Coversheet

This is the accepted manuscript (post-print version) of the article.
Contentwise, the accepted manuscript version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication
Please cite the final published version:


Publication metadata

**Title:** A Danish version of self-efficacy in diabetes self-management: A valid and reliable questionnaire affected by age and sex

**Author(s):** Lindkvist, Emilie Bundgaard; Kristensen, Lene Juel; Sildorf, Stine Møller; Kreiner, Svend; Svensson, Jannet; Mose, Anne Hvarregaard; Birkebaek, Niels; Thastum, Mikael.

**Journal:** Pediatric Diabetes

**DOI/Link:** [https://doi.org/10.1111/pedi.12601](https://doi.org/10.1111/pedi.12601)

**Document version:** Accepted manuscript (post-print)

General Rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

If the document is published under a Creative Commons license, this applies instead of the general rights.
Statistical analysis of a Danish SEDM

Corresponding Author

*Mikael Thyssum,

School of Business and Social Sciences

Department of Psychology

Bartholins Allé 9, Dk-8000 Aarhus C

Telephone-number: +45 4023 8559

E-mail address: mikael@psy.au.dk
A Danish version of Self-Efficacy in Diabetes Self-Management: A valid and reliable questionnaire affected by age and sex.

Authors: Emilie Bundgaard Lindkvist1, Lene Juel Kristensen2, Stine Møller Sildorf3, Svend Kreiner4, Jannet Svensson1, Anne Hvarregaard. Mose4, Niels Birkebaek4, Mikael Thastum2

(Shared first authorship by Emilie and Lene)

Affiliations:

1. Copenhagen Diabetes Research Center (CPH-DIRECT), Department of Paediatrics, Herlev University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark

2. Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus, Denmark

3. Section of Biostatistics, Department of Public Health, Faculty of Health and Medical Sciences, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen K, Denmark

4. Department of Pediatrics, Aarhus University Hospital, Aarhus N, Denmark

Word count: 5.326
Abstract:

**Background**

Managing the chronic illness Type 1 Diabetes (T1D) is extremely demanding, especially during adolescence. Self-efficacy is belief in one’s own capabilities and this is crucial for diabetes management. Having a valid method for measuring self-efficacy is important.

**Objective**

Our aims were to psychometrically validate a Danish version of the Self-Efficacy for Diabetes Management (SEDM) questionnaire, and to examine the relationship between background variables and self-efficacy.

**Methods**

All Danish adolescents with T1D (n = 1,075) were invited to participate in our study. In total, 689 agreed to participate and 602 completed the study. Data were collected using a web-based survey. All participants were asked to provide a blood sample for HbA1c measurement. Graphical Log Linear Rasch Modeling (GLLRM) was used to validate the questionnaire and its reliability was assessed using Monte Carlo simulation.

**Results**

We found the questionnaire to be valid and reliable, but it had a dual structure that suggested a need for two separate subscales. One subscale related to practical (SEDM1) and the other to emotional (SEDM2) aspects of diabetes management. Both subscales were targeted toward
adolescents with lower self-efficacy and were associated with HbA1c. SEDM1 was influenced by treatment modality and age. In SEDM2 we found an interaction between age and sex.

**Conclusion**

The Danish version of the SEDM questionnaire should be divided into two parts, each with a valid and reliable subscale for self-efficacy measurement. The relationship between self-efficacy and age seems to differ between boys and girls.

**Keywords**

Adolescent, diabetes mellitus type 1, reproducibility of results, self-efficacy, sex characteristics
**Abbreviations**

T1D: Type 1 diabetes

SEDM: Self-efficacy in Diabetes Management

GLLRM: Graphical Log Linear Rasch Modelling

IRT: Item Response Theory

DIF: Differential Item Functioning

ADQ: Adherence in Diabetes Questionnaire

CLR: Conditional Likelihood Ratio
Introduction

A growing number of children and adolescents live with Type 1 Diabetes (T1D) (prevalence ~0.4%) (1), which is a complex and demanding disease requiring extensive self-care. Adherence to treatment and maintenance of tight metabolic control are essential to prevent complications from developing (2, 3). The high demands of self-managing the disease can affect the individual’s quality of life and psychosocial health (4). Deteriorations in adherence and metabolic control is particularly common during adolescence (5). Social Cognitive Theory defines self-efficacy as an individual’s belief in their own ability to exercise control over the difficulties they face (6). Self-efficacy is correlated with quality of life (QoL) (7, 8), metabolic control (9–11), and adherence to treatment (12, 13). Age, sex, and diabetes duration all influence HbA1c (14), but the association between these variables and self-efficacy has not been adequately explored.

Inspired by the self-efficacy questionnaire developed by Grossman et al. (15), Iannotti et al. developed an updated, shortened scale that measured self-efficacy in adolescents with T1D called the Self-Efficacy in Diabetes Management (SEDM) questionnaire (16). They tested the psychometric properties of the SEDM questionnaire using principal factor analysis and found a one-factor solution. Reliability was tested using Cronbach’s alpha (0.90). Prior to our study, there was no Danish self-efficacy questionnaire.

The aim of this study was to translate the English version of the SEDM questionnaire into Danish and test its criterion-related validity and reliability using the Rasch model (17, 18). We then examined the relationship between self-efficacy and background variables (age, sex, HbA1c, diabetes duration, and treatment). We predicted an increase in self-efficacy with age and higher self-efficacy in girls.
Methods

Overall design

Our study was conducted in 2009 as part of a nationwide web-survey that included all 1,997 Danish children and adolescents (2 to 17 years old) with T1D (13, 19, 20). It contained 15 different questionnaires (including SEDM) and assessed the influence of psychosocial variables on adherence, glycemic control, and quality of life and was conducted in collaboration with the Danish Society for Diabetes in Childhood and Adolescence, which administers the Danish Registry for Childhood and Adolescent Diabetes (DanDiabKids). This registry contains information dating back to 1996 on every child and adolescent with T1D in Denmark. A total of 258 families were excluded because they were unwilling to participate in scientific research, had a protected address, or were not living at their Danish Civil Registration System recorded address. Of the remaining 1,739 two- to seventeen-year-olds, the 1,075 eligible adolescents (aged 12 to 17 years) all received a written invitation to complete the SEDM questionnaire (Fig. 1). The children (aged 2 to 11 years) were excluded because the SEDM questionnaire was developed for and tested to fit adolescents only. The participants were asked to provide a small blood sample for HbA1c measurement. This measurement was performed using a high-pressure liquid chromatographic method (Tosoh Bioscience, South San Francisco, CA, USA) with a working range of 23.5–40 mmol/mol (4.3–5.8%). Informed consent was obtained from participants’ parents and assent was also provided by those adolescents aged from 15 to 17 years old. The study was approved by the Danish Data Protection Agency.

Self-Efficacy of Diabetes Management (SEDM)
The English version of the 10-item SEDM questionnaire (16) was translated into Danish using multiple forward-translations (21). A panel of three health psychologists and one pediatric endocrinologist performed the translations. Their translations were compared and discrepancies were discussed. A consensus version was tested by adolescents with T1D and their caregivers, and a final version generated incorporating their feedback. Responses were provided on an ordinal scale and ranged from 1 (not sure at all) to 10 (completely sure), with a high overall score indicating a high level of self-efficacy. The overall score ranged from 10 to 100 points.

**Statistical method**

For comparison of background variables simple t-tests were applied. The SEDM scale was subjected to a test-of-fit using the Rasch model and the graphical log linear Rasch model (GLLRM). Reliability was tested using Cronbach’s alpha and Monte Carlo simulation (22). The effect of background variables was examined by linear regression with SEDM scores adjusted for differential item functioning (DIF) as dependent variables (23).

**The Rasch Model**

The Rasch model is an item response theory (IRT) model particularly appropriate for the assessment of summated scales (17, 24). In addition to addressing the same issues as classical test theory (CTT) (e.g., reliability, measurement error) it also addresses issues that CTT cannot, for instance local dependency between items and differential item functioning. For an item to fit a Rasch model it must exhibit different properties (i.e., criterion-related constructs). These include:

i) unidimensionality, ii) monotonicity, iii) local independence, and iv) no differential item functioning (DIF) (Table 1). Health-related scales seldom conform to all four criterion-related
constructs. In these cases, the GLLRM, which relaxes the necessity for local independence and no DIF, may be applied instead (18, 24). A fifth property, known as statistical sufficiency, requires the total score to capture all the information available on the person parameters and differentiates the Rasch model from other IRT models.

**Reliability**

Reliability assessed by Cronbach’s alpha produces an inflated value when there is local dependence among the items. An unbiased estimate of true reliability can be obtained using the Monte Carlo method (22). Different reliabilities must be calculated for groups affected by DIF or different background variables.

**Targeting**

Targeting tells us to what degree the included items uncover the desired participant information. It also tells us whether the items are skewed toward very large or small participant values. To assess whether targeting was adequate, the average test information for the population was calculated and compared with the largest possible test information (23).

**Correlation between self-efficacy and adherence**

The Adherence in Diabetes Questionnaire (ADQ) (13) was also included in the web-survey and the correlation between this and the SEDM questionnaire was examined using Pearson’s correlation.

**Assessment of significance**
Significance was assessed at the 5% level following adjustment for multiple testing using the Benjamini-Hochberg procedure (25). We distinguished between weak to moderate evidence with $p$-values ranging from 0.01–0.05, and strong evidence with $p$-values below 0.01 (26).

**Statistical software**

The item analysis by Rasch models and GLLRM was performed using DIGRAM (27, 28).
Results

Demographic data

A total of 689 (64.1%) of the 1,075 eligible adolescents participated in the study and the remaining 386 failed to respond. Of the 689 participants, complete SEDM scores were obtained for 602 (87.4%) and data for 87 adolescents (12.6%) were missing. Blood samples were provided by 580 (84.2%) participants. There were no significant differences between the HbA1c values generated for the study (M = 8.23 % (~ 66.4 mmol/mol), standard deviation [SD] = 1.24) and those obtained from the DanDiabKids registry for the same participants (M = 8.21 % (~ 66.2 mmol/mol), SD = 1.28, t(579) = 0.65, p = 0.52). Therefore, the DanDiabKids’ HbA1c values were used for those participants who did not provide a blood sample. We found no differences between those completing the entire SEDM (responders) and those with missing SEDM values (incomplete-responders) at the 0.01 level of significance (Table 4a and 4b).

Validity

Unidimensionality

Attempts to fit a unidimensional model to the ten items failed because there was evidence of positive and negative local dependence, indicating a multidimensional structure (29). Therefore, the items were separated into two sets based on question type. The first set of questions (SEDM1; items 1–4) related to practical diabetes management, whereas the second set (SEDM2; items 5–10) related to the emotional capacity of each adolescent to cope with managing their condition. For unidimensionality, the expected correlation between the subscales would have been 0.53. However, the observed correlation was significantly lower (0.46, p < 0.001), confirming that more
than one dimension of self-efficacy was being measured by the ten items. Nonetheless, unidimensionality was confirmed within each subscale (29) and so the remaining statistical analyses were performed on each of the two subscales separately.

**Monotonicity**

The assessment of monotonicity is presented in terms of item-fit statistics (Table 2). The Rasch model assessment showed large differences between observed and expected $\gamma$-values indicating that monotonicity was not fulfilled and so this model was rejected. Applying the GLLRM resulted in a significant difference for item ten, although the item still functioned at an acceptable level and was not excluded. No significant differences were found for the remaining items, meaning that monotonicity was fulfilled when the GLLRM was applied.

**Local dependence and DIF**

Conditional likelihood ratio (CLR) testing identified local dependence between SEDM1 items 1 and 4 (Fig. 2a). For SEDM2, all items showed local dependence except item 5 (Fig. 2b). No DIF was found in SEDM1 (i.e., the model fitted all items equally well within the different subpopulations; Table 4a). In SEDM2, DIF relating to sex was found for items 5 and 9 (Table 4b).

**Reliability**

Cronbach’s alpha values were measured for both subscales (SEDM1 = 0.74, SEDM2 = 0.88). SEDM1 was associated with both HbA1c and treatment, and we expected reliability to differ in subpopulations defined by these variables. The Monte Carlo estimates of reliability in different subpopulations ranged from 0.44 among adolescents with HbA1c $\leq$ 7 and no insulin pump
(variation among these adolescents was very low) to 0.83 among adolescents with HbA1c ≥ 10 who had an insulin pump. For SEDM2, Monte Carlo estimates were made in subpopulations defined by sex and HbA1c. These calculations estimated reliability at 0.77–0.84 depending on the subpopulation.

**Targeting**

Both SEDM subscales were off target and aimed at adolescents with lower self-efficacy than our study population. Therefore, the average test information obtained was approximately 75% of what SEDM items could have provided with better targeting (23). Consequently, SEDM measurements were less precise for adolescents with a relatively high degree of self-efficacy (i.e., the ceiling effect), but almost optimal for adolescents with low self-efficacy.

**Clinical outcome**

**The SEDM1 score**

Analyzing the effect of background variables on SEDM1 score demonstrated that sex ($p = 0.74$) and diabetes duration ($p = 0.75$) were not associated with SEDM1. There was no evidence of any interaction between age and sex ($p = 0.13$). SEDM1 score was negatively associated with HbA1c and reduced by 0.8 points (95% CI: –1.27, –0.36) when HbA1c increased 1% (approximately 11 mmol/mol; $p = 0.01$). The SEDM1 score was 1.4 (95% CI: 0.21, 2.53) points greater on average in pump than in pen users ($p = 0.021$). SEDM1 score also increased by 0.36 points (95% CI: 0.01, 0.72) for each additional year of age ($p = 0.046$), depending on HbA1c and treatment (Fig. 2a).
The SEDM2 score

After adjustment for DIF, the associated background variable analysis demonstrated that SEDM2 score was associated with HbA1c ($p < 0.001$) and decreased by 3.0 points (95% CI: 2.28, 3.74) when HbA1c increased by 1% (approximately 11 mmol/mol). Treatment ($p = 0.37$) and diabetes duration ($p = 0.66$) were not associated. Any association between SEDM2 score and sex or age was marginal and insignificant. However, multivariate regression analysis provided evidence of an interaction between the effects of age and sex ($p = 0.004$) suggesting that sex modified the effect of age (and vice versa). Therefore, the effect of age on SEDM2 score was analyzed separately in girls and boys (Fig. 3a and 3b). The average SEDM2 score for boys was 42.1 (including all ages), but was higher in girls (average SEDM2 = 45.9) than in boys at age 12. In girls, age had a significantly negative effect on SEDM2 ($p = 0.011$), which decreased by 1.1 points per year of age. The opposite effect was observed in boys, with SEDM2 increasing insignificantly by 0.68 points per year of age ($p = 0.12$). Therefore, at 17 years old, the average SEDM2 score for girls (40.5) was lower than that of boys of the same age. The seen associations were independent of HbA1c values.

Correlation between SEDM1/SEDM2 and the ADQ

Both subscales were significantly correlated with the ADQ (SEDM1: 0.5, $p < 0.001$; SEDM2: 0.6, $p < 0.001$) (13). The differences in correlation were insignificant and had no practical implications.
**Discussion**

The Danish version of the SEDM questionnaire proved to be a statistically valid and reliable tool. There were two correlated constructs being measured, resulting in a separation of the scale. Both subscales correlated strongly with HbA1c. Diabetes-management self-efficacy is affected by age, and the effects of age on the emotional aspects of self-efficacy differed between boys and girls.

**Psychometric validation**

*Validity*

The assessment of unidimensionality demonstrated that two different constructs were being measured and that the scale should be split. This finding differs from the original scale validation (16) and could be explained by cultural, linguistic or medical variances or, more probably, by differences in the statistical approach. Originally, principal factor analysis was applied (16) and revealed a one-factor structure (minimum loading of 0.53) and a one-construct scale. Principal factor analysis is primarily used when developing questionnaires (30–32), whereas we used Rasch and GLLRM to test the basic assumptions, e.g. unidimensionality (31). The current SEDM scale originated as a 35-item scale with three conceptual subscales: diabetes-specific, medical-related, and general self-efficacy (15). This was then revised to a 42-item questionnaire with five conceptual subscales: insulin-administration routines, self-care adjustment, dietary-routine, blood glucose testing frequency, and exercise frequency. Finally, it was reduced to a ten-item scale, loading to a single factor (16). Unfortunately, it is not possible to determine whether the two facets of the SEDM, revealed by the test for unidimensionality, are products of the earlier, longer questionnaires with different subscales because the original item subdivisions are not specified in the existing literature. However, it is clear that the ten items examine different aspects of self-
efficacy (i.e., practical and emotional) and this is supported by our statistical analysis (see appendix).

Because monotonicity was demonstrated, removing items from the scales is unnecessary. Item ten differed from the others, and if it differs in future studies, perhaps it should be excluded or rephrased.

Local dependence and DIF

Local dependence was observed in both subscales and probably results from similarities among the items (see appendix). This is very common in health-related scales, and is not a problem when the correct statistical analyses are applied (17). It tells us that item-overlap is preventing us obtaining 100% of the intended information. Scales with excessive amounts of local dependence (e.g., SEDM2) exhibit signs of statistical redundancy. This means that the scale may be shortened without threatening its criterion-related validity. However, this must be done cautiously and with due regard for content validity, which is the degree to which the scale represents the trait it was designed to sample (33).

Our analysis demonstrated that DIF (relating to sex) was associated with items 9 and 5 (see appendix). This means that a participant’s sex affects how they understand/respond to these two items and that a statistical adjustment may be necessary when comparing scores between the sexes.

Reliability

The reliabilities measured by Cronbach’s alpha (0.74 and 0.88) were satisfactory, comparable with previous studies (30–32), and similar to the original value of 0.90 (16). When there is DIF, true
reliability should be calculated by applying the Monte Carlo method. Despite greater variation, the two subscales still had acceptable reliability. However, this finding cannot be compared with earlier studies because the method has not been applied before.

Targeting was not optimal for either of the subscales and in both cases the population was toward the upper end of the scale. An optimal questionnaire is targeted mid-scale, but this is often impossible in questionnaires where the likelihood of answering ‘always’ or ‘never’ is high. Other self-efficacy studies had the same targeting problems (34–36). The sub-optimal targeting is a limitation when the scale is used to show changes in individuals with high self-efficacy. However, most frequently, the aim is to show an improvement in individuals with low self-efficacy and our sub-optimal targeting is less of a problem in this case (9–11). Therefore, we believe it is unnecessary to adjust the scale.

Subgroups and self-efficacy

Both subscales showed significant correlations with HbA1c, corresponding to the findings of Iannotti et al. (16) and other studies on self-efficacy (9–11). Tests for an association between the English version of the SEDM questionnaire and treatment type were not performed, but our study was consistent with a Norwegian study in finding more negative perceptions of treatment control (SEDM1) in pen-users compared with pump-users (37). The SEDM1 score, which mainly measures practical aspects of diabetes management, was indirectly associated with age. An increasing self-efficacy in older individuals is desirable because adolescents must learn to take responsibility for their own self-management and liberate themselves from their parents during this transitional phase (38).
It is worrying that the emotional capacity for diabetes management (SEDM2) in girls appears to decline with increasing age. This finding is in line with other studies demonstrating that girls have a more negative perception of their illness, more concerns about insulin treatment (37), more severe diabetes distress (feeling overwhelmed, lacking motivation, and feelings of failure) (39), more worries, and poorer health perceptions (8). These factors have also been linked to increased levels of depression (40). Therefore, girls may be more vulnerable than boys and this should be taken into account in the clinical setting.

Association with adherence

In a previous article we tested the psychometric validity of the ADQ, and its association with the SEDM questionnaire (13). Because of the division of the SEDM into the two subscales, we wished to retest the association. As the ADQ is concerned primarily with practical adherence to the diabetes treatment regime, we speculated that SEDM1 (which relates to the practical aspects of diabetes management) would be more closely associated with the ADQ than SEDM2. However, we found no signs of a difference in correlation between the two subscales.

Strengths and limitations

The main strength of our study is the sample size, which provides increased statistical power. We also found no signs of selection bias among the participating responders and incomplete-responders. In addition, we used robust statistical methods that included the Rasch model and GLLRM.

In all questionnaire studies, there is a risk of selection bias. In the grand study including children and adolescents, we found that non-participants had higher HbA1c levels, indicating that those
with the poorest ability to regulate their condition, and possibly the worst psychosocial health, did not take part (13). However, any subgroup selection will predominantly influence targeting, not test validity or reliability. The translational process did not include backwards translation as is recommended (21), which could introduce a small bias.

Clinical implications and relevance

Psychosocial factors affect the extent to which adolescents manage their diabetes, and being able to measure these factors with properly tested and validated tools is essential (17, 21, 41). Self-efficacy is a particularly important psychological construct because psychological interventions can improve self-efficacy, enhancing metabolic control and quality of life (42, 43). Implementing these scales in the clinic is easy because very little training of healthcare professionals is required. Additionally, the scale is short and therefore the burden on adolescents is small. However, it is essential that the scale is split into two because the subscales measure different aspects of self-efficacy. Therefore, separate scores should be calculated and evaluated. One or both subscales may be used, depending on what each individual center needs to determine. In addition, DIF means that SEDM2 scores from boys should be adjusted using the table provided (see appendix) when comparing boys with girls. Before using the scale in non-Danish settings, it should be retested using the same statistical approach as to identify potential necessary country specific adjustments.

Conclusion

In conclusion, the Danish version of the SEDM questionnaire had high validity and reliability when adjusted for DIF and local dependency. Although further studies are needed to replicate our
findings, the differences in self-efficacy relating to age and sex generate valuable insight in self-efficacy research.

Acknowledgements

The authors thank the Danish Society for Diabetes in Childhood and Adolescence for their cooperation and for their contribution of data from the Danish Registry of Childhood Diabetes. The authors are particularly grateful to all the families who took the time to participate in this survey. The authors also thank Novo Nordisk for supporting the survey and the Research Centers at Herlev and Gentofte Hospitals.
17. Kreiner S, Christensen KB (2007) Validity and Objectivity in Health-Related Scales:
Table 1 Description of the criterion-related-constructs

<table>
<thead>
<tr>
<th>Criterion-related-construct</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unidimensionality</strong></td>
<td>If the scale is unidimensional, it only measures one latent variable (e.g., self-efficacy).</td>
</tr>
<tr>
<td><strong>Monotonicity</strong></td>
<td>If items are monotonic, the item score should increase when the latent variable increases.</td>
</tr>
<tr>
<td><strong>Local independence</strong></td>
<td>If the scale is locally independent, then the item responses are independent of each other but dependent on the latent variable.</td>
</tr>
<tr>
<td><strong>Differential item functioning</strong></td>
<td>In a scale with no DIF, all item responses are independent of background variables and only dependent on the latent variable.</td>
</tr>
</tbody>
</table>

*Note: DIF stands for Differential Item Functioning.*
Table 2a Distribution of participants according to sex and treatment

<table>
<thead>
<tr>
<th></th>
<th>Total  (n = 689)</th>
<th>Responders (n = 602)</th>
<th>Non-responders (n = 87)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>49.2%</td>
<td>49.7%</td>
<td>46.0%</td>
<td>0.52</td>
</tr>
<tr>
<td>Girls</td>
<td>50.8%</td>
<td>50.3%</td>
<td>54.0%</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pen</td>
<td>62.8%</td>
<td>61.6%</td>
<td>71.3%</td>
<td>0.082</td>
</tr>
<tr>
<td>Pump</td>
<td>37.2%</td>
<td>28.4%</td>
<td>28.7%</td>
<td></td>
</tr>
</tbody>
</table>
Table 2b Distribution of participants according to age, HbA1c, and duration of diabetes

<table>
<thead>
<tr>
<th></th>
<th>Total  (n = 689)</th>
<th>Responders (n = 602)</th>
<th>Non-responders (n = 87)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>Mean (sd)</td>
<td>14.62 (1.63)</td>
<td>14.67 (1.60)</td>
<td>14.28 (1.75)</td>
</tr>
<tr>
<td><strong>HbA1C</strong></td>
<td>Mean (sd)</td>
<td>8.25 (1.25)</td>
<td>8.21 (1.26)</td>
<td>8.51 (1.16)</td>
</tr>
<tr>
<td><strong>Diab. Duration (years)</strong></td>
<td>Mean (sd)</td>
<td>6.00 (3.46)</td>
<td>6.09 (3.48)</td>
<td>5.40 (3.30)</td>
</tr>
</tbody>
</table>

sd: standard deviation
Table 3: Item fit statistics, observed vs. expected

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Item</th>
<th>Observed $\gamma$</th>
<th>Expected $\gamma$</th>
<th>$p$</th>
<th>Observed $\gamma$</th>
<th>Expected $\gamma$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale 1</td>
<td>SEDM_1</td>
<td>0.422</td>
<td>0.398</td>
<td>0.43</td>
<td>0.422</td>
<td>0.474</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>SEDM_2</td>
<td>0.425</td>
<td>0.441</td>
<td>0.027</td>
<td>0.425</td>
<td>0.393</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>SEDM_3</td>
<td>0.414</td>
<td>0.454</td>
<td>0.013</td>
<td>0.414</td>
<td>0.406</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>SEDM_4</td>
<td>0.524</td>
<td>0.423</td>
<td>&lt;0.001</td>
<td>0.524</td>
<td>0.477</td>
<td>0.09</td>
</tr>
<tr>
<td>Scale 2</td>
<td>SEDM_5</td>
<td>0.459</td>
<td>0.586</td>
<td>&lt;0.001</td>
<td>0.459</td>
<td>0.466</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>SEDM_6</td>
<td>0.567</td>
<td>0.601</td>
<td>0.10</td>
<td>0.567</td>
<td>0.553</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>SEDM_7</td>
<td>0.622</td>
<td>0.574</td>
<td>0.042</td>
<td>0.622</td>
<td>0.621</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>SEDM_8</td>
<td>0.708</td>
<td>0.577</td>
<td>&lt;0.001</td>
<td>0.708</td>
<td>0.674</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>SEDM_9</td>
<td>0.610</td>
<td>0.594</td>
<td>0.46</td>
<td>0.610</td>
<td>0.657</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>SEDM_10</td>
<td>0.649</td>
<td>0.576</td>
<td>0.002</td>
<td>0.649</td>
<td>0.584</td>
<td>0.008</td>
</tr>
</tbody>
</table>
Table 4a: Test of fit, SEDM1

<table>
<thead>
<tr>
<th>Test criterion</th>
<th>Rasch model</th>
<th>Graphical log linear Rasch model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CLR</td>
<td>Df</td>
</tr>
<tr>
<td>Low and high score groups</td>
<td>49.2</td>
<td>34</td>
</tr>
<tr>
<td>HbA1c</td>
<td>114.2</td>
<td>102</td>
</tr>
<tr>
<td>Treatment</td>
<td>35.2</td>
<td>34</td>
</tr>
<tr>
<td>Age</td>
<td>57.3</td>
<td>68</td>
</tr>
<tr>
<td>Sex</td>
<td>53.0</td>
<td>34</td>
</tr>
<tr>
<td>Diabetes duration</td>
<td>117.5</td>
<td>102</td>
</tr>
</tbody>
</table>
### Table 4b: Test of fit, SEDM2

<table>
<thead>
<tr>
<th>Test criterion</th>
<th>Rasch model</th>
<th>Graphical log linear Rasch model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CLR</td>
<td>Df</td>
</tr>
<tr>
<td><strong>Low and high score groups</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLR</td>
<td>126.6</td>
<td>53</td>
</tr>
<tr>
<td>HbA1c</td>
<td>185.9</td>
<td>159</td>
</tr>
<tr>
<td>Treatment</td>
<td>46.6</td>
<td>53</td>
</tr>
<tr>
<td>Age</td>
<td>112.1</td>
<td>106</td>
</tr>
<tr>
<td>Sex</td>
<td>99.9</td>
<td>53</td>
</tr>
<tr>
<td>Diabetes duration</td>
<td>185.0</td>
<td>159</td>
</tr>
</tbody>
</table>
Figure 1 Flowchart depicting inclusion of study participants
Figure 2a IRT-graph for SEDM1.
Figure 2b IRT-graph for SEDM2.
Figure 3: DIF equated score for boys and girls.
Legends:

Table 1:

Description of the criterion-related-constructs

Table 2a+b:

The tables show the differences in background variables for the responders and non-responders (i.e. those responding to all the SEDM-items and those with missing items) among participants. Apart from HbA1c none of the variables show any statistically significant differences as calculated by t-tests.

Table 3:

The table shows the item-fit statistics for each question in the SEDM scale. Question number ten seems to measure a different construct than item one to nine; however, the item still functioned at an acceptable level and was not excluded.

Table 4a+b:

The table shows the differential item function (DIF) for the background variables. Only sex showed DIF for item five and nine. This means that adjustment is needed if self-efficacy is compared between boys and girls.

Figure 1: Shows the flowchart of participants in the study.

Figure 2a+b: The figure shows the item response theory (IRT) graph for the two subscales. The SEDM1 scale has local dependence between item 4 and item 1 but no DIF. The scale is dependent on HbA1c and treatment (pen/pump), age and diabetes duration was indirectly associated with self-efficacy mediated through HbA1c and treatment. The SEDM2 scale has local dependence among every item except item 5. There is DIF between sex and item 5 and item 9. The scale is dependent on HbA1c, the effect of all other background variables is mediated through HbA1c and each other.

Figure 3:

Figure 3 shows the association between age and the DIF-equated SEDM2 score for (a) girls and (b) boys. The scatterplots show the estimated linear regression with SEDM2 as the dependent, and age as the independent variables. The horizontal lines in the upper and lower plots correspond to the average SEDM score for girls and boys, respectively.