Psychometric properties of the Swedish version of the
General Self-Efficacy Scale in stroke survivors

Short title: Psychometric properties of the General Self-Efficacy Scale

Emma Carlstedt¹, Eva Månsson Lexell¹,², Hélène Pessah-Rasmussen¹,² and Susanne
Iwarsson¹

¹Department of Health Sciences, Lund University, Sweden
²Department of Neurology and Rehabilitation Medicine, Skåne University Hospital,
Sweden

Corresponding author:
Emma Carlstedt, Department of Health Sciences,
Lund University, SE-221 00 Lund, Sweden.

Email: Emma.Carlstedt@med.lu.se

Conflicts of Interest: None declared.

Source of Funding: This work was supported by the Swedish Stroke Association
[Strokeförbundet], the Swedish Research Council for Health, Working Life and Welfare
(Forte; Grant No. 2006-1613), and the Ribbingska Foundation in Lund, Sweden.
Abstract

The objective of this study was to assess the psychometric properties of a Swedish version of the General Self-Efficacy Scale (GSE) in stroke survivors. The GSE was administered by the same assessor on two occasions three weeks apart with 34 stroke survivors (21 men, 13 women; mean age= 68.1) six to ten month after stroke. Psychometric properties including targeting and scaling assumptions, and several reliability indices, were calculated. The mean score was well above the midpoint of the scale, and the total scores spanned almost the whole scale range. Floor and ceiling effects were within the limits of 15–20% for total scores (0% and 8.8%, respectively), but not for each item individually. Total skewness was estimated at -1.02 and skewness for individual items as -1.55 to -0.33. The corrected item-total correlations were all above 0.3 except for one item. Cronbach’s alpha was high (0.92) and the test–retest reliability was acceptable (ICC2,1 =0.82). The mean difference (d) was -0.68 (n.s.). The SEM was 2.97 (SEM%; 9.40). In conclusion, although targeting in relation to skewness and ceiling effects was seen in some items, the GSE was reliable for use in mobile stroke survivors six to ten month post stroke.

Keywords

Cerebrovascular disorders, reliability, reproducibility of results, self-efficacy
Introduction

Annually, around 25,000 people experience a stroke in Sweden, and most of them (80%) are 65 years old or older (The Swedish Stroke Register, 2013). Stroke is among the most common cause of disability in adults and can result in various types of physical, cognitive and psychological impairments. Physical impairments, such as hemiparesis and reduced balance can often impact the individual’s ability to move around (Michael et al., 2005) and different kinds of mobility devices are common. Cognitive impairments – for example, aphasia, reductions in memory, and ability to engage in abstract thinking (Claesson et al., 2005) – are experienced by 64% of those aged 65+ (Jin et al., 2006). A psychological impairment such as depression occurs in approximately one-third of all stroke survivors (Hackett et al., 2005). The different types of impairments following stroke can play a crucial role in an individual’s ability to manage everyday activities in the home as well as in society (Mayo et al., 2002; Combs et al., 2013; Adamit et al., 2014).

One concept that is attracting interest is self-efficacy, which is considered an important aspect of a person’s ability to manage every day activities post stroke (Korpershoek et al., 2011). Self-efficacy is related to both cognitive function (Lewin et al., 2013) and depression and seems to play an important role in the measurement of rehabilitation outcomes post stroke (Jones and Riazi, 2011; Lewin et al., 2013). Self-efficacy is a component of Bandura’s social cognitive theory and can be explained as a person’s belief or confidence in his or her capability to realize a specific task (Bandura, 1997; Bandura, 2006). It reflects the statement that if people do not believe that they have the power to perform a specific assignment, they will not do it (Bandura, 1997). Self-efficacy has become a key factor in self-management programmes to improve health-related behaviour among people with chronic diseases (Lorig and Holman, 2003).
including stroke (Jones et al., 2013). It also seems that improved self-efficacy decreases the use of healthcare among individuals with chronic diseases (Lorig et al., 2001). There are incentives to mitigate the consequences of stroke, both for the individuals affected and for the healthcare cost involved (Claesson et al., 2005; Husaini et al., 2013). Strengthening self-efficacy through self-management programmes might be one way of attaining this, but to do so we need reliable instruments that can evaluate the effects of such programmes, targeting specific groups of the population.

There are different ways to measure self-efficacy. Bandura advocated focusing on self-efficacy in a specific task or situation (Bandura, 1997; Bandura 2006) – for example, by using the Chronic Disease Self-Management Program (CDSMP) (Lorig et al., 1996) or the Stroke Self-efficacy Questionnaire (Jones et al., 2008). Other argues that self-efficacy also can be targeted on a generalized level, thus referring to the individual’s overall confidence in his or her ability to manage different situations in life. The widely accepted General Self-Efficacy Scale (GSE) (Schwarzer and Jerusalem, 1995) is an example of such an instrument. The psychometric properties of the GSE have been analysed using data from 25 countries, showing good internal consistency (α=0.86, range=0.75-0.91) in a large cross-national sample (n=19,120) (Scholz et al., 2002). However, the instrument has been only sparsely used in a Swedish context (Löve et al., 2012; Nilsson et al., 2015). Even though psychometric properties of the GSE have been studied in different languages and study populations, to the best of our knowledge no study has investigated such properties in the Swedish version of the GSE in post stroke survivors.

Accordingly, the overall aim of this study was to assess the psychometric properties of the Swedish version of the GSE in a sample of stroke survivors six to ten month after
stroke. More specifically, the aims were to 1) investigate the targeting and scaling assumptions of the GSE, and 2) to explore the reliability of the GSE.

**Methods**

*Participants*

Thirty-four participants were recruited from a university hospital in the south of Sweden during October 2012-February 2014, including those who had sustained a first-ever stroke as well as those with recurrent strokes. The inclusion criteria were: 1) six to ten months since the most recent stroke; 2) a self-report stating independent indoor as well as outdoor mobility three months post stroke; and 3) an age of 55 or older. Exclusion criteria were insufficient language skills to participate in the data collection. The mean age of the participants were 68.1 years and they were all living in ordinary housing in the community. The sample consisted of 21 men and 13 women, people with/without mobility devices, with/without impaired cognitive function and with/without depression. Participant characteristics are presented in Table 1.

*Insert Table 1 about here*

*Ethics*

All participants received verbal and written information about the study, and gave their informed consent to participate. They were aware that they could withdraw at any time without any consequences for their future care. The Regional Ethical Review Board in Lund, Sweden, approved the study (Ref. No. 2012/174).
The General Self-Efficacy Scale

The GSE scale consists of ten statements such as “I can solve most problems if I invest the necessary effort.” Possible responses are scored 1–4, where 1 represents “Not at all true”, 2 “Hardly true”, 3 “Moderately true” and 4 “Exactly true”. This gives a total score ranging from 10–40; higher scores indicate a greater sense of general self-efficacy (Schwarzer and Jerusalem, 1995). The GSE has been translated into Swedish (Koskinen-Hagman et al., 1999) and has been demonstrated to be valid (Löve et al., 2012).

Procedure

The data collection was made as a part of a wider data collection of several questionnaires. It was performed by the same assessor (first author; EC), on two occasions three weeks apart (± 2 weeks), at the outpatient clinic at the Department of Neurology. At the beginning of each first occasion, the assessor described the background of the study. She then explained the purpose of the questionnaire, read the questions out loud and recorded the participant’s answers on the printed form. Each occasion lasted approximately 30–60 min, whereof the data collection regarding GSE lasted about 10 minutes. All 34 participants completed the test at both T1 and T2 and there were no missing data.

Statistical methods

Descriptive statistics were used to present participant characteristics, including the use of walking devices indoors and outdoors, cognitive impairments and depressive symptoms. Targeting and scaling assumptions, and the Cronbach’s alpha, were calculated based on data from the first test occasion (T1). All other calculations were made based on data from T1 and T2.
Targeting was studied by assessing the score distribution, floor and ceiling effects, and skewness of the scale (Hobart and Cano, 2009). The full scale range should be used in the total sample and the total mean score should stay close to the midpoint of the scale (in the GSE = 25). Floor and ceiling effects should not exceed 15–20 %. The recommended range for skewness statistics is ±1 (Hobart et al., 2004). Floor and ceiling effects and skewness were calculated for the total GES score as well as for the individual items.

Scaling assumptions were evaluated, calculating the distribution of the mean score and the SD of each item—that is, to estimate if the items were taken at the same point of the scale and demonstrated equivalent variance (Hobart and Cano, 2009). Also, the corrected item-total correlations were calculated as an indication of whether each of the items represented the same underlying construct (should exceed 0.3) (Nunnally and Bernstein, 1994; De Vet et al., 2014).

Reliability was examined in different ways. In terms of the internal consistency of the scale, Cronbach’s alpha was calculated, the desired value being between 0.70 and 0.90 (De Vet et al., 2014). The intra-class correlation coefficient (ICC) (Shrout and Fleiss, 1979; McGraw and Wong, 1996) was used to explore test–retest reliability. Two-way mixed (ICC_{2,1}) single measures with absolute agreement (McGraw and Wong, 1996) were used. The ICC should reach 0.80 to be acceptable for group comparisons but must exceed 0.90 when individual scores are of interest (Nunnally and Bernstein, 1994). To capture any systematic or random differences, the mean difference (Δ) between T1 and T2 was measured.

The standard error of measurement (SEM) was computed using the formula $SD_{\text{baseline}} \times \sqrt{1 - \text{reliability}}$ (Weir, 2005; Streiner and Norman, 2008). Since SEM is expressed in
absolute values, $\text{SEM}\% = \left( \frac{\text{SEM}}{\text{mean}} \right) \times 100$ were calculated in order to report a relative value that might be easier to comprehend (Lexell and Downham, 2005).

$P$-values below 0.05 indicate statistical significance. The IBM SPSS Statistics 22.0 software was used for the analyses.

**Results**

The total mean score (31.7) was above the midpoint of the scale, although the midpoint was within the SD (6.95). The total scores spanned almost the whole scale range (13–40). The floor and ceiling effects were 0 and 8.8 %, respectively. When looking at each item separately, the floor effects were 0%–14.7% (mean 7.4 %) and the ceiling effects were 26.5%–61.8% (mean 45.7 %). The ten percent lowest values contained 2.9% of the answers and the ten percent highest values contained 26.5%. The distribution of the total score showed skewness at -1.02. On the item level skewness ranged from -1.55 to -0.33. The means of the ten items were distributed between 2.74 and 3.47. The corrected item-total correlation was satisfying for all but one item, which did not exceed the value of 0.3 (see Table 2).

*Insert Table 2 about here*

All reliability indices are shown in Table 3. The internal consistency illustrated by Cronbach’s alpha was 0.92 and the ICC$_{2,1}$ was 0.82 (CI: 0.67, 0.90). The mean differences ($d$) between T1 and T2 were slightly below zero (-0.68; CI: -2.2, 0.88), although not statistically significant ($p=0.38$). The SEM was 2.97 and the SEM% was 9.40.
Discussion

Overall, the analyses of a set of psychometric properties of a Swedish version of the GSE in stroke survivors six to ten month post stroke suggest that the GSE is reliable for this population, even though not all of the properties were optimal.

The Cronbach’s alpha coefficient was high in the present study and high levels have also been estimated in GSE studies targeting other populations (Scholz et al., 2002; Luszczynska et al., 2005; Nilsson et al., 2015). It should be kept in mind that according to some authors (Steiner and Norman, 2008; De Vet et al., 2014) Cronbach’s alpha values exceeding 0.90 may indicate a redundancy among the items of a scale, and this could be taken as an indication to consider whether all items should necessarily be included. However, with an alpha value just slightly above 0.90, this is not an issue in the present study. The test–retest – that is, ICC₂,₁ – was estimated at an acceptable level at a group level; optimally, however, it should be slightly higher to be seen as acceptable at an individual level. Still, our results are comparable with the results of a recent study (Nilsson et al., 2015) involving people with Parkinson’s disease, in which the estimated ICC was just below our value (0.69–0.80). Furthermore, the mean difference (d) between T1 and T2 in the present study indicated that the participants tended to score lower at T2, but the difference was not statistically significant. While the interpretation of SEM in relation to change scores is intricate, in the sample of the present study a difference of at least 2.97 points is needed to identify a meaningful change, that is, a change that exceeds the measurement error. However, given that measurement error is a random variable with a different value at each measurement occasion and also the large proportions of ceiling effects (discussed below) this result
has to be interpreted with caution, not the least as the SEM in our sample was lower than that found in another study on the GSE (Nilsson et al., 2015).

The fact that the total mean score was above the midpoint of the scale in the present sample is in line with studies involving other stroke samples (Lewin et al., 2013; Omu and Reynolds, 2013). The fact that the item-total correlation exceeded the desirable level for all except one item is positive, but raises some questions of that item (Item 3 – “It is easy for me to stick to my aims and accomplish my goals”). For instance, that question is not as concrete as the other items, but rather abstract. It might be particularly difficult for an individual who recently suffered a brain injury to answer an abstract question regarding aims and goals in life. Furthermore, Scholz et al. (2002) have also reported similar problems, with item-total correlations of less than 0.3 on one item of the GSE in six out of 25 countries, and in two of these cases, it was item 3. As to ceiling effects, which occur separately for all individual items (but not for the total scale), they were present particularly for the items “I can always manage to solve difficult problems if I try hard enough”, “If someone opposes me, I can find the means and way to get what I want” and “I can usually handle whatever comes my way”. Such results are not desirable because the scale gives no room for improvement for items that receive high ratings. Moreover, in the present study the score distribution of the GSE total score as well as the scores for several of the individual items were somewhat negatively skewed, which was not seen at this level in previous studies (Scholz et al., 2002; Nilsson et al., 2015). Taken together, these concerns indicate that further evaluations of the scale in this particular target group are needed. Regarding potential revisions of the instrument, it is important to remember that any diagnosis-specific revisions will form the basis of subsequent comparisons between different populations.
Turning to methodological concerns, this study included a relatively small sample, even if it was within the recommended number of 20–50 participants for reliability studies (Hobart et al., 2012). Still, our sample seems to represent the whole span of the studied subpopulation of stroke survivors – that is, both people with physical, cognitive, and psychological impairments as well as those without such impairments. For example, the sample had the same proportion of participants with reported depressive symptoms (Hackett et al., 2005) and only slightly fewer with cognitive impairments, compared to those of other studies (Jin et al., 2006).

In this study, it was the assessor, rather than the participants themselves, who completed the GSE forms, which could explain the fact that no instances of missing data were found. This could in turn affect reliability, if the questionnaire were used in a different way. However, since we include people with different types of impairments, in order to reduce the sources of variation during administration of the questionnaire, we decided to use the same administration mode throughout.

In conclusion, this study indicates that the Swedish version of the GSE is reliable for use in mobile stroke survivors six to ten month post stroke, although not all of the psychometric properties studied reached the optimal levels. Further research should assess the GSE in suitable samples of stroke survivors to establish if the existing GSE can be used with confidence in the whole stroke population. Nevertheless, since self-efficacy is gaining increasing attention as a target and outcome of rehabilitation interventions, our study contributes to the development of instruments that target self-efficacy.
Acknowledgements

The authors wish to thank all respondents as well as the employees at the Department of Neurology for assistance during the data collection especially Dr Eva Ask and Dr Maria Macek. This study was accomplished within the context of the Centre for Ageing and Supportive Environments (CASE), Lund University, Sweden.
**References**


Table 1. Sample characteristics, N=34.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Men, n (%)</td>
<td>21 (61.8)</td>
</tr>
<tr>
<td>Women, n (%)</td>
<td>13 (38.2)</td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
</tr>
<tr>
<td>Mean (minimum–maximum)</td>
<td>68.1 (58–86)</td>
</tr>
<tr>
<td>Time post stroke, month</td>
<td></td>
</tr>
<tr>
<td>Mean (minimum–maximum)</td>
<td>8.0 (6–10)</td>
</tr>
<tr>
<td>First-ever stroke</td>
<td></td>
</tr>
<tr>
<td>Yes, n (%)</td>
<td>29 (85.3)</td>
</tr>
<tr>
<td>No, n (%)</td>
<td>5 (14.7)</td>
</tr>
<tr>
<td>Stroke subtype</td>
<td></td>
</tr>
<tr>
<td>Infarction, n (%)</td>
<td>33 (97.1)</td>
</tr>
<tr>
<td>Haemorrhage, n (%)</td>
<td>1 (2.9)</td>
</tr>
<tr>
<td>Living condition</td>
<td></td>
</tr>
<tr>
<td>Living with partner, n (%)</td>
<td>22 (64.7)</td>
</tr>
<tr>
<td>Living alone, n (%)</td>
<td>12 (35.3)</td>
</tr>
<tr>
<td>Children living at home, n (%)</td>
<td>3 (8.8)</td>
</tr>
<tr>
<td>Housing condition</td>
<td></td>
</tr>
<tr>
<td>Multifamily house, n (%)</td>
<td>21 (68.8)</td>
</tr>
<tr>
<td>Single-family house, n (%)</td>
<td>13 (38.2)</td>
</tr>
<tr>
<td>Mobility device indoors, n (%)</td>
<td>2 (6.1)</td>
</tr>
<tr>
<td>Mobility device outdoors, n (%)</td>
<td>9 (27.3)</td>
</tr>
<tr>
<td>Self-rated cognitive impairment</td>
<td></td>
</tr>
</tbody>
</table>
Any cognitive impairment\(^b\), \(n\) (%)

19 (55.9)

Number of self-rated cognitive impairment\(^b\),

mean *(minimum–maximum)*

3.2 (0–20)

Geriatric depression scale (GDS)\(^c\)

Total score, mean *(minimum–maximum)*

5.3 (0–19)

Possible depression >5p\(^d\), \(n\) (%)

12 (35.3)

\(^a\) One missing

\(^b\) Study-specific self-rated cognitive impairment questionnaire. According to Wendel *et al.* (2008) this assessment has showed cognitive impairment to a greater extent compared to an objective assessment.

\(^c\) GDS (*Gottfries et al.*, 1997)

\(^d\)>5p= possible depression
Table 2. Means (SD) and corrected item-total correlations (CI) for the ten GSE items, N=34.

<table>
<thead>
<tr>
<th>No.</th>
<th>Item</th>
<th>Mean (SD)</th>
<th>Corrected item-total correlation (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>I can always manage to solve difficult problems if I try hard enough</td>
<td>3.41 (0.86)</td>
<td>0.69 (0.46, 0.84)</td>
</tr>
<tr>
<td>2.</td>
<td>If someone opposes me, I can find the means and ways to get what I want</td>
<td>3.41 (0.86)</td>
<td>0.55 (0.27, 0.75)</td>
</tr>
<tr>
<td>3.</td>
<td>It is easy for me to stick to my aims and accomplish my goals</td>
<td>3.47 (0.71)</td>
<td>0.25 (-0.10, 0.54)</td>
</tr>
<tr>
<td>4.</td>
<td>I am confident that I could deal efficiently with unexpected events</td>
<td>2.74 (1.02)</td>
<td>0.68 (0.45, 0.83)</td>
</tr>
<tr>
<td>5.</td>
<td>Thanks to my resourcefulness, I know how to handle unforeseen situations</td>
<td>3.18 (1.00)</td>
<td>0.66 (0.42, 0.82)</td>
</tr>
<tr>
<td>6.</td>
<td>I can remain calm when facing difficulties because I can rely on my coping</td>
<td>3.15 (0.93)</td>
<td>0.52 (0.22, 0.73)</td>
</tr>
<tr>
<td>7.</td>
<td>I can solve most problems if I invest the necessary effort</td>
<td>3.03 (1.00)</td>
<td>0.68 (0.45, 0.83)</td>
</tr>
<tr>
<td>8.</td>
<td>When I am confronted with a problem, I can usually find several solutions</td>
<td>3.06 (0.85)</td>
<td>0.75 (0.56, 0.87)</td>
</tr>
<tr>
<td>9.</td>
<td>If I am in trouble, I can usually think of a solution</td>
<td>2.94 (1.04)</td>
<td>0.80 (0.64, 0.90)</td>
</tr>
<tr>
<td>10.</td>
<td>I can usually handle whatever comes my way</td>
<td>3.26 (0.86)</td>
<td>0.60 (0.32, 0.78)</td>
</tr>
</tbody>
</table>
Table 3. Reliability indices of the GSE, N=34.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach’s alpha (95% CI)</td>
<td>0.92 (0.86, 0.95)</td>
</tr>
<tr>
<td>a ICC 2,1 (95% CI)</td>
<td>0.82 (0.67, 0.90)</td>
</tr>
<tr>
<td>Mean difference, d (95% CI)</td>
<td>-0.68 (-2.23, 0.88)</td>
</tr>
<tr>
<td>b SEM</td>
<td>2.97</td>
</tr>
<tr>
<td>c SEM%</td>
<td>9.4</td>
</tr>
</tbody>
</table>

*a ICC=Intra-class correlation coefficient (two-way random model, absolute agreement, single measure)*

*b SEM=Standard error of measurement, defined as SEM= SD_{baseline} \times \sqrt{1 - \text{reliability}}*

*c SEM%= (SEM/mean) \times 100*