

Self-Regulated Learning among Gifted Learners in Madrasah Acceleration Programs: Developing a Context-Specific Measure

Fany Rifqoh^{1*}, Muhammad Nur Wangid², Suwarjo³, Yulia Ayriza⁴

^{1,2,3,4}Bimbingan dan Konseling, Fakultas Ilmu Pendidikan, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

*Penulis koresponden, e-mail: fanyrifqoh.2023@student.uny.ac.id

Abstract: Self-regulated learning (SRL) refers to learners' capacity to actively direct, monitor, and regulate their own learning activities. Although a number of self-regulated learning scales (SSRL) have been developed, instruments specifically designed and empirically validated for gifted students in madrasah settings remain scarce. This study aimed to examine the validity and measurement quality of an SSRL scale developed by integrating components of the Motivated Strategies for Learning Questionnaire (MSLQ) and the Self-Regulated Learning Interview Schedule (SRLIS). The study involved 360 gifted students enrolled in acceleration programs at madrasahs in Indonesia. Data were analyzed using two-level confirmatory factor analysis (CFA) conducted with R and Rasch model analysis. The CFA results indicated an adequate construct structure across the four SRL dimensions, as reflected by CFI and TLI values above 0.90 and RMSEA and SRMR values below 0.08. The Rasch analysis further identified 67 items that demonstrated acceptable model fit, with no items classified as overfit or misfit. Nevertheless, several items exhibited non-ideal threshold characteristics and therefore require further refinement. Overall, the findings indicate that the developed SSRL scale demonstrates satisfactory validity and measurement quality and can be used to assess self-regulated learning among gifted students participating in madrasah acceleration programs.

Keywords: Accelerated Programs, Self-Learning, Student Excellence, CFA, Rasch Model.

Abstrak: *Self-Regulated Learning (SRL) merupakan keterampilan yang menekankan kemampuan individu dalam mengelola dan mengendalikan proses belajarnya secara mandiri. Meskipun sudah ada beberapa Skala SRL (SSRL), instrumen yang valid dan sesuai untuk anak berbakat di madrasah masih terbatas. Penelitian ini bertujuan menguji validitas SSRL untuk siswa berkemampuan unggul yang mengikuti program percepatan di madrasah. Skala yang digunakan merupakan adaptasi dari SRL Scale Erdogan, yang menggabungkan elemen Motivated Strategies for Learning Questionnaire (MSLQ) dan Self-Regulatory Learning Interview Schedule (SRLIS). Sampel terdiri dari 360 siswa berbakat di Indonesia. Analisis data menggunakan confirmatory factor analysis tingkat dua dengan bantuan program R dan rasch model. Hasil uji CFA menunjukkan validitas yang baik untuk keempat dimensi SRL, dengan nilai CFI dan TLI > 0,90 serta RMSEA dan SRMR < 0,08. Selain itu, analisis Rasch menghasilkan 67 item dengan semua butir menunjukkan kesesuaian model (tidak overfit atau misfit). Namun, beberapa butir dengan threshold tidak ideal disarankan untuk direvisi. Secara keseluruhan, SSRL terbukti valid untuk mengukur SRL siswa berbakat yang mengikuti program percepatan di madrasah*

Kata kunci: Program Akselerasi, Pembelajaran Mandiri, Kemampuan Unggul Siswa, CFA, Model Rasch.

PENDAHULUAN

Self-Regulated Learning (SRL) is a learning approach that emphasizes the ability of individuals to self-regulate and control their learning process (Wong et al., 2019). In the scope of learning in Madrasah, this will relate to the competence of students excel with ability (Ridgley, DaVia Rubenstein, & Callan, 2022, Butler, 2023 Siekanska, Wilson, Blecharz, & Young,

2023; Zimmerman, 2010). State that academically successful students are those who have adequate self-confidence, perseverance, and learning skills. They view learning as a structured process and take full responsibility in achieving their academic goals (Schraw, Kauffman, & Lehman, 2006). From this context, the role of SRL becomes important for gifted students (Ridgley, DaVia Rubenstein, & Callan, 2022). SRL is particularly important when gifted students face complex learning-related tasks and simultaneously, the learning environment provides high autonomy through SRL (Hertel, et al., 2024). One of the madrasah education programs to serve gifted children is acceleration services. In achieving this acceleration, gifted children need the right concept of self-management in learning, one of which is SRL (Zimmerman, Bonnner, Kovach, 1996; Wangid, 2004; Wahab, 2010; Nilson, 2013; Hertel, 2024). Self-Regulated Learning (SRL) ability is one of the key indicators of academic success, especially for students who participate in accelerated programs. Therefore, understanding the level of SRL ability of learners is not only important, but crucial as a basis for preparing appropriate educational intervention strategies, so an appropriate measurement tool is needed, namely the SRL scale for superior ability learners who take part in accelerated classes in madrasah. However, the right instrument has not yet been found.

There is a relevant and previous research conducted by (Zimmerman, 2023) the research has developed a SRL scale, but the resulting instrument is in the form of an interview guide and the results of the study have not explained the validity of the product, so it is considered less effective for measuring large numbers of students. Pintrich, et al's research has developed SRL instruments, in general, the instruments developed have been validated constructs and are widely used to assess student motivation and learning strategies by self-report. (Pintrich, Smith, Garcia, & McKeachie, 1991). MSLQ has major weaknesses in the form of dependence on self-report which is prone to bias, does not capture the dynamics of the learning process, and needs cultural adaptation for cross-context validity. Furthermore, (Erdogan & Senemoglu, 2016) the development and implementation of the instrument is still limited for use in Turkey, and access is still limited, so flexibility is needed regarding access to the instruments developed.

According to Hakim (2024), difficulties in meeting academic targets or experiencing failure can trigger stress and contribute to academic anxiety. This issue is also relevant to gifted students, who may experience particular academic pressures despite their superior abilities. Interviews with school counselors indicated that some high-ability students still required remedial support, while students participating in accelerated programs more frequently expressed concerns related to anxiety than to their academic effort. These findings suggest that high academic ability does not necessarily protect students from academic anxiety and that their emotional experiences need to be identified alongside their academic performance. In this context, an appropriate and sensitive instrument is important for identifying the specific dimensions of academic anxiety experienced by gifted students. One relevant dimension is reflected in the anxiety scale developed by Erdogan et al., which has also been examined in relation to academic anxiety among students (Himawan & Nurgiyantoro, 2022). Therefore, identifying the dimensions of anxiety through an appropriate instrument is important for providing a more comprehensive understanding of the psychological experiences of gifted students in academically demanding programs.

Finally, Larasati et al. (2020) reported that 11 instruments had been developed to measure self-regulated learning. However, their review indicated that the existing instruments provided limited depth in measuring self-regulated learning among gifted students. This gap emphasizes the need for an appropriate and culturally sensitive approach in Indonesia to create an accurate and credible scale that can reflect the superior abilities of students participating in acceleration programs in madrasahs. The instrument must be validated and ready for use, and it must be tailored to children with superior abilities in Indonesia. Overall, the researchers aim to develop an appropriate and accurate scale to measure the SRL of gifted students participating in acceleration programs through the development of an SRL scale. The validated scale is still very limited in scope, as supported by several opinions stating that the references for developing the instrument are still very limited (Qin et al., 2024; Carter et al., 2020; Ridgley et al., 2022). Irrelevant or poorly standardized instruments have the potential to produce biased and misleading data in decision-making. Therefore, the development of an SRL scale specifically designed for gifted in madrasah, as well as validated with

a strong methodological approach, is a scientific and practical urgency that needs to be answered immediately. In line with its purpose, this study will try to develop an SRL scale instrument and test the validity of the instrument to achieve the optimization of superior ability children in madrasah.

This research presents novelty through the development of a Self-Regulated Learning (SRL) scale specifically aimed at superior ability students in accelerated programs in madrasah. This scale is a contextual adaptation of Erdogan's SSRL (2016). This research presents novelty through the development of a Self-Regulated Learning (SRL) scale specifically designed for high-ability students participating in accelerated programs in madrasahs. The development of this scale is based on the consideration that existing SRL instruments may not fully reflect the academic demands and learning context experienced by high-ability students in accelerated Islamic education settings. Students in accelerated programs are expected to manage their learning more independently, regulate their motivation and learning strategies, and adapt to a relatively intensive academic environment. At the same time, learning in madrasahs is situated within a particular educational, cultural, and religious context that may shape students learning experiences and self-regulation. Therefore, the scale requires contextual adaptation rather than direct adoption from an instrument developed in a different educational setting. This study adapts Erdogan's SSRL scale (2016) by considering the language, cultural context, learning characteristics, and specific educational context of madrasahs. The adaptation is intended to improve the contextual relevance of the instrument and enable a more accurate assessment of SRL among high-ability students in accelerated programs.

The main novelty lies in the use of dual validation methods, namely Second-Order Confirmatory Factor Analysis and Rasch Model Analysis, which are rarely used simultaneously in similar studies. The resulting instrument is valid, reliable, and culturally and pedagogically relevant, and can be used for assessment purposes as well as the development of more effective self-directed learning interventions for gifted students in madrasah.

From this description, to conduct construct validity of the developed product, the use of appropriate statistical software is very important, to ensure accurate results (Haerudin & Nur, 2020). Studio was chosen to support the required multivariate statistical analysis, namely confirmatory factor analysis (CFA), and using Rasch model analysis, which is considered a more modern development of the SRL Scale adapted from SSRL developed by Erdogan. As part of the validation process, the scale will be used on a sample of high-ability students enrolled in an accelerated program in Indonesia.

METODE

Participants were recruited from madrasahs offering accelerated programs at the MTs level in Yogyakarta and East Java. The study involved gifted learners enrolled in these programs. In total, data were collected from 361 respondents across nine madrasahs located in seven districts/cities. In Yogyakarta, the participating institutions were MTs N 1 Yogyakarta in Yogyakarta City; MTs N 6 Sleman, MTs N 4 Sleman, and MTs Sunan Pandanaran in Sleman Regency; MTs N 4 Bantul in Bantul Regency; MTs 1 Kulonprogo in Kulonprogo Regency; and MTs N 4 Gunungkidul in Gunungkidul Regency. In East Java, participants were recruited from MTs N 3 Malang in Malang City and MTs Unggulan PP Amanatul Ummah Pacet in Mojokerto Regency. The research permission procedure differed slightly between the two provinces. In Yogyakarta, the permission request was initially submitted to the Regional Office of the Ministry of Religious Affairs. After provincial approval was obtained, the researchers provided the official permission letter to the respective madrasahs. In East Java, the research permission process was conducted directly with the participating madrasahs. The questionnaire was administered using two options: participants could complete the scale directly in printed form or respond electronically through Google Forms. This flexible administration procedure was used to accommodate the conditions of each participating madrasah. The sample size was also considered in relation to the requirements for confirmatory factor analysis (CFA). Following Jöreskog (1968), the recommended sample size is at least five times the number of parameters when maximum likelihood estimation is employed, or ten times the number of parameters when another estimation

method is used. Based on this criterion, the number of participants in the present study was considered sufficient for the CFA procedure.

The instrument employed in this study was a self-regulated learning scale consisting of 67 items. Responses were recorded using a five-point Likert scale ranging from 1 (*Never*) to 5 (*Always*). The development of the scale was informed by the Self-Regulation in Learning Scale (SSRL) proposed by Erdogan (2016), which integrates components of the Motivated Strategies for Learning Questionnaire (MSLQ) developed by Pintrich and De Groot and the Self-Regulated Learning Interview Schedule (SRLIS) developed by Zimmerman and Martinez-Pons. Erdogan's SSRL conceptualizes self-regulated learning as a multidimensional construct comprising 17 dimensions and 67 items that address both learning-related skills and motivational processes. The cognitive component covers three stages of learning: before learning, during learning, and after learning. The before-learning stage includes aspects such as environmental structuring and planning, whereas the during-learning stage involves organizing information, seeking assistance, and self-monitoring. The after-learning stage encompasses self-evaluation and reflections on the consequences of success or failure. The motivational component includes task value, self-efficacy, anxiety, attribution of failure, and goal orientation. Permission to use the original Erdogan instrument was formally requested and obtained from the original author. In response to the request, Erdogan provided three versions of the instrument: the initial version, the validated version, and a version classified according to item polarity. The instrument was subsequently contextualized for gifted learners in Indonesia through a systematic adaptation process involving translation, contextual adjustment, and expert judgment. The resulting items were reviewed to ensure that their wording and content were appropriate for the characteristics and learning context of gifted students participating in madrasah acceleration programs.

Steps	Original Item	Adjustments for Gifted-Accelerated Students
A. Before Study	4. Saya kesulitan menggunakan waktu secara efektif.	4. Saya kesulitan membagi waktu belajar karena materi pada kelas percepatan menuntut penguasaan lebih cepat.
	5. Saya berkonsultasi dengan guru tentang mata pelajaran yang penting untuk jangka panjang.	5. Saya berkonsultasi dengan guru mengenai mata pelajaran yang penting untuk jangka panjang/ selama mengikuti kelas percepatan pada jenjang SMP/MTs.
	9. Di awal semester saya merencanakan apa yang harus dilakukan dan kapan harus melakukannya.	9. Saya merancang strategi belajar semesteran agar dapat menyelesaikan materi lebih cepat dan mendalam sesuai ritme pribadi.
B. During Study	20. Saya hanya membaca buku referensi yang diberikan oleh guru	Saya hanya membaca buku UKBM (Unit Kegiatan Belajar Mandiri) sebagai referensi yang diberikan oleh guru.
C. Motivation	46. Saya yakin dapat memahami materi pelajaran yang paling sulit sekalipun.	46. Saya yakin dapat memahami materi kelas percepatan yang menantang dengan usaha dan strategi belajar yang tepat.

Table 1. Instrument SRL
Sources: Erdogan (2016) & Personal data (2025)

After developing the 65 items, the instrument was reviewed by three experts-one linguist and two instrument experts-who evaluated the relevance of each item to measure self-regulated learning in high-ability students using a four-point scale. They also provided qualitative feedback for improvement. Quantitative analysis using Gregory's formula was conducted to assess content

validity, while qualitative feedback helped to improve item clarity and grammar. After revision, the final instrument consisted of 67 items for the pilot study.

The pilot study used both offline (paper-based) and online (Google Form) methods. Offline data collection involved coordination with schools and classroom-based distribution, where students completed the questionnaire within 20-25 minutes. Online data collection took place over three weeks in November 2024 using a dedicated link. Participants could only respond once and were informed of confidentiality and voluntary participation to ensure ethical standards were upheld.

There are two stages involved in analyzing the research data. First, the analysis was conducted with descriptive statistics to determine the feasibility and characteristics of the data to be analyzed. Second, factor analysis was conducted to prove the construct validity of self-regulated learning. This study used Second-Order Confirmatory Factor Analysis (CFA) and Rasch Model Analysis to test the hypothesis. Through second-order CFA, this research will evaluate the goodness of fit of the conceptual model and the loading factor for each manifest variable. Goodness of fit of the statistical model is used to evaluate how well the theoretical model fits the empirical data. Good model fit is indicated by several parameters: the ratio of chi-square to degree of freedom is less than 3 ($\chi^2/df < 3$) (Kline, 2016), comparative fit index is more or equal to 0.90 ($CFI \geq 0.90$) (Hu & Bentler, 1999), Tucker-Lewis Index is more or equal to 0.90 ($TLI \geq 0.90$) (Hu square error of 0.90 ($TLI \geq 0.90$) (Hu & Bentler, 1999), and root mean square residual less than or equal to 0.08 ($SRMR \leq 0.08$) (Hu & Bentler, 1999). To evaluate the standardized factor loading of each manifest variable, we used the criteria $\beta > 0.3$ (Gustafson et al, 1988 & de Oliveira et al, 2006), $p < 0.05$ (Gore et al., 2017 in Khampirat). Meanwhile, further data analysis was conducted using the Rasch model with Winsteps software version 3.73 to analyze Unidimensionality, Local Independence, Item Measure, Reliability (Person & Item Separation Reliability), Wright Map, and Information Function Test.

Following the initial development of 65 items, the instrument underwent expert evaluation involving three reviewers: one expert in linguistics and two experts in educational measurement. The experts assessed the relevance of each item to the measurement of self-regulated learning among high-ability students using a four-point rating scale. In addition to the quantitative ratings, the experts provided qualitative comments concerning item wording, clarity, and grammatical accuracy. Content validity was examined quantitatively using Gregory's formula, while the qualitative recommendations were considered in revising the wording and formulation of the items. Following the expert review and revision process, the instrument contained 67 items and was subsequently administered in the pilot study.

Data collection was conducted through both paper-based and online administration. For the offline procedure, the researchers coordinated with the participating madrasahs and distributed the questionnaires during classroom sessions. Students were given approximately 20–25 minutes to complete the instrument. The online administration was conducted over a three-week period in November 2024 using a designated Google Forms link. The online system was configured to allow each participant to submit a response only once. Before completing the questionnaire, participants were informed about the purpose of the study, the voluntary nature of their participation, and the confidentiality of the information provided. These procedures were implemented to maintain the ethical standards of the research.

The data analysis consisted of two main stages. The first stage involved descriptive statistical analysis to examine the general characteristics and preliminary suitability of the collected data for further analysis. The second stage focused on testing the construct validity and measurement properties of the self-regulated learning scale using second-order confirmatory factor analysis (CFA) and the Rasch model.

Second-order CFA was employed to examine the extent to which the empirical data supported the proposed conceptual structure of self-regulated learning. The analysis evaluated both the overall model fit and the factor loadings of the observed variables. Model fit was assessed using several commonly reported indices, including the chi-square to degrees-of-freedom ratio ($\chi^2/df < 3$) (Kline, 2016), comparative fit index ($CFI \geq 0.90$), Tucker–Lewis index ($TLI \geq 0.90$), root mean square error

of approximation ($RMSEA < 0.06-0.08$), and standardized root mean square residual ($SRMR \leq 0.08$) (Hu & Bentler, 1999; Schreiber et al., 2006). Standardized factor loadings were also examined, with coefficients greater than 0.30 considered indicative of an acceptable contribution of the observed variable to its corresponding construct (Gustafson et al., 1988; de Oliveira et al., 2006). Statistical significance was evaluated at the 0.05 level.

Following the CFA, Rasch analysis was conducted using Winsteps version 3.73 to obtain further evidence regarding the measurement quality of the scale. The Rasch analysis examined several measurement properties, including unidimensionality, local independence, item measures, person and item separation reliability, the Wright map, and the test information function. These analyses were used to evaluate whether the items functioned consistently within the measurement model and to identify items requiring further examination or refinement.

RESULT AND DISCUSSION

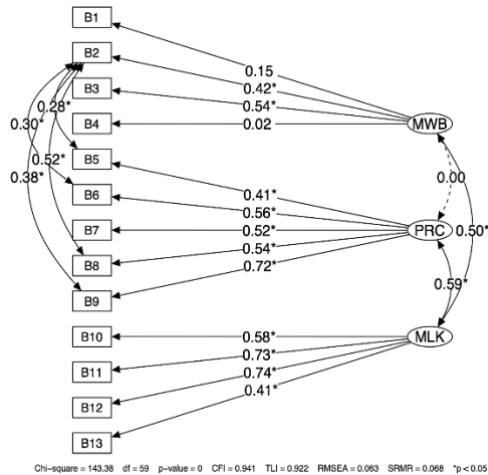
Descriptive Statistics

The self-regulated learning scale consists of two subscales, namely independent learning skills and motivation. Self-regulated learning skills (cognitive factors) are grouped into three main dimensions, namely before study (organizing/managing learning time, planning, and managing the environment), during study (organizing and transforming, searching for appropriate information, searching for easily accessible information, seeking help from peers, teachers, or adults, self-monitoring and practicing and memorizing), and after study (self-evaluation, self-consequences after success, and self-consequences after failure). Motivational factors include task value, self-efficacy, anxiety, failure attribution, and goal orientation. There are a total of 17 dimensions consisting of 67 statement items.

The self-regulated learning (SRL) scale comprises two overarching components: independent learning skills and motivation. The learning skills component represents the cognitive aspects of self-regulated learning and is organized into three sequential phases: before study, during study, and after study. The before-study phase covers the regulation of learning activities through organizing and managing study time, planning, and managing the learning environment. The during-study phase includes organizing and transforming information, locating appropriate information, accessing readily available information, seeking assistance from peers, teachers, or other adults, self-monitoring, and practicing or memorizing. The after-study phase focuses on self-evaluation and reflections on the consequences of success and failure. The second component concerns motivational processes, which include task value, self-efficacy, anxiety, attribution of failure, and goal orientation. Overall, the scale encompasses 17 specific dimensions represented by 67 statement items, covering both the cognitive regulation of learning across different phases and the motivational processes underlying students' self-regulated learning.

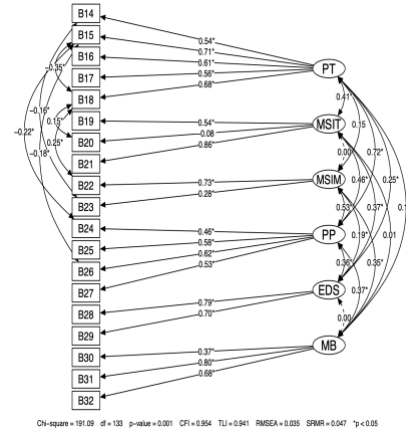
To prove the validity of the self-regulated learning scale construction, we conducted two types of analysis, namely CFA, the first being second-order confirmatory factor analysis. To evaluate whether the conceptual model fits the empirical data, we used several fit statistics. All fit indices of the conceptual model (χ^2 , RMSEA, CFI, TLI, and SRMR) meet all criteria of goodness of model fit. Thus, it can be concluded that the constructs of the self-regulated learning scale measured through two subscales and 17 dimensions fit the empirical data.

The results of the CFA analysis for the before study main dimension show that the measurement model meets the *model fit* criteria ($CFI = 0.941$, $TLI = 0.941$, $RMSEA = 0.063$, $SRMR = 0.068$), so it fits the empirical data. In the MWB dimensions, only two of the four items (B2 and B3) have a loading factor > 0.3 , while B1 and B4 can be considered for deletion. In the PRC dimensions, all items (5 items) have a loading factor > 0.3 , with B9 as the highest contribution and B5 the lowest. Similarly, in the MLK dimensions, all four items had a loading factor > 0.3 , with the highest contribution in B12 and the lowest in B13. Overall, these results prove that the before study main dimension measurement scale is valid and supported by empirical data.



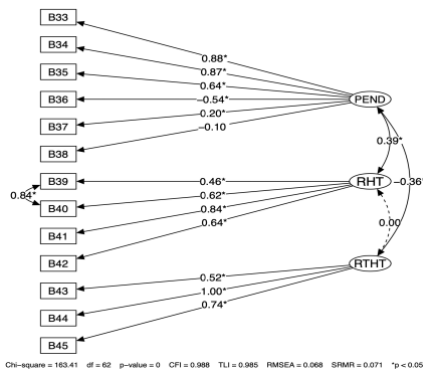
Sources: Personal data (2025)

Figure 1. CFA Report Before Study



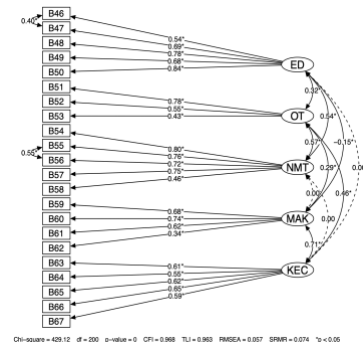
Sources: Personal data (2025)

Figure 2. CFA Report During Study



Sources: Personal data (2025)

Figure 3. After Study



Sources: Personal data (2025)

Figure 4. Motivation

Main Dimensions During Study

The analysis results show that the measurement model of the during study main dimension meets the model fit criteria (CFI = 0.954, TLI = 0.941, RMSEA = 0.035, SRMR = 0.047), so it fits the empirical data. In the PT, PP, EDS, and MB dimensions, all items had a loading factor > 0.3, indicating significant contributions. In the MSIT dimensions, only two items (B19 and B21) are significant, while B20 can be considered for deletion. In the MSIM dimensions, only B22 is > 0.3, but B23 remains statistically significant ($p < 0.05$) so it can still be retained. Overall, the during study main dimension scale is valid for use because it is supported by empirical data.

Fit Index	Obtained Value	Recommended Criterion	Interpretation
CFI	0.954	≥ 0.90	Good fit
TLI	0.941	≥ 0.90	Good fit
RMSEA	0.035	≤ 0.08	Good fit
SRMR	0.047	≤ 0.08	Good fit

Table 2. Model Fit Indices of the Measurement Model

Main Dimension After Study

The results of the analysis show that the measurement model for the After Study main dimension demonstrated an acceptable level of fit with the empirical data, based on the overall pattern of fit indices (CFI = 0.988, TLI = 0.850, RMSEA = 0.068, SRMR = 0.0071). Although the TLI value (0.850) was below the commonly recommended threshold of 0.90, the high CFI value (0.988), RMSEA value of 0.068, and very low SRMR value of 0.0071 provided additional evidence of adequate model-data correspondence. Therefore, the model should not be characterized as demonstrating a uniformly good fit across all indices; rather, the combination of the fit indices suggests an acceptable overall fit that warrants further consideration of the measurement structure.

For the PEND dimension, three of the six items (B33, B34, and B35) had factor loadings above 0.30 and contributed significantly to the construct. In the RHT dimension, all four items (B39–B42) contributed significantly, with B41 showing the highest loading and B39 the lowest. Similarly, all items in the RTH dimension contributed significantly, with B44 showing the highest loading and B43 the lowest. Overall, the findings provide empirical support for the measurement structure of the After Study main dimension, although the TLI value indicates that the model fit is not uniformly strong across all fit indices.

Main Dimensions of Motivation

The results of the analysis show that the measurement model of the motivation main dimension meets the model fit criteria (CFI = 0.968, TLI = 0.963, RMSEA = 0.057, SRMR = 0.074), which indicates a good fit with the empirical data. In all dimensions (ED, OT, NMT, MAK, and KEC), all items have a loading factor > 0.3, so they all contribute significantly. The highest and lowest contributions of each dimension are as follows: ED: Highest B50, lowest B46, OT: Highest B51, lowest B53, NMT: Highest B54, lowest B58, MAK: Highest B60, lowest B62, KEC: Highest B66, lowest B64

Validity Based on Rasch Model

To further validate the construction of the self-regulated learning scale, we compared CFA analysis with Rasch model analysis, measured through two subscales and 17 dimensions, with the empirical data. The first assumption examined in the Rasch model is unidimensionality, which is used to determine whether the items in an instrument predominantly measure a single underlying construct. Following the criterion proposed by Holster and Lake (2016), a raw variance explained by the measure of greater than 40% was considered evidence of acceptable unidimensionality. The results showed that the Main dimension of Before Study and the Main dimension of Motivation met this criterion, with raw variance explained by the measure of 42.0% and 51.3%, respectively. In contrast, the Main dimension of During Study and the Main dimension of After Study yielded raw variance explained values of 37.5% and 30.2%, respectively, which were below the 40% criterion.

These findings indicate that the evidence for unidimensionality was not consistent across all main dimensions. The lower variance explained in the During Study and After Study dimensions suggests that the items in these dimensions may not sufficiently represent a single dominant latent construct. This result may be related to the multidimensional nature of self-regulated learning, in which different regulatory processes can occur across different phases of learning. Nevertheless, the results should not be interpreted as demonstrating satisfactory unidimensionality for these two dimensions. Instead, the findings indicate that the dimensional structure of the During Study and After Study dimensions requires further examination, particularly through analysis of item content, local item dependence, residual contrasts, and the conceptual coherence of the indicators. Therefore, the evidence of unidimensionality should be interpreted separately for each main dimension rather than generalized across the entire scale.

Item measure

Based on the Rasch analysis using Winsteps, the item difficulty in the Before Study–Main dimension ranged from –1.11 to +1.53 logits, with a mean of 0.00 logits, indicating a centered distribution of item difficulty. Most items conformed to the Rasch model; however, Item 9 showed

misfit characteristics and a negative correlation with the measured construct. Therefore, Item 9 was identified for further review rather than being immediately removed from the instrument.

In the During Study–Main dimension, item difficulty ranged around a mean of 0.00 logits with an SD of 0.63 logits, indicating a centered scale with adequate variation in item difficulty. Item 9 again showed misfit, with an Infit MNSQ of 1.57, Outfit MNSQ of 1.70, and correlation of -0.16 . These results indicate that the item requires further examination, particularly with respect to its wording and conceptual alignment with the intended construct. The average item correlation was 0.43, while the Exact Match rate was 37.8%, which was close to the expected value of 36.3%.

For the After Study–Main dimension, item difficulty ranged from -0.44 to $+0.86$ logits (mean = 0.00; SD = 0.44), suggesting that the items were centered around participants with average ability. However, Item 7 showed misfit and was therefore identified for further evaluation. Similarly, in the Motivation–Main dimension, item difficulty ranged from -1.51 to $+2.03$ logits (SD = 0.90), with high item separation (14.70) and reliability (1.00). Most items conformed to the model, although Item 6 showed misfit (Infit MNSQ = 1.45; Outfit MNSQ = 1.74; high ZSTD) and was therefore flagged for review.

The misfitting items were not automatically interpreted as evidence that the entire instrument was invalid. Instead, they were identified as items requiring substantive and statistical examination, including a review of item wording, conceptual relevance, and potential ambiguity. Because the present analysis identifies these items as requiring refinement, the instrument should be regarded as demonstrating generally acceptable Rasch measurement characteristics, with specific items requiring revision and subsequent retesting. Therefore, the findings do not support an unrestricted claim that the instrument is fully valid and reliable in its current form.

Local independen

In local independence, it is an assumption that must be met, which requires that each item developed is not significantly related to the other items. A residual correlation value above 0.30 indicates local item independence (LID), which is supported by K. B. Christensen et al. (2017), Andrich (2018), and Yen (1984), who state that residual correlation values above 0.30 above the average can signal local dependence between items. LID occurs due to the influence of other responses. This finding can affect the accuracy of the SRL scale. The LID items can be seen in the table 2 below:

Sub Dimension	Residual Correlin	Entry Number	Entry Number
After Study	0.73	39	40
	0.63	44	45
	0.57	33	34
	0.31	41	42
Motivation	0.64	55	56
	0.47	49	50
	0.45	46	47
	0.42	54	55
	0.42	54	57
	0.40	56	57
	0.40	48	50
	0.39	54	56
	0.39	55	57
Before study	0.33	11	12
During Study	0.48	31	32
	0.43	28	29
	0.32	22	23
	0.31	19	21
	0.25	30	31

0.25	15	16
------	----	----

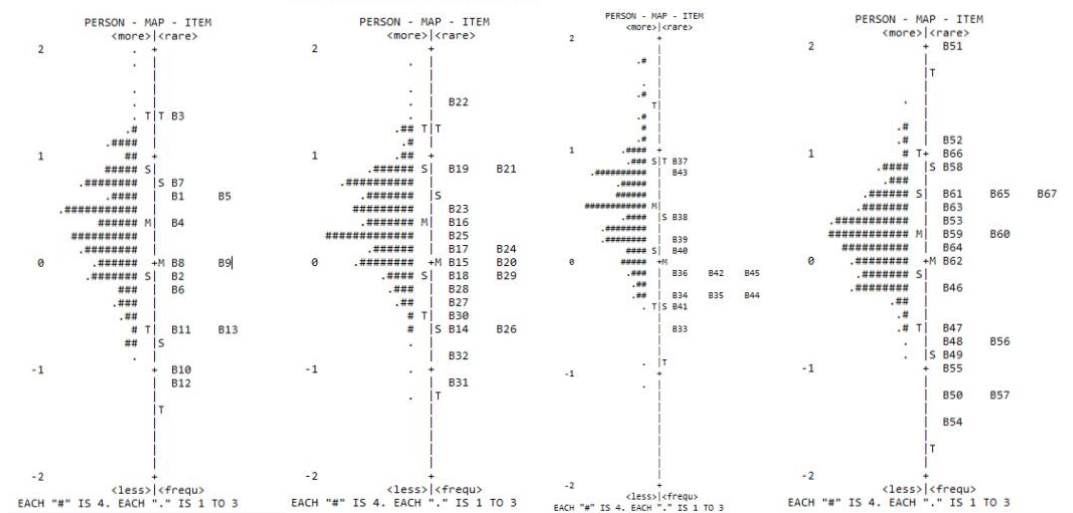
Table 2. Local Item Independen
Sources: Personal data (2025)

Reliability (Person dan Item Separation Reliability)

The results of Rasch analysis on the main dimensions before learning showed very high item reliability (0.99) with an item separation value of 11.87, indicating that the items were of very good quality in distinguishing levels of difficulty. However, respondent reliability of 0.62 and person separation of 1.29 indicate that the instrument's ability to distinguish participants' ability levels is still limited. A similar pattern emerged in the dimension during learning, with item reliability of 0.99 and item separation of 10.94 indicating very good item quality, but person reliability of 0.68 and separation of 1.46 only able to distinguish participants into two large ability groups. In the post-learning dimension, item reliability remained high (0.99) with item separation of 8.13, but participant reliability was only 0.53 and person separation 1.07, indicating that the instrument's sensitivity in distinguishing participants abilities is still low. Meanwhile, in the motivation dimension, item reliability reached 1.00 and item separation 14.70, indicating a very strong item structure; however, person reliability of 0.67 and separation of 1.42 still showed limitations in distinguishing participants abilities in more detail. Thus, although the instrument has a very good and stable item structure, improvements are still needed in the distribution and variation of items to make it more sensitive in measuring differences in participants ability levels in greater depth.

Wright map

Key dimensions after learning: The Wright Map shows that participants abilities are spread from -2 to +2 logits, with the majority falling within the range of 0 to +1 logits. The distribution of items is fairly even, allowing for effective measurement. Item B3 is classified as very difficult and its effectiveness needs to be reviewed. Items B10 and B12 reach participants with low abilities, although adding more items is still recommended. Overall, there is a good balance between participants abilities and item difficulty levels. Primary dimensions during learning: Most participants are in the 0 to +1 logit range, consistent with the concentration of items in that range, indicating a good fit between item difficulty levels and participants abilities. Item B22 is beneficial for high-ability participants, while B31 and B32 reach low-ability participants. This distribution supports optimal measurement, though adding items remains an option. Third dimension after learning: Most participants are in the logit range of 0 to +1, reflecting moderate to slightly high ability. Item distribution aligns with this pattern, enabling effective participant differentiation. The most difficult items (B37, B43) reach high-ability participants, while the easiest items (B33) reach low-ability participants, although their numbers are limited. Overall, the item distribution is fairly balanced and supports good measurement functionality. Main dimension of motivation: Participants abilities are concentrated in the logit range from 0 to +1, with few at very high or low levels. Most items fall within an appropriate range, enabling optimal differentiation. The most difficult items (B51, B52) and easiest items (B54, B50, B57) reach participants at both ends of the ability spectrum. This distribution reflects a balanced, valid, and reliable instrument.



Sources: Personal data (2025)

Figure 5. Wright Map (Before studi, during study, after study, motivational)

Based on the Test Information Function (TIF) graph from the Rasch analysis results, the four dimensions (before learning, during learning, after learning, and motivation) show a symmetrical normal curve pattern with the highest information peak around 0 logit, indicating that the highest measurement accuracy occurs in participants with moderate abilities. In the pre-learning dimension, information values reach 13–14; during learning, around 19–20; after learning, approaching 18–19; and motivation reaches 22–23, indicating low measurement error within that range. However, at the extreme ends of ability (very low or very high), information values drop sharply, leading to reduced measurement accuracy. Therefore, this instrument is highly effective for measuring participants with average ability, but it is recommended to add easier and more difficult items to expand the range of abilities that can be accurately measured and reduce measurement errors at the extreme ends. Based on Figure 3, the information function test curve shows that item separation has a high value of 17. Thus, the 13 questions in the After Study dimension given to 360 students indicate that they are suitable for determining the SRL scale.

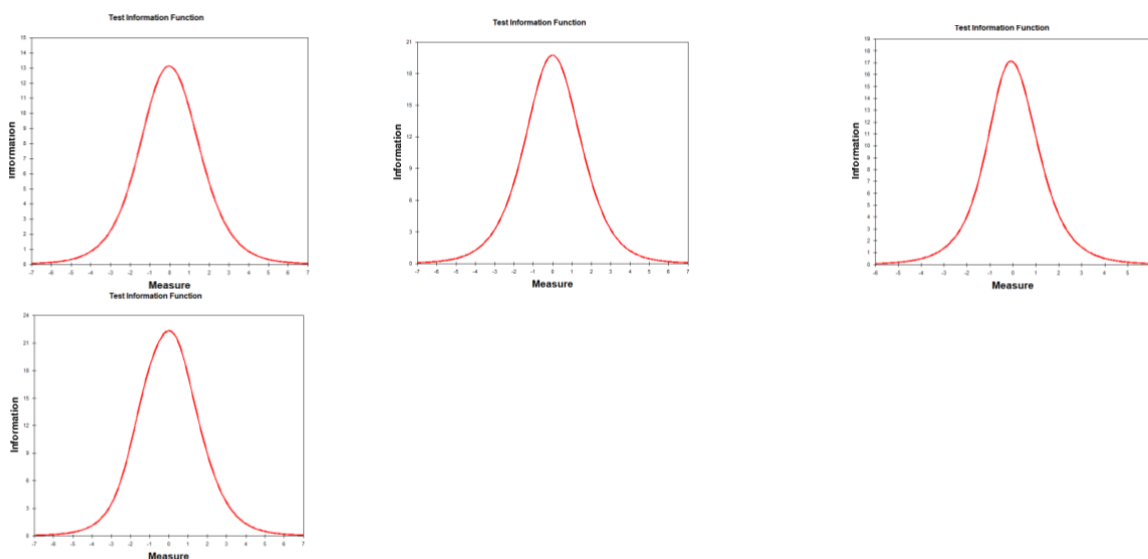


Figure 6. TIF (Before Studi, During Study, After Studi, and Motivation)

Nevertheless, this study has several limitations that should be considered when interpreting the findings. First, the development and validation of the SRL scale were conducted specifically with high-ability students participating in accelerated programs in madrasahs; therefore, the findings may not be directly generalizable to students in regular programs, different educational levels, or other educational contexts. Second, although the measurement model demonstrated acceptable overall fit and the four main factors were supported, several items showed misfit characteristics and require further substantive and statistical examination. These items should not be considered evidence of overall instrument invalidity, but they indicate the need for item refinement and subsequent retesting. Third, the contextual adaptation of the scale was based on the language, learning characteristics, and educational context of madrasahs, which means that further validation may be necessary when the instrument is applied in different cultural or institutional settings. Future studies should therefore involve more diverse samples and conduct further item refinement and validation, including cross-context testing, to strengthen the evidence for the validity and applicability of the SRL scale.

CONCLUSION

CFA results recommended revisions to specific items (B1, B4, B20, B36, B38). The sample consisted of 360 gifted students in Indonesia. Data analysis used two-level confirmatory factor analysis with the help of the R program. Test results showed good validity for all four SRL dimensions, with CFI and TLI values >0.90 and RMSEA and SRMR <0.08 . Based on Rasch analysis, several items need to be revised because they show misfit or local item dependence (LID) that can interfere with the validity and accuracy of the SRL instrument. Items with significant misfit are Item 6 (motivation), Item 7 (after learning), and Item 9 (before and during learning), because they have inappropriate Infit, Outfit, and correlation values. Additionally, there are item pairs with residual correlations >0.30 , namely: 33–34, 39–40, 41–42, 44–45 (after learning); 46–47, 48–50, 49–50, 54–55, 54–56, 55–56, 55–57, 56–57 (motivation); 11–12 (before learning); and 15–16, 19–21, 22–23, 28–29, 30–31, 31–32 (during learning). These items need to be reviewed or restructured to avoid redundancy and improve construct validity and measurement accuracy.

After being adapted to Indonesian culture and the characteristics of accelerated students, The scale comprises two subscales: independent learning skills and motivation. Independent learning skills are divided into three phases: before study, during study, after study and main dimension Motivational⁶⁷ valid items were obtained based on CFA and Rasch analysis, with some items revised. This instrument is intended for high-ability junior high school students in accelerated programs and can be completed in 20–25 minutes, either online or offline. It is suitable for large-scale research, but it is recommended to be supplemented with interviews or observations for a deeper understanding. This study used a combination of Confirmatory Factor Analysis (CFA) and Rasch Model to analyze the SRL scale for high-ability students, with the aim of improving the validity and accuracy of the scale. The suggestion for further research is that after item improvement, retesting can be conducted.

REFERENSI

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Wadsworth Publishing Company.

- Andrich, D. (2018). Controlling response dependence in the measurement of change using the Rasch model. *Statistical Methods in Medical Research*, 27(12), 3709–3725. <https://doi.org/10.1177/0962280217710834>
- Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31(6), 445–457. [https://doi.org/10.1016/S0883-0355\(99\)00014-2](https://doi.org/10.1016/S0883-0355(99)00014-2)
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamentals of measurement in the human sciences* (3rd ed.). Routledge.
- Butler, D. L. (2023). Qualitative approaches to investigating self-regulated learning: Contributions and challenges. *Using Qualitative Methods To Enrich Understandings of Self-Regulated Learning: A Special Issue of Educational Psychologist*, 1520 (January), 59–63. <https://doi.org/10.4324/9781410608529-7>
- Carter, R. A., Rice, M., Yang, S., & Jackson, H. A. (2020). Self-regulated learning in online learning environments: strategies for remote learning. *Information and Learning Science*, 121(5-6), 311–319. <https://doi.org/10.1108/ILS-04-2020-0114>
- Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical Values for Yen's Q3: Identification of Local Dependence in the Rasch Model Using Residual Correlations. *Applied Psychological Measurement*, 41(3), 178–194. <https://doi.org/10.1177/0146621616677520>
- Erdogan, T., & Senemoglu, N. (2016). Development and validation of a scale on self-regulation in learning (SSRL). *SpringerPlus*, 5(1). <https://doi.org/10.1186/s40064-016-3367-y>
- Haerudin, & Nur, I. R. D. (2020). Analysis of Student Difficulties in Inferential Statistics Courses. *Lebesgue Journal: Scientific Journal of Mathematics Education, Mathematics and Statistics*, 1(3), 208–215. <https://doi.org/10.46306/lb.v1i3.35>
- Hertel, S., Reschke, K., & Karlen, Y. (2024). Are profiles of self-regulated learning and intelligence mindsets related to students' self-regulated learning and achievement? *Learning and Instruction*, 90, Article 101850. <https://doi.org/10.1016/j.learninstruc.2023.101850>
- Holster, T. A., & Lake, J. (2016). Guessing and the Rasch Model. *Language Assessment Quarterly*, 13(2), 124–141. <https://doi.org/10.1080/15434303.2016.1160096>
- Himawan, R., & Nurgiyantoro, B. (2022). Analisis butir soal latihan penilaian akhir semester ganjil mata pelajaran bahasa Indonesia kelas VIII SMPN 1 Bambanglipuro Bantul menggunakan program ITEMAN. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, Dan Pengajarannya*, 8(1), 160–180. <https://doi.org/10.22219/kembara.v8i1.20530>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Larasati, P. E., Supahar, & Yunanta, D. R. A. (2020). Validity and reliability estimation of assessment ability instrument for data literacy on high school physics material. *Journal of Physics: Conference Series*, 1440(1). <https://doi.org/10.1088/1742-6596/1440/1/012020>
- Linacre, J. M. (2002) *What do infit and outfit, mean-square and standardized mean? Rasch Measurement Transactions*, 16(2), 878–879.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- Pintrich, P. R. R., Smith, D., Garcia, T., & McKeachie, W. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). *Ann Arbor. Michigan*, 48109 (August 2016), 1259. <https://doi.org/ED338122>
- Qin, J., Qin, J., Qiu, J., Liu, Q., Li, M., & Ma, Q. (2024). SRL-ORCA: A socially aware multi-agent mapless navigation algorithm in complex dynamic scenes. *IEEE Robotics and Automation Letters*, 9(1), 143–150. <https://doi.org/10.1109/LRA.2023.3331621>
- Ridgley, L. M., DaVia Rubenstein, L., & Callan, G. L. (2022a). Are Gifted Students Adapting Their Self-Regulated Learning Processes When Experiencing Challenging Tasks? *Gifted Child Quarterly*, 66(1), 3–22. <https://doi.org/10.1177/00169862211025452>
- Ridgley, L. M., DaVia Rubenstein, L., & Callan, G. L. (2022b). Are Gifted Students Adapting Their Self-Regulated Learning Processes When Experiencing Challenging Tasks? *Gifted Child Quarterly*, 66(1), 3–22. <https://doi.org/10.1177/00169862211025452>

- Schraw, G., Kauffman, D. F., & Lehman, S. (2006). Self-Regulated Learning. *Encyclopedia of Cognitive Science*, (January). <https://doi.org/10.1002/0470018860.s00671>
- Siekanska, M., Wilson, S. G., Blecharz, J., & Young, B. W. (2023). Examining the validity of the Polish short form version of the self-regulated learning-sport practice survey among competitive athletes. *Frontiers in Psychology*, 14(February), 1-10. <https://doi.org/10.3389/fpsyg.2023.1132608>
- Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting Self-Regulated Learning in Online Learning Environments and MOOCs: A Systematic Review. *International Journal of Human-Computer Interaction*, 35(4-5), 356-373. <https://doi.org/10.1080/10447318.2018.1543084>
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- Yen, W. M. (1984). Effects of Local Item Dependence on the Fit and Equating Performance of the Three-Parameter Logistic Model. *Applied Psychological Measurement*, 8(2), 125-145. <https://doi.org/10.1177/014662168400800201>
- Zimmerman, B. J., & Pons, M. M. (1986). Development of a Structured Interview for Assessing Student Use of Self-Regulated Learning Strategies. *American Educational Research Journal*, 23(4), 614-628. <https://doi.org/10.3102/00028312023004614>
- Zimmerman, B. J. (2010). Becoming a Self-Regulated Learner: An Overview Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, 5841(2002), 64-70. <https://doi.org/10.1207/s15430421tip4102>