing the tools of disability outcomes research.

There were some limitations to this study. It should be noted that upon inclusion in the study, for practical reasons no standardized assessment was used to evaluate cognitive function. However, the case managers had good knowledge about the potential informants, and with their longstanding professional experience, we believe that they were well equipped to make a valid selection. Regarding possible changes in the informants' situations between the test and retest, opposed to in Denmark no such information was collected in Finland. Changes could possibly have affected the informants' responses (Streiner and Norman, 2008), and these informants should then have been excluded from the study. In Denmark, to assure that the informants did not have access to the results from the first interview when the interviews were repeated, their SATS forms were returned by mail immediately after the first interview. In Finland this procedure did not apply, but a closer inspection of the data collected revealed that no informants gave exactly the same responses in the two interviews. Consequently, we have no reason to believe that the slightly different modes of administration affected the results.

Conclusion

The reason for constructing the SATS was to provide an instrument with the ability to assess different aspects of the SDP for quality development purposes in research and practice. The present study indicates that the SATS may be reliably administered by means of telephone interviews among adult PWC and scooter users in two national contexts. The ICC values, percentage agreement and internal consistency coefficients indicate that the Danish and Finnish versions of the SATS can be used as intended. It should, however, be kept in mind that as only two informants in the Danish sample had received a PWC, it cannot be concluded that the Danish version of the SATS is applicable for studies within this user group. Further psychometric testing of the SATS is certainly required, as are studies in other national contexts.

Acknowledgements

The authors wish to thank all informants and data collectors.

Declaration of interest

The authors report no declarations of interest.

The study was financed by the Norwegian Labour and Welfare Service, Oslo, Norway; the Health and Social Services Development Unit, Service system department, National Institute for Health and Welfare, Helsinki, Finland; Academy of Finland; the National Resource Centre on Disability, Assistive Technology and Psychiatry, The National Board of Social Services, Odense, Denmark; the Department of Occupational Therapy, University College Northern Jutland, Aalborg, Denmark; and the Ribbing Foundation, Lund, Sweden. The study was conducted in the context of the Centre for Ageing and Supportive Environment (CASE) at Lund University, funded by the Swedish Council for Working Life and Social Research.

Table 1: The User Satisfaction with Assistive Technology Services (SATS) instrument*

		Five-point rating scale**	Does not know
Items	**How satisfied are you with:**		
1	the accessibility to the professionals?	1 2 3 4 5	
	Comments:		
2	the information given about application and assistive device solution?	1 2 3 4 5	
	Comments:		
3	the coordination between the professionals?	1 2 3 4 5	
	Comments:		
4	the knowledge of the professionals?	1 2 3 4 5	
	Comments:		
5	the waiting time from application to delivery of the assistive device?	1 2 3 4 5	
	Comments:		
6	the possibility to participate in decision about the assistive device solution?	1 2 3 4 5	
	Comments:		
7	the instruction and training given in the use of the assistive device?	1 2 3 4 5	
	Comments:		
8	the follow-up services after the delivery of the assistive device?	1 2 3 4 5	
	Comments:		
9	All in all, how satisfied are you with the service delivery in this case?	1 2 3 4 5	
	Comments:		

* This version is for presentation purposes only; the SATS has not been translated into English according to scientific recommendations.

**Five-point rating scale: 1=very dissatisfied; 2=dissatisfied; 3=neither dissatisfied nor satisfied; 4=satisfied; 5=very satisfied

Table 2: Characteristics of the national samples of powered wheelchair and scooter users (N=94)

	Danish sample, n=40	Finnish sample, n=54	Difference between samples, p value	Total sample, n=94
Mean age in years (SD)	67.5 (13.09)	55.6 (12.09)	p<0.001	60.7 (13.78)
Age in years, range	30-93	33-77		30-93
Men, n (%)	20 (50)	20 (37)	p=0.341	40 (43)
Place of living, n (%)			p=0.650	
city	21 (53)	24 (45)		45 (48)
suburb	8 (20)	15 (28)		23 (25)
rural area	11 (27)	14 (26)		25 (27)
Living alone, n (%)	20 (50)	51 (94)	p<0.001	71 (76)
Powered wheelchair, n (%)	2 (5)	13 (24)	p<0.001	15 (16)
Other mobility assistive devices, n (%)	35 (88)	53 (98)	p=0.037	88 (94)
Occupation, n (%)			p=0.973	
Employed	2 (5)	3 (6)		5 (5)
Retired/Pensioner	35 (88)	43 (84)		78 (86)
Redundant	2 (5)	3 (6)		5 (5)
Other	1 (2)	2 (4)		3 (4)
Self-reported impairments, n (%)*				
Reduced vision	27 (68)	12 (22)	p<0.001	39 (42)
Reduced hearing/deafness	13 (33)	4 (7)	p=0.002	17 (18)
Reduced balance and/or vertigo	16 (40)	37 (69)	p=0.034	53 (56)
Reduced endurance	29 (73)	30 (56)	p=0.237	59 (63)
Reduced function in arms	19 (48)	34 63)	p=0.326	53 (56)
Reduced function in back and/or legs	34 (85)	51 (94)	p=0.265	85 (90)
Problems with coordination of movements	11 (28)	25 (46)	p=0.141	36 (38)
Problems with movements of head/neck	6 (15)	6 (11)	p=0.746	12 (13)
Memory problems	11 (28)	10 (19)	p=0.292	21 (22)
Tiredness	28 (70)	29 (54)	p=0.238	57 (61)
Number of self-reported impairments, mean (SD)	5.1 (2.2)	4.8 (2.1)	p=0.422	4.9 (2.2)
Self-reported main diagnoses, n (%)**			p=0.012	
Diseases of the nervous system	15 (21)	35 (29)		50 (26)
Diseases of the circulatory system	13 (19)	20 (16)		33 (17)
Diseases of the musculosceletal system and connective tissue	21 (30)	14 (12)		35 (18)

Diseases of the respiratory system	11 (16)	10 (8)	21 (11)
Psychiatric diseases	1 (1)	10 (8)	11 (6)
Injury, poisoning etc.	2 (3)	3 (2)	5 (2)
Other diseases	7 (10)	31 (25)	38 (20)

*Data collected by means of the Nordic Mobility Related Outcome Evaluation of Assistive Device Intervention (NOMO) instrument.

**Data collected by means of study-specific ICD-10 based questionnaire.

Table 3: Number and percentage* of "does not know" responses in the Danish and Finnish samples (N=94)

	Danish (n=40)		Finnish (n=54)	
	Test	Retest	Test	Retest
Item**	Number (%) of "does not know" responses	Number (%) of "does not know" responses	Number (%) of "does not know" responses	Number (%) of "does not know" responses
Accessibility to	2 (5.0)	0 (0.0)	1 (1.9)	2 (3.7)
Information	0 (0.0)	1 (2.5)	2 (3.7)	1 (1.9)
Coordination	3 (7.5)	2 (5.0)	11 (20.4)	6 (11.1)
Knowledge	2 (5.0)	2 (5.0)	1 (1.9)	1 (1.9)
Waiting time	4 (10.0)	0 (0.0)	0 (0.0)	0 (0.0)
Participation	0 (0.0)	0 (0.0)	2 (3.7)	0 (0.0)
Instruction	5 (12.5)	5 (12.5)	1 (1.9)	1 (1.9)
Follow-up services	11 (27.5)	14 (35.0)	14 (25.9)	14 (25.9)
All in all satisfaction	2 (5.0)	3 (7.5)	0 (0.0)	0 (0.0)

*Number of responses in percentage of all possible responses

** See Table 1 for full item presentation

Table 4: Intraclass correlation coefficient (ICC) and percentage agreement of data collected by means of the SATS instrument* in the Danish and Finnish samples.

	Danish sample (n=40)			Finnish sample (n=54)		
Item	n**	ICC (CI)	Percentage agreement	n**	ICC (CI)	Percentage agreement
Accessibility to	38	0.73 (0.54 - 0.85)	71.1	52	0.65 (0.47-0.78)	71.2
Information	39	0.93 (0.86 - 0.96)	79.5	52	0.77 (0.62-0.86)	69.2
Coordination	36	0.85 (0.74 - 0.92)	72.2	42	0.85 (0.74-0.92)	81.0
Knowledge	36	0.85 (0.73 - 0.92)	69.4	53	0.62 (0.42-0.76)	69.8
Waiting time	36	0.86 (0.74 - 0.92)	72.2	54	0.93 (0.88-0.96)	79.6
Participation	40	0.57 (0.32 - 0.74)	70.0	52	0.88 (0.81-0.93)	76.9
Instruction	35	0.85 (0.72 - 0.92)	77.1	53	0.41 (0.16-0.61)	77.4
Follow-up services	24	0.71 (0.45 - 0.86)	54.2	36	0.81 (0.66-0.90)	72.2
All in all satisfaction	36	0.72 (0.52 - 0.85)	63.9	53	0.74 (0.59-0.84)	81.1

* Items rated on a five point scale: 1) Very dissatisfied, 2) Dissatisfied, 3) Neither dissatisfied nor satisfied, 4) Satisfied, 5). Very satisfied. (See Table 1)

** Because of internal drop-out n varies.

Table 5: Floor- and ceiling effects of the test-retest SATS data of Danish and Finnish samples among users of powered wheelchairs and scooters

	Danish sample (n=40)													
	Test							Retest						
Item	n**	Median (IQR)	1* (%)	2* (%)	3* (%)	4* (%)	5* (%)	n**	Median (IQR)	1* (%)	2* (%)	3* (%)	4* (%)	5* (%)
Accessibility to	38	5.0 (4.0-5.0)	0.0	5.3	7.9	31.6	55.3	40	4.0 (4.0-5.0)	0.0	5.0	7.5	42.5	45.0
Information	40	4.0 (4.0-5.0)	7.5	5.0	7.5	32.5	47.5	39	4.0 (4.0-5.0)	5.1	7.7	5.1	38.5	43.6
Cooperation	37	4.0 (3.5-5.0)	8.1	5.4	10.8	40.5	35.1	38	4.0 (4.0-5.0)	5.3	7.9	2.6	47.4	36.8
Knowledge	38	4.5 (4.0-5.0)	5.3	7.9	7.9	28.9	50.0	38	4.0 (4.0-5.0)	5.3	2.6	13.2	36.8	42.1
Waiting time	36	4.5 (3.0-5.0)	5,6	8.3	19.4	16.7	50.0	40	4.0 (2.0-5.0)	10.0	17.5	7.5	20.0	45.0
Participation	40	4.0 (4.0-5.0)	5.0	5.0	10.0	40.0	40.0	40	4.0 (4.0-5.0)	0.0	5.0	17.5	42.5	35.0
Instruction	35	5.0 (4.0-5.0)	5.7	2.9	5.7	25.7	60.0	35	5.0 (4.0-5.0)	5.7	8.6	2.9	31.4	51.4
Follow-up	30	4.0 (2.8-5.0)	6.7	16.7	10.0	33.3	33.3	26	4.0 (3.0-5.0)	11.5	3.8	26.9	30.8	26.9
All in all satisfaction	38	4.0 (4.0-5.0)	5.3	13.2	2.6	31.6	47.4	37	4.0 (4.0-5.0)	2.7	2.7	8.1	40.5	45.9
	Finnish sample (n=54)													
	Test							Retest						
Accessibility to	53	5.0 (4.0-5.0)	1.9	1.9	3.8	35.8	56.6	52	5.0 (4.0-5.0)	3.8	1.9	7.7	26.9	59.6
Information	52	4.5 (4.0-5.0)	0.0	5.8	7.7	36.5	50.0	53	4.0 (4.0-5.0)	0.0	3.8	7.5	41.5	47.2
Cooperation	43	4.0 (4.0-5.0)	4.7	7.0	2.3	37.2	48.8	48	5.0 (4.0-5.0)	4.2	4.2	10.4	25.0	56.3
Knowledge	53	5.0 (4.0-5.0)	0.0	1.9	3.8	26.4	67.9	53	5.0 (4.0-5.0)	0.0	1.9	0.0	30.2	67.9
Waiting time	54	5.0 (4.0-5.0)	7.4	5.6	3.7	24.1	59.3	54	5.0 (4.0-5.0)	5.6	9.3	1.9	25.9	57.4
Participation	52	4.5 (4.0-5.0)	0.0	9.6	5.8	34.6	50.0	54	5.0 (4.0-5.0)	1.9	7.4	9.3	24.1	57.4
Instruction	53	5.0 (4.0-5.0)	0.0	0.0	3.8	30.2	66.0	53	5.0 (4.0-5.0)	0.0	1.9	3.8	22.6	71.7
Follow-up	40	5.0 (4.0-5.0)	0.0	5.0	7.5	30.0	57.5	40	5.0 (4.0-5.0)	0.0	5.0	10.0	17.5	67.5
All in all satisfaction	54	5.0 (4.0-5.0)	0.0	7.4	9.3	24.1	59.3	53	5.0 (4.0-5.0)	0.0	1.9	5.7	34.0	58.5

* SATS response alternatives: 1=very dissatisfied; 2=dissatisfied; 3=Neither dissatisfied nor satisfied; 4=Satisfied; 5=Very satisfied.

** Because of internal drop-out, n varies.

References

Ahtola S, Heinonen A, Haikonen K, Anttila H 2011 Adaptation and validation of the modified KWAZO and EATS-2D instruments into Finnish circumstances. Everyday technology for independence and care AAATE 2011, Assistive Technology series. IOS Press BV, Amsterdam, Netherlands 29: 300-307.

Andresen E 2000 Criteria for assessing the tools of disability outcomes research. Archives of Physical Medicine and Rehabilitation 81: S15-S20.

Auger C, Demers L, Gelinas I, Routhier F, Mortenson WB, Miller WC 2010 Reliability and validity of telephone administration of the wheelchair outcome measure for middle-aged and older users of power mobility devices. Journal of Rehabilitaion Medicine 42: 574-581.

Berndt T, van der Pijl D, De Witte LP 2009 Existing models and instruments for the selection of assistive technology in rehabilitation practice. Scandinavian Journal of Occupational Therapy 16: 146-158.

Bland J, Altman D 1997 Statistics notes: Cronbach's alpha. British Medical Journal 314: 275.

Brandt Å 2005 Outcomes of rollator and powered wheelchair interventions. User satisfaction and participation (Dissertation) Faculty of Medicine, Division of Occupational Therapy, Lunds University. Lund, Sweden.

Brandt Å 2005 Translation, cross-cultural adaptation, and content validation of the QUEST. Technology and Disability 17: 205-216.

Brandt Å, Kreiner S, Iwarsson S 2010 Mobility-related participation and user satisfaction: Construct validity in the context of powered wheelchair use." Disability and Rehabilitation: Assistive Technology 5: 305-313.

Brazil K, Cupido C, Taniguchi A, Howard M, Akhtar-Denesh N, Frolic A 2013 Assessing familiy members' satisfaction with information sharing and communication during hospital care at the end of lIfe. Journal of Palliative Medicine 16: 82-86.

Cook A, Polgar J (Eds.) 2008 Cook & Hussey. Assistive Technologies. Principles and Practice. 3rd. edition, Mosby Elsevier.

De Vet H, Terwee C, Knol DL, Bouter LM 2006 When to use agreement versus reliability measures. Journal of Clinical Epidemiology 59: 1033-1039.

De Vet H, Terwee C, Mokkink LB, Knol DL 2011 Measurement in medicine. Practical Guides To Biostatistics and Epidemiology, Cambridge.

Demers L, Monette M, Descent M, Jutai J, Wolfson C 2002 The psychosocial impact of assistive devices scale (PIADS): Translation and preliminary psychometric evaluation of a Canadian-French version. Quality of life Research 11: 583-592.

Demers L, Weiss-Lambrou R, Ska B 1996 Development of the Quebec user evaluation of satisfaction with assistive technology (QUEST). Assistive Technology 8: 3-13.

Demers L, Weiss-Lambrou R, Ska B 2000 Quebec user evaluation of satisfaction with assistive technology QUEST version 2.0. An outcome measure for assistive technology devices. New York, The Institute for Matching Person & Technology.

Demers L, Weiss-Lambrou R, Ska B 2002 The Quebec user evaluation of satisfaction with assistive technology (QUEST 2.0): An overview and recent progress. Technology and Disability 14: 101-105.

Demers L, Wessels R, Weiss-Lambrou R, Ska B, De Witte LP 2001 Key dimensions of client satisfaction with assistive technology: A cross-validation of Canadian measure in the Netherlands. Journal of Rehabilitation Medicine 33: 187-191.

Dijcks B, Wessels R, De Vlieger SLM, Post MWM 2006 KWAZO, a new instrument to assess the quality of service delivery in assistive technology provision. Disability and Rehabilitation 28: 909-914.

Donabedian A 2003 An introduction to quality assurance in health care. New York, Oxford University Press.

Gerdes N, Funke UN, Schuwer U, Themann P, Pfeiffer G, Meffert C 2012 Scores of independence for neurologic and geriatric rehabilitation (SINGER) - Development and validation of a new assessment instrument." Rehabilitation (Stuttgart) Epub ahead of print.

Heart-Line C 1994 European Service Delivery Systems in Rehabilitation Technology. iRv, Hoensbroek.

Holbrook A, Green M, Krosnick J 2003 Telephone versus face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. Public Opinion Quarterly 67: 79-125.

Iwarsson S, Isacsson Å 1996. Development of a novel instrument for occupational therapry assessment of the physical environment in the home - A methodologis study on "The Enabler". Occupational Therapy Journal of Research, 16 (4):227-244.

Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, Roberts C, Shoukri M, Streiner DL 2011 Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. Journal of Clinical Epidemiology 64: 96-106.

Lenker J, Fuhrer M, Jutai J, Demers L, Scherer M, De Ruyter F 2010 Treatment theory, intervention specification, and treatment fidelity in assistive technology outcomes research. Assistive Technology 22: 129-138.

Lenker J, Scherer M, Fuhrer MJ, Jutai JW, De Ruyter F 2005 Psychometric and administrative properties of measures used in assistive technology device outcomes research. Assistive Technology 17: 7-22.

Mokkink L, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, De Vet HCW 2012 COSMIN checklist manual. Available at: www.cosmin.nl, www.emgo.nl. Accessed June 28th, 2012.

Mokkink L, Tewee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, De Vet HCW 2010 The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of Clinical Epidemiology 63: 737-45.

Nunnaly J, Bernstein L 1997 Psychometric theory. New York, McGraw-Hill Higher, INC.

Rao K, Peters DH, Bandeen-Roche K 2006 Towards patient-centered healch services in India - a scale to measure patient perceptions of quality. International Journal of Quality in Health Care 18: 414-421.

Sackett D, Rosenberg WM, Gray JA, Haynes RB, Richardson WS 2007 Evidence based medicine: What is and what it isn't. 1996. Clinical Orthopedics and Related Research 455: 3-5.

Samuelsson K, Wressle E 2008 User satisfaction with mobility assistive devices: An important element in the rehabilitation process. Disability and Rehabilitation 30: 551-558.

Scherer M 2002 The change in emphasis from people to person: introduction to the special issue on assistive technology. Disability and Rehabilitation 24: 1-4.

Slaug B, Schilling O, Helle T, Iwarsson S, Carlsson G, Brandt Å 2012 Unfolding the phenomenon of inter-rater agreement: a multi-component approach for in-depth examination was proposed. In Exploration and Development of Methology for Accessibility Assessments (Doctoral thesis) Lund University, Sweden.

SPSS Inc 2009 Introduction to Statistical Analysis Using SPSS Statistics. Chicago, U.S.A.

Steeh C 2008 Telephone surveys. Chapter 12. International handbook of survey methology. E.D.d. Leeuw, J.J. Hox and D.A. Dillman. New York, NY, Lawrence Erlbaum Associates.

Streiner D 2003 Statistical development and applications. Starting at the beginning: An introduction to coefficient alpha and internal consistency. Journal of Personality Assessment 80: 99-103.

Streiner D, Norman G 2008 Health measurement scales - a practical guide to their development and use. Oxford University Press, New York, U.S.A.

Sund T 2004 Når rullestolen er en del av livet. En studie om fornøydhet blant brukere av manuelle rullestoler i aldersgruppen 18-65 år. Hovedfagsoppgave (When the wheelchair is part of life. A study on satisfaction among users of manual wheelchairs in the age group 18-65 years. Master thesis). Oslo, University of Oslo, Oslo, Norway.

Sund T, Iwarsson S, Andersen MC, Brandt Å 2013 Documentation of and satisfaction with the service delivery process of electric powered scooters among adult users in different national contexts. Disability and rehabilitation: Assistive Technology 8: 151-160.

Swaine-Verdier A, Doward L, Hagell P, Thorsen H, McKenna SP 2004 Adapting quality of life instruments. Value Health 7: 27-30.

The Norwegian Labour and Welfare Service NAV. Statistics of assistive technology in Norway. Departement of Assistive Technology NAV, The Norwegian Labour and Welfare Service NAV, Oslo, Norway.

Weir J, 2005 Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. Journal of Strength and Conditioning Research 19: 231-240.

Willis G, Reeve B, Barofsky I 2005 The use of cognitive interviewing techniques in quality-of-life and patient-reported outcomes assessment. In Outcomes Assessment in Cancer: Measures, Methods and Applications, Cambridge University Press.

World Health Organization WHO, The World Bank 2011 The World Report on Disability.

Wressle E, Samuelsson K 2004 User satisfaction with mobility assistive devices. Scandinavian Journal of Occupational Therapy 11: 143-150.