Teaching and Learning in Nursing 21 (2026) e823–e829

Contents lists available at ScienceDirect

Teaching and Learning in Nursing

journal homepage: www.journals.elsevier.com/teaching-and-learning-in-nursing

Research

# Measuring self-regulation to inform interventions: The Icelandic adaptation of the self-regulation subscale (SRS) for higher education students in health sciences

Asta B. Schram[1]*, Sumeyra Sahbaz[2]

[1] School of Health Sciences, University of Iceland, Reykjavik, Iceland
[2] Bern University of Teacher Education, Switzerland

ARTICLE INFO

*Article History:*
Accepted 10 January 2026

*Keywords:*
health sciences
higher education
nursing
self-control
self-regulated learning
self-regulation
SPOCK

ABSTRACT

*Background:* Self-regulation is the key to successful learning. It is important to gain students´ level of self-regulation using tools with valid and reliable scores that could provide insights for developing and testing intervention strategies that enhance performance.
*Aim:* The purpose was to adapt the eight-item self-regulation scale (SRS), a part of the student perceptions of classroom knowledge-building scale (SPOCK), for Icelandic health sciences students.
*Methods:* 234 students in nursing and pharmacy responded to an online survey. Classical test theory (CTT) was applied to examine the reliability and validity of the scores and item response theory (IRT) to conduct an item-level evaluation of the psychometric properties, focusing on item difficulty and item discrimination. Differential item functioning across disciplines were assessed.
*Results:* The findings indicated that the adapted SRS is a robust, internally consistent scale. It demonstrates a strong unidimensional structure with high factor loadings, highly discriminative items, and uses response options that effectively assess self-regulation.
*Conclusion:* The SRS-Icelandic should prove useful in testing interventions developed to increase self-regulation for nursing and pharmacy students. Studies are needed in other fields.


## Introduction

Self-regulation plays a valuable role in successful learning within higher education (Russell et al., 2020). Given its importance, instruments that yield reliable and valid scores to assess students´ self-regulation skills and to inform the design of targeted interventions are needed. There is a shortage of research on how to design interventions that can foster self-regulation in learning for university students (Russel et al., 2020). For example, little is known about how clinical teachers in nursing can help their students become self-regulated, although it is clear that if the teachers themselves are self-regulated, they are more effective (Filice et al., 2020). Self-regulated learning stresses students´ autonomy or active roles and is strongly related to student motivation (Efkliedes, 2011; Zimmerman, 2008). Health sciences students need to become autonomous, self-motivated learners who can pursue their fields of study and become able practitioners. Measures that gain students´ levels of self-regulation could aid in the development and testing of interventions that might strengthen students´ participation and raise their level of self-regulatory skills, such as task analysis, goal setting, self-talk, self-reinforcement, and self-monitoring (Schunk & Zimmerman, 2023). The aim of this study was to translate and adapt the self-regulation scale (SRS) from the student perceptions of classroom knowledge-building scale (SPOCK) (Shell et al., 2005; Shell & Husman, 2008), for use with Icelandic university students in health sciences. This scale was selected because it aligned with our objectives for a larger, longitudinal study that we were conducting, one that focused on motivation and self-regulation.

Self-regulation in learning has been defined by Zimmerman (2000, p. 14) as "self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals." Building on this definition, Schunk et al. (2014, p. 158) described self-regulation as "the process whereby students activate and sustain cognitions, behaviors, and affects that are systematically oriented toward attainment of their goals." The self-regulation process is often conceptualized as comprising three cyclical phases: forethought phase, performance phase, and self-reflection phase (Zimmerman & Schunk, 2001). In the forethought phase, students analyze the task, set goals, and select appropriate learning strategies. During the performance phase, they implement these

*Corresponding author.
*E-mail address:* astabryndis@hi.is (A.B. Schram).

https://doi.org/10.1016/j.teln.2026.01.005

strategies—such as time management, self-instruction, imagery, environmental structuring, and help-seeking—to work toward their goals. In the self-reflection phase, students evaluate and monitor their learning progress and decide on future actions based on their perceived success. This evaluation then informs the next cycle, as students re-enter the forethought phase and set new goals. Thus, the self-regulation process is an ongoing, dynamic process.

Research on self-regulation often emphasizes the role of metacognition and the importance of strategy use in learning—for example, through techniques such as question-asking, note-taking, and progress monitoring. Effective self-regulation requires students to develop self-awareness, enabling them to recognize their emotions, thoughts, and reactions, monitor these internal processes, and select appropriate actions in response. Key self-regulatory skills include task analysis, goal setting, self-talk, self-reinforcement, and self-monitoring. However, students differ significantly in their ability to apply these skills. Assessing students' self-regulation and goal-pursuing skills can help educators identify those who may benefit from targeted interventions. Certain interventions have been shown to improve learning outcomes (Filice et al., 2020; Lin et al., 2011; Russel et al., 2020; Schunk & Zimmerman, 2023; Zimmerman, 2013). For instance, supporting students in identifying short- and long-term goals, breaking down complex tasks into manageable steps, and providing regular, constructive feedback has been found to enhance self-regulatory behavior (Cleary & Zimmermann, 2004). A quasi-experimental study on the effect of a metacognitive intervention (concept mapping) in an undergraduate nursing program showed a deeper approach to learning and an increase in self-regulation on behalf of the experimental group in comparison to the control group (August-Brady, 2005). In another study, a pilot randomized controlled trial, the innovative strategy of augmented reality used as an intervention in a nursing program (distance learning), significantly improved some subdomains of SRL competency, e.g., task strategies, time management, help seeking, and self-evaluation (An et al., 2022). Moreover, various motivational beliefs influence the self-regulation process. Investigating self-regulation in conjunction with other motivational factors across different course environments can deepen our understanding of students' learning behaviors and inform strategies to foster greater engagement and self-directed learning.

Various instruments have been developed to assess self-regulation, including the motivated strategies for learning questionnaire (MSLQ) (Pintrich et al., 1993), the learning and study strategies inventory (the LASSI) (Weinstein et al., 1988), the Study Process Questionnaire (Biggs et al., 2001), and the SPOCK (Shell et al., 2005; Shell & Husman, 2008). These instruments differ in their scope and focus. For instance, LASSI shares similarities with SPOCK in measuring students' control of learning and initiative, although SPOCK is course-specific, whereas LASSI provides a more holistic perspective. Researchers often select or combine these inventories based on their research questions (e.g., Hamman, 1998). The purpose of this paper is to describe the translation and adaption of the eight-item Self-Regulation Subscale (SRS) from SPOCK (Shell et al., 1996, 2005; Shell & Husman, 2008). As far as we know, the SRS subscale has not been translated into other languages, but the original English version been used, for example in Taiwan and Singapore (Lin et al., 2011).

## Methods

### *Instrument*

SPOCK was developed as part of Project CIRCLE (Shell et al., 1996) and demonstrates satisfactory internal consistency, with Cronbach`s alpha coefficients ranging from 0.65 to 0.91. These values are comparable to those of similar subscales in other instruments mentioned above. SPOCK assesses four dimensions of students' classroom perceptions: self-regulation, knowledge building, question asking, and lack of regulation. The self-regulation subscale captures strategic self-regulation by asking students about behaviors such as planning, goal setting, monitoring, and evaluating learning within a specific course. The knowledge building subscale assesses the extent of students' seeking and integration. The question asking subscale measures students' efforts to obtain information. The lack of regulation subscale evaluates students` difficulties in knowing how to study and their perceived need for support and guidance. Responses are recorded on a 5-point Likert scale ranging from 1 (Almost never) to 5 (Almost always). Scores for each subscale are calculated by summing the item responses and dividing by the number of items in the subscale. Scores for self-regulation, knowledge building, and lack of regulation have been found to correlate with corresponding LASSI subscales, providing evidence for convergent validity (Hamman, 1998; Weinstein et al., 1988). SPOCK has been widely used to examine self-regulation and its associations with various learning-related constructs. Hamman (1998), for example, employed all four subscales to investigate the extent to which pre-service teachers applied self-regulation strategies in their learning. However, it is common for researchers to use individual subscales independently. Flanigan et al. (2023) utilized only the lack of regulation subscale to assess ineffective self-regulated learning. Stump et al. (2009) used two subscales to examine the relationship between students` beliefs about intelligence and their learning. Lee & Turner (2018) applied four items from the knowledge building subscale to measure pre-service teachers' use of extensive knowledge integration strategies. Similarly, Shell et al. (2013) used three subscales to study the relationships among motivation, self-regulation, and creative competence in relation to long-term learning and achievement. Similarly, Shell and Soh (2013) also utilized subscales to explore the connection between strategic self-regulation and motivation. For the present study, the selected subscale for translation and adaptation was the eight-item SRS from SPOCK (Shell et al., 1996, 2005; Shell & Husman, 2008). This scale assesses students` use of strategies such as planning, goal setting, and monitoring progress in learning contexts. The items align with those from other prominent self-regulated learning models (e.g., Pintrich, 2004; Pressley et al., 1987; Weinstein & Mayer, 1986). The scale has demonstrated strong internal consistency, with Cronbach's alpha values of 0.81 (Shell & Husman, 2008) and 0.84 (Jones & Carter, 2019).

### *Participants and sample size*

An online survey was sent to the 2nd year nursing students ($N$ = 171) and all undergraduate pharmacy students ($N$ = 157). Respondents were 234 students with a nearly equal distribution from the two academic disciplines: nursing ($N$ = 133; 78% response rate) and pharmacy ($N$ = 101; 64% response rate). Of the participants, 90% identified as females, and 78% were between 20 and 27 years of age. The sample was considered representative of the target population, as more than 90% of nursing students identify as females, and students in both departments are predominantly in their twenties. By opening the link to the survey and answering the questions, students agreed to anonymous participation and to the use of the data for publication purposes. Determining an appropriate sample size is critical, particularly when evaluating the validity and reliability of scores generated by self-report measures. However, there is no clear consensus regarding the minimum acceptable sample size. Some experts recommend a minimum of 100-200 participants per group (Lai et al., 2005; Zumbo, 1999), while others suggest that a minimum of 200 respondents per group is necessary to ensure adequate statistical power (>80%) and acceptable model performance (Scott et al., 2009). For the adaptation of a simple model—such as one with one to three factors and fewer than 30 items—a sample size of 200-300 participants is generally considered sufficient (Scott et al., 2009). Given

that the SRS is a one-factor model consisting of only eight items, the group size in the present study was deemed adequate for the intended analysis.

### *Scale development and quality criteria*

Translating a measure from one language to another and adapting it to a specific cultural context presents significant challenges. Cultural nuances and literal translations can substantially affect the validity of the results and may lead to misinterpretations (Beck et al., 2003). The etic approach, which involves translating and adapting an existing instrument, is generally less demanding than the emic approach, where researchers immerse themselves in the target culture to develop a new measure (Villagran & Lucke, 2005). The etic approach is also considered less time-consuming and more cost-effective (Church & Lonner, 1998), and practical to use when the source and target cultures are similar.

A major challenge in the translation process is ensuring that the words selected in the translation reflect the meaning of the original instrument and that the content and structure of the items are appropriate for the target culture. Experts have identified two key criteria for high-quality translation: semantic equivalence, addressing the development of the instrument, and translation invariance, i.e., how well the instrument functions for the target group (Villagran & Lucke, 2005). Semantic equivalence involves both the denotative (literal) and connotative (implied) meanings of words (Beck et al., 2003; Behling & Law, 2000). The connotative meaning includes cultural and emotional associations. Both aspects must be carefully considered, as constructs often vary in meaning across cultures and may evolve over time. Failure to address connotative differences can lead to inaccurate interpretations of results (Beck et al., 2003). To enhance semantic equivalence, Villagran and Lucke (2005) recommend one of the most frequently used back-translation method for cross-cultural instrument translation (Behling & Law, 2000; Brislin, 1986).

In the present study, this method was employed: an initial translation of the SRS from English to Icelandic was completed by a bilingual translator. A second bilingual individual then translated the Icelandic version back into English. A third bilingual expert, familiar with both cultures, compared the two English versions. Discrepancies were resolved collaboratively by all three translators to ensure accuracy and cultural relevance.

The second key criterion, translation invariance, referring to the extent to which the translated instrument functions equivalently across cultures, is evaluated by examining the psychometric properties of the translated measure, typically through confirmatory factor analysis (CFA) (Widaman & Reise, 1997) and item response theory (IRT) (Moustaki & Knott, 2000).

In the present study, both classical test theory (CTT) and IRT were applied to assess the psychometric properties of the SRS. CTT was used to evaluate the scale-level properties, while IRT was used for item-level analyses. Specifically, we examined reliability, validity, factor structure, item parameters (discrimination and difficulty), and differential item functioning (DIF). To provide an initial overview, descriptive statistics (means and standard deviations) were calculated, inter-item correlations were examined to identify potential redundancy among items and the internal consistency of the scale assessed.

The factor structure of the SRS items was evaluated using CFA with the Maximum Likelihood estimator. Model fit was assessed using a combination of absolute and relative fit indices, including the chi-square test ($\chi^2$), root mean square error of approximation (RMSEA), and standardized root mean residual (SRMR) as well as relative fit indices such as the Tucker–Lewis Index (TLI) and comparative fit index (CFI). Acceptable model fit is indicated by SRMR values less than 0.05, CFI and TLI values greater than 0.90 and RMSEA values below 0.08 (Hu & Bentler, 1999).

As a model-based framework, IRT assumed the presence of an underlying latent trait that influences both individuals` responses and the item characteristics. This unique feature allows for the simultaneous estimation of item parameters and individual ability levels (Embretson & Reise, 2000). Since the SRS uses polytomous response format, the graded response model was employed (Samejima, 1969; 1997).

One key IRT parameter is item discrimination (or slope), which indicates how well and item differentiates between individuals with low and high levels of the latent construct (Baker, 2001). Although item discrimination values can range from $+\infty$ to $-\infty$, they rarely exceed +2. Higher values suggest greater sensitivity to individual differences. In other words, items with higher discrimination parameters exhibit a more rapid change in the probability of selecting a response category as the latent construct (self-regulation) change (Hays et al., 2000). Negative discrimination values, however, may indicate a problem with the item—and may distort interpretation as they suggest that individuals with increasing levels of the latent construct are less likely to endorse more severe response options (Yang & Kao, 2014). According to Baker, 2001, discrimination values above 0.65 are acceptable, values greater than 1.34 are high; and values greater than 1.69 are considered very high (Baker, 2001).

IRT also provides item information functions (IIFs), which quantify the precision of measurement across different levels of the latent trait. Higher information indicates more reliable measurement at a particular trait level (Reise et al., 2005). Therefore, items with high information at a certain level of the latent construct offer precise estimates of person parameters at that particular level of the continuum (Baker, 2001). Summing the IFFs across all items yields the test information function (TIF), which reflects the scale`s overall precision. In this study, the SRS provided the most information for individuals whose self-regulation levels ranged from the two standard deviations below the mean to one standard deviation above. In contrast, the scale yielded less information for individuals with very high self-regulation (two standard deviations above the mean).

To examine whether the scale functioned equally well across subgroups, we evaluated differential item functioning (DIF) using IRT. DIF analysis identifies whether individuals from different groups—who are otherwise matched on the latent trait—respond differently to certain items (Harris, 1989). The presence of DIF suggests potential bias and threatens the validity of cross-group comparisons (Thissen et al., 1993).

All preliminary analyses for the SRS were conducted using IBM SPSS 26. Mplus 8.4 (Muthén & Muthén, 2017) was used for CFA, and Stata 17 (StataCorp, 2021) was employed for all IRT and DIF analyses.

### *Ethical considerations*

The study was conducted in accordance with the Helsinki guidelines (World Medical Association, 2024). Participants were informed about the purpose of the study, data handling procedures, voluntary participation, anonymity, and their right to withdraw at any time. The study was reviewed by the Ethics Committee of the Universities for Scientific Research in accordance with Icelandic regulations.

## Results

### *Preliminary analysis*

The mean scores of the eight SRS items ranged from 3.62 to 4.55 on a 5-point Likert scale from *Almost never* (1) to *Almost always* (5). When examined by discipline, nursing students reported mean scores ranging from 3.77 to 4.70, while pharmacy students` scores

ranged from 3.40 to 4.53 (see Table 1). The composite mean score for nursing students was 34.46, compared to 32.26 for pharmacy students. The overall scale score was calculated by averaging the item responses. The SRS demonstrated strong internal consistency, with a Cronbach's alpha of 0.83. Furthermore, inter-item correlations did not exceed $r$ = 0.70, suggesting no redundancy among the items (see supplementary material).

### *Confirmatory factor analysis and model fit*

Following the SRS scoring algorithm (Shell & Husman, 2008), all eight items were modeled to load onto a single latent factor. Factor loadings ranged from 0.62 to 0.72, with the item *Monitor my progress when studying* showing the highest loading (Table 1). CFA supported the unidimensional structure of the scale and indicated excellent model fit: CFI = 0.99, TLI = 0.98, RMSEA = 0.05 (90% CI = [0.04, 0.12]), and standardized root mean square residual (SRMR) = 0.02, $\chi^2$ (15, $N$ = 234) = 534.73, p = 0.001.

### *Item parameters*

The item parameters for discrimination (slope) and difficulty (threshold between response categories) are presented in Table 2. All discrimination estimates were significantly different from zero, indicating that each item effectively differentiated between participants with varying levels of self-regulation. Discrimination values ranged from 1.34 to 2.42, with *Make plans for how to study* and *Check how well I understand* showing the highest discrimination values. All items exceeded the minimum acceptable discrimination threshold of 0.65 (Baker, 2001), and most items were classified as highly or very highly discriminative (values > 1.69). The relative consistency in discrimination values across items supports the use of an unweighted composite score (Cheng et al., 2012).

Item difficulty, also referred to as category threshold, represents the trait level at which as respondent has a 50% chance of endorsing a specific response category or higher. For a 5-point Likert scale, each item has four difficulty thresholds (b1 to b4). Most item difficulty estimates fell between −2 and +2, consistent with expectations for a normally distributed latent trait such as self-regulation (Hays et al., 2000). The lowest threshold (b1) ranged from −3.59 to −2.25, with *Focus on understanding important ideas* showing the lowest value. This indicates that relatively low levels of self-regulation are required to select a response category higher than "Almost never." The thresholds increased progressively across categories, and standard errors ranged from 0.10 to 0.65, indicating adequate precision in the estimates.

To further evaluate how response categories functioned across the continuum of the latent trait, Category Characteristic Curves were analyzed (de Ayala, 2013). Ideally, CCCs should be clearly peaked and dispersed across the trait continuum. Items such as *Monitor my progress when studying, Use different ways to organize thoughts*), and *Take notes when I read class materials* displayed the most dispersed CCCs. In contrast, items like *Determine best approach for how to study, Make plans for how to study,* and *Check how well I understand* exhibited distinctly peaked curves, suggesting that these items are particularly effective in predicting specific response options based on participants` self-regulation levels. Notably, for two items—*Determine best approach for how to study* and *Check how well I understand*—the lowest response category ("Almost never") was rarely selected, indicating limited use of this category among respondents.

Figures 1 and 2 display the item information functions (IIF) and the test information function (TIF). The x-axis represents self-regulation levels, while the y-axis indicates the amount of information provided. The IFFs demonstrate that *Make plans for how to study* and *Check how well I understand* were the most informative items across the self-regulation continuum. Conversely, *Use different ways to organize thoughts* and *Take notes when I read class materials* contributed the least (For full CCC plots, see Supplementary Materials).

Overall, the SRS provided the highest level of information for individuals whose self-regulation levels were between two standard deviations below the mean and one standard deviation above the mean. This suggests that the scale performs optimally for students

**Table 1**
Descriptive Statistics and Confirmatory Factor Analysis Parameter Estimates.

| Self-Regulation Scale Items | Mean (StDev) | Estimate/ Factor Loadings | Standard Errors |
|---|---|---|---|
| Determine best approach for how to study | 4.55 (.71) | 0.67 | 0.05 |
| Monitor my progress when studying | 3.96 (1.15) | 0.72 | 0.04 |
| Make plans for how to study | 4.35 (.89) | 0.71 | 0.03 |
| Use different ways to organize thoughts | 3.62 (1.24) | 0.62 | 0.04 |
| Check how well I understand | 4.30 (.78) | 0.71 | 0.04 |
| Take notes when I read class materials | 3.76 (1.22) | 0.62 | 0.05 |
| Focus on understanding important ideas | 4.50 (.73) | 0.66 | 0.04 |
| Set goals for myself | 4.46 (.79) | 0.69 | 0.05 |

For all factor loadings, the p-value is less than 0.001. For the CFA, modification indices have been used that indicated that residuals of some variables were correlated.

**Table 2**
Self-Regulation Scale Item Difficulty and Discrimination Parameter Estimates.

| Items | $b_1$ | $b_2$ | $b_3$ | $b_4$ | a |
|---|---|---|---|---|---|
| Determine best approach for how to study | - | −2.85 (0.41) | −1.67 (0.19) | −0.58 (0.11) | 2.12 (0.34) |
| Monitor my progress when studying | −2.54 (0.32) | −1.60 (0.19) | −0.74 (0.13) | 0.13 (0.12) | 1.78 (0.24) |
| Make plans for how to study | −3.23 (0.52) | −2.12 (0.24) | −1.09 (0.13) | −0.30 (0.10) | 2.42 (0.37) |
| Use different ways to organize thoughts | −2.58 (0.37) | −1.27 (0.19) | −0.35 (0.13) | 0.66 (0.16) | 1.34 (0.19) |
| Check how well I understand | - | −2.38 (0.28) | −1.27 (0.15) | 0.04 (0.10) | 2.33 (0.34) |
| Take notes when I read class materials | −2.25 (0.28) | −1.41 (0.18) | −0.57 (0.12) | 0.47 (0.13) | 1.63 (0.23) |
| Focus on understanding important ideas | −3.59 (0.65) | −2.67 (0.37) | −1.81 (0.22) | −0.42 (0.11) | 1.93 (0.30) |
| Set goals for myself | −3.46 (0.59) | −2.36 (0.28) | −1.59 (0.18) | −0.42 (0.11) | 2.12 (0.32) |

b indicates difficulty parameter, and a indicates discrimination. Standard error estimates appear in parentheses.

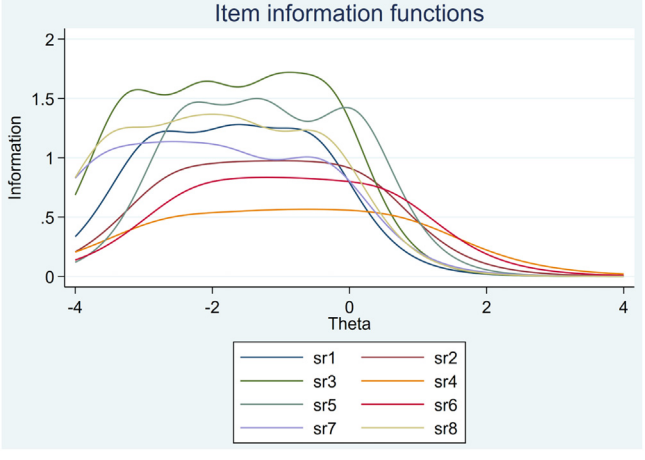


**Fig. 1.** Item information functions.

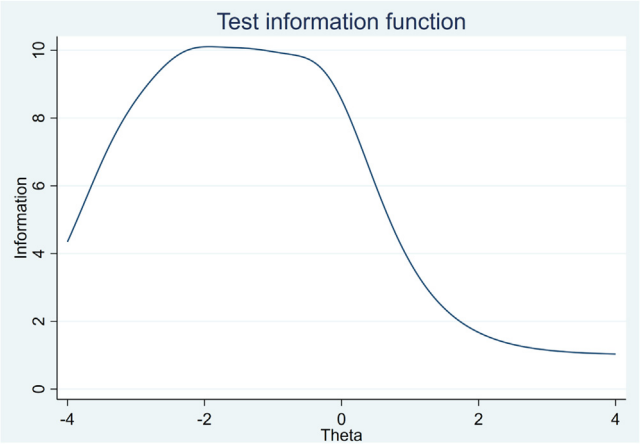


**Fig. 2.** Test information function.

with low to moderate self-regulation and less effectively for those with very high self-regulation levels.

*Differential item functioning*

DIF tests were conducted to assess whether students from different academic programs (nursing vs. pharmacy) matched on the underlying level of self-regulation, had an equal probability of endorsing specific response options for each item (Clauser & Mazor, 1998). To detect DIF, likelihood ratio tests were employed to compare the goodness-of-fit—between models that constrained item parameters to be equal across groups and those that allowed parameters to vary. The results of the DIF analyses (see Table 3) revealed that six of the eight items in the SRS exhibited significant DIF. The two items that did not show significant DIF were *Make plans for how to study* (p = 0.62) and *Take notes when I read class materials* (p = 0.48). This suggests that, controlling for their level of self-regulation, nursing and pharmacy students responded similarly to these two items. In contrast, the remaining six items showed differential response patterns between the two groups, despite equivalent latent trait levels. This discrepancy may be partially explained by sparse data in some response categories—specifically, the *"Almost never"* category for items 1 and 5, which was rarely selected.

## Conclusion

Having instruments generating reliable and valid scores of university students' self-regulation is useful, as research has demonstrated that targeted interventions for students can enhance academic success (Schunk & Zimmerman, 2023; Zimmerman, 2013). Research on how to design interventions has however been limited but is increasing (Filice et al., 2020; Russel et al., 2020). The SRS is a widely used instrument designed to assess students' perceptions of their self-regulatory behaviors in academic settings. The aim of the present study was to culturally adapt the SRS for use with Icelandic health science students and evaluate its psychometric properties. The findings indicate that the Icelandic version of the SRS is internally consistent, with a Cronbach's alpha of 0.83. This aligns with previous studies reporting similar reliability estimates ($\alpha$ = 0.81 by Shell & Husman (2008); $\alpha$ = 0.84 by Jones & Carter (2019)). Inter-item correlations revealed no issues of redundancy, and CFA supported the scale unidimensional structure with good model fit indices. Furthermore, IRT analyses showed that the scale comprises highly discriminative items, capable of effectively distinguishing between individuals with low and high levels of self-regulation. The item difficulty parameters increased gradually across response options, suggesting that the scale`s Likert format appropriately captures varying degrees of the construct. All the items contributed meaningfully to the scale, further confirming the absence of redundant content.

However, the TIF revealed that the SRS provides the greatest statistical precision for individuals with low to moderate levels of self-regulation, but less precision for those at higher levels. This limitation should be considered in future applications of the scale, particularly when evaluating high-performing student populations.

**Table 3**
Differential Item Functioning Results.

| Determine Best Approach for How to Study | Pharmacy | Nursing | Difference |
|---|---|---|---|
| a | 1.35 | 4.87 | −3.52 |
| $b_1$ | - | - | - |
| $b_2$ | −3.11 | - | −3.11 |
| $b_3$ | −1.42 | −1.76 | 0.34 |
| $b_4$ | −0.33 | −0.64 | 0.31 |
| LR test | $\chi^2(3) = 16.19$, p = 0.001 | | |
| Monitor my progress when studying | | | |
| a | 2.19 | 1.39 | 0.8 |
| $b_1$ | −2.35 | −2.97 | 0.62 |
| $b_2$ | −1.19 | −2.23 | 1.04 |
| $b_3$ | −0.50 | −1.07 | 0.57 |
| $b_4$ | 0.04 | 0.24 | −0.2 |
| LR test | $\chi^2(5) = 14.02$, p = 0.015 | | |
| Make plans for how to study | | | |
| a | 2.15 | 2.99 | −0.84 |
| $b_1$ | −2.99 | - | −2.99 |
| $b_2$ | −1.76 | −2.78 | 1.02 |
| $b_3$ | −0.50 | −1.74 | 1.24 |
| $b_4$ | 0.13 | −0.64 | 0.77 |
| LR test | $\chi^2(4) = 2.62$, p = 0.62 | | |
| Use different ways to organize thoughts | | | |
| a | 1.46 | 1.12 | 0.34 |
| $b_1$ | −1.93 | −3.90 | 1.97 |
| $b_2$ | −0.89 | −1.83 | 0.94 |
| $b_3$ | −0.17 | −0.58 | 0.41 |
| $b_4$ | 0.87 | 0.55 | 0.32 |
| LR test | $\chi^2(5) = 12.07$, p = 0.04 | | |
| Check how well I understand | | | |
| a | 2.54 | 1.76 | 0.78 |
| $b_1$ | - | - | - |
| $b_2$ | −2.05 | −3.31 | 1.26 |
| $b_3$ | −0.76 | −2.36 | 1.6 |
| $b_4$ | 0.24 | −0.22 | 0.46 |
| LR test | $\chi^2(4) = 35.02$, p = 0.0001 | | |
| Take notes when I read class materials | | | |
| a | 1.53 | 1.73 | −0.2 |
| $b_1$ | −2.30 | −2.23 | −0.07 |
| $b_2$ | −1.39 | −1.45 | 0.06 |
| $b_3$ | −0.35 | −0.75 | 0.4 |
| $b_4$ | 0.59 | 0.39 | 0.2 |
| LR test | $\chi^2(5) = 4.48$, p = 0.48 | | |
| Focus on understanding important ideas | | | |
| a | 2.56 | 2.00 | 0.56 |
| $b_1$ | −2.73 | - | −2.73 |
| $b_2$ | −2.38 | −2.42 | 0.04 |
| $b_3$ | −1.46 | −1.79 | 0.33 |
| $b_4$ | −0.59 | −0.12 | −0.47 |
| LR test | $\chi^2(4) = 51.95$, p = 0.001 | | |
| Set goals for myself | | | |
| a | 1.98 | 2.06 | −0.08 |
| $b_1$ | −3.14 | - | −3.14 |
| $b_2$ | −2.19 | −2.55 | 0.36 |
| $b_3$ | −1.37 | −1.78 | 0.41 |
| $b_4$ | −0.16 | −0.52 | 0.36 |
| LR Test | $\chi^2(4) = 17.28$, p = 0.001 | | |

To provide additional evidence of validity, the relationship between self-regulation and effort was examined. The correlation between these two constructs was statistically significant, $r$ = 0.49, supporting the concurrent validity of the SRS scores.

Another noteworthy finding is that six of the eight items exhibited DIF between students in nursing and pharmacy disciplines. This indicates that certain items may function differently across academic disciplines, potentially affecting cross-group comparisons. Further investigation of DIF across other health science disciplines is warranted to ensure the scale`s fairness and validity in diverse student populations.

The goal of this study was to adapt the SRS for Icelandic health sciences students. Participants in our sample were from two of the five departments of the School of Health Sciences, namely pharmacy and nursing, making the sample rather homogeneous. This limits the generalizability of the Icelandic version. Our future studies include collecting and testing responses from the other departments and possibly from a wider student population in Iceland. Overall, the SRS captures a critical aspect of student learning in higher education by measuring self-regulatory behaviors. Expanding this work by translating, adapting and validating the remaining three SPOCK subscales would provide researchers and practitioners with a more comprehensive set of tools for assessing self-regulation and related constructs.

## Declaration of competing interest

The authors declare that they have no competing relationships or interests that could have influenced the study.

## CRediT authorship contribution statement

**Asta B. Schram:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Formal analysis, Data curation, Conceptualization. **Sumeyra Sahbaz:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or non-for-profit sectors.

## Data availability

Data is available upon the approval of the Research Ethics Committee for Public Higher Education Institutions in Iceland.

## Acknowledgments

We express our gratitude to the students who participated in the study.

## Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.teln.2026.01.005.

## References

An, J., Oh, J., & Park, K. (2022). Self-regulated learning strategies for nursing students: A pilot randomized controlled trial. *International Journal of Environmental Research and Public Health, 19*(15), 9058. doi:10.3390/ijerph19159058.

August-Brady, M. M. (2005). The effect of a metacognitive intervention on approach to and self-regulation of learning in baccalaureate nursing students. *Journal of Nursing Education, 44*(7), 297–304. doi:10.3928/01484834-20050701-02.

de Ayala, R. J. (2013). *The theory and practice of item response theory*. Guilford Publications.

Baker, F. B. (2001). The basics of item response theory *(ED458219)*. ERIC. https://files.eric.ed.gov/fulltext/ED458219.pdf.

Beck, C. T., Bernal, H., & Froman, R. D. (2003). Methods to document semantic equivalence of a translated scale. *Research in Nursing & Health, 26*(1), 64–73. doi:10.1002/nur.10066.

Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments*. Sage.

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology, 71*, 133–149. doi:10.1348/000709901158433.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner, & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137–164). Sage.

Cheng, Y., Yuan, K.-H., & Liu, C. (2012). Comparison of reliability measures under factor analysis and item response theory. *Educational and Psychological Measurement, 72*(1), 52–67. doi:10.1177/0013164411407315.

Church, A. T., & Lonner, W. J. (1998). The cross-cultural perspective in the study of personality: Rationale and current research. *Journal of Cross-Cultural Psychology, 29*(1), 32–62. doi:10.1177/0022022198291003.

Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice, 17*, 31–44. doi:10.1111/j.1745-3992.1998.tb00619.x.

Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools, 41*, 537–550. doi:10.1002/pits.10177.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*, 6–25. doi:10.1080/00461520.2011.538645.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory*. Psychology Press.

Filice, S., Tregunno, D., Edge, D., & Egan, R. (2020). Re-imaging Clinical education: The interdependence of the self-regulated clinical teacher and nursing student. *International Journal of Nursing Education Scholarship, 17*,(1) 20190056. doi:10.1515/ijnes-2019-0056.

Flanigan, A. E., Peteranetz, M. S., Shell, D. F., & Soh, L. K. (2023). Relationship between implicit intelligence beliefs and maladaptive self-regulation of learning. *ACM Transactions on Computing Education, 23*(3), 32. doi:10.1145/3595187.

Hamman, D. (1998). Preservice teachers' value for learning-strategy instruction. *The Journal of Experimental Education, 66*(3), 209–221. doi:10.1080/00220979809604405.

Harris, D. (1989). Comparison of 1-, 2-, and 3-parameter IRT models. *Educational Measurement: Issues and Practice, 8*, 157–163. doi:10.1111/j.1745-3992.1989.tb00313.x.

Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical Care, 38*(9). doi:10.1097/00005650-200009002-00007 SupplII-28-II-42.

Hu, L.- T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. doi:10.1080/10705519909540118.

Jones, B. D., & Carter, D. (2019). Relationships between students' course perceptions, engagement, and learning. *Social Psychology of Education: An International Journal, 22*, 819–839. doi:10.1007/s11218-019-09500-x.

Lai, J.-S., Teresi, J., & Gershon, R. (2005). Procedures for the analysis of differential item functioning (DIF) for small sample sizes. *Evaluations and the Health Professions, 28*(3), 283–294. doi:10.1177/0163278705278276.

Lee, J., & Turner, J. E. (2018). Extensive knowledge integration strategies in pre-service teachers: The role of perceived instrumentality, motivation, and self-regulation. *Educational Studies, 44*(5), 505–520. doi:10.1080/03055698.2017.1382327.

Lin, K. Y., Hong, H. Y., & Chai, C. S. (2011). *Exploring college students' perceptions of classroom learning in a knowledge building environment* (pp. 591–597). Asia-Pacific Society for Computers in Education. https://repository.nie.edu.sg/entities/publication/e20ec54f-2d97-4385-b8e8-2e95b269c676. Accessed February 3, 2026.

Moustaki, I., & Knott, M. (2000). Generalized latent trait models. *Psychometrika, 65*(3), 391–411. doi:10.1007/BF02296153.

Muthén, B., & Muthén, L. (2017). Mplus. In W. J. van der Linden (Ed.), *Handbook of item response theory: Volume 3: applications* (pp. 507–518). Chapman and Hall/CRC.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385–407. doi:10.1007/s10648-004-0006-x.

Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. (1993). Predictive validity and reliability of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement, 53*, 801–813. doi:10.1177/0013164493053003024.

Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta, & G. Whitehurst (Eds.), *Annals of child development* (5th ed.). (pp. 89–129). JAI Press.

Reise, S. P., Ainsworth, A. T., & Haviland, M. G. (2005). Item response theory: Fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science, 14*(2), 95–101. doi:10.1111/j.0963-7214.2005.00342.x.

Russell, J. M., Baik, C., Ryan, A. T., & Molloy, E. (2020). Fostering self-regulated learning in higher education: Making self-regulation visible. *Active Learning in Higher Education, 23*(2), 97–113. doi:10.1177/1469787420982378.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika, 34*, 1–97. doi:10.1007/BF03372160.

Samejima, F. (1997). Graded response model. In W. J. Linden, & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85–100). Springer.

Schunk, D. H., Meece, J. L., & Pintrich, P. R. (2014). *Motivation in education: Theory, research & applications*. Pearson.

Schunk, D. H., & Zimmerman, B. J. (2023). Self-regulation in education: Retrospect and prospect. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications*. Routledge [eBook edition]Original work published 1994.

Scott, N. W., Fayers, P. M., Aaronson, N. K., Bottomley, A., de Graeff, A., Groenvold, M., Gundy, C., Koller, M., Petersen, M. A., Sprangers, M. A. G, & on behalf of the FORTC Quality of Life Cross-Cultural Meta-Analysis Group. (2009). A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scales. *Journal of Clinical Epidemiology, 62*(3), 288–295. doi:10.1016/j.jclinepi.2008.06.003.

Shell, D. F., Hazley, M. P., Soh, L. K., Ingraham, E., & Ramsey, S. (2013). Associations of student creativity, motivation, and self-regulation with learning and achievement in college computer science courses. *CSE Conference and Workshop Papers* (p. 261). http://digitalcommons.unl.edu/cseconfwork/261.

Shell, D. F., & Husman, J. (2008). Control, motivation, affect, and strategic self–regulation in the college classroom: A multidimensional phenomenon. *Journal of Educational Psychology, 100*(2), 443–459. doi:10.1037/0022-0663.100.2.443.

Shell, D. F., Husman, J., Turner, J. E., Cliffel, D. M., Nath, I., & Sweany, N. (2005). The impact of computer-supported collaborative learning communities on high school students' knowledge building, strategic learning, and perceptions of the classroom. *Journal of Educational Computing Research, 33*(3), 327–349. doi:10.2190/787L-BCBQ-20FN-FW6C.

Shell, D. F., & Soh, L.-K. (2013). Profiles of motivated self-regulation in college computer science courses: Differences in major versus required non-major courses. *Journal of Science Education and Technology, 22*, 899–913. doi:10.1007/s10956-013-9437-9.

Shell, D. F., Turner, J. E., Husman, J., Droesch-Cliffel, D. M., Nath, I., & Sweany, N. I. (1996). *Effects of collaborative, computer-supported, knowledge-building communities on high school students' knowledge building and intentional learning* (ED395570). *ERIC*. *https://files.eric.ed.gov/fulltext/ED395570.pdf*.

StataCorp. (2021). *Stata statistical software: release 17*.

Stump, G., Husman, J., Chung, W. T., & Done, A. (2009). Student beliefs about intelligence: Relationship to learning. 39th Annual Frontiers in Education Conference: Imagining and Engineering Future CSET Education. Frontiers in Education. doi:10.1109/FIE.2009.5350426.

Thissen, D., & Steinberg, L. (1993). Detection of DIF using the parameters of item response models. In H. Wainer, P. W. Holland, H. Wainer (Eds.), *Differential item functioning* (pp. 67–113). Erlbaum.

Villagran, M. M., & Lucke, J. F. (2005). Translating communication measures for use in non-English-speaking populations. *Communication Research Reports, 22*(3), 247–251. doi:10.1080/00036810500230743.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). (pp. 315–327). Macmillan.

Weinstein, C., Zimmermann, S., & Palmer, D. (1988). Assessing learning strategies: Design and development of the LASSI. In C. Weinstein, E. Goetz, P. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 279–306). Erlbaum.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). American Psychological Association.

World Medical Association. (2024). *WMA declaration of helsinki – Ethical principles for medical research involving human participants*. https://www.wma.net/policies-post/wma-declaration-of-helsinki/

Yang, F. M., & Kao, S. T. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry, 26*(3), 171–177. doi:10.3969/j.issn.1002-0829.2014.03.010.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–40). Academic Press.

Zimmerman, B. J. (2013). From cognitive modelling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135–147. doi:10.1080/00461520.2013.794676.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*, 166–183. doi:10.3102/0002831207312909.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). *Self-regulated learning and academic achievement: theoretical perspectives*. (2nd ed). Erlbaum.

Zumbo, B. D. (1999). A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores. *Directorate of human resources research and evaluation*. Department of National Defense. https://faculty.educ.ubc.ca/zumbo/DIF/handbook.pdf.