Journal of Educational Sciences
Journal homepage: https://jes.ejournal.unri.ac.id/index.php/JES

P-ISSN 2581-1657
E-ISSN 2581-2203

Development and Rasch Validation of An Emotional Intelligence Scale for Secondary School Students

Indi Siti Nurihsani*, Mubiar Agustin, Idat Muqodas
Educational Psychology, Postgraduate School, Universitas Pendidikan Indonesia, Bandung, 40154, Indonesia

ARTICLE INFO

Article history:

Received: 29 May 2026
Accepted: 20 July 2026
Available Online: 25 July 2026

Keywords:

Emotional Intelligence,
Rasch Model,
Scale Development,
Secondary School,
Student Assessment

* *Corresponding author:*

E-mail: indiihsani28@upi.edu

Article Doi:

https://doi.org/10.31258/jes.10.7.p.1195-1221

This is an open access article under the CC BY-SA license.

A B S T R A C T

This study aims to develop and validate an emotional intelligence scale for secondary school students to support comprehensive assessment beyond cognitive achievement. Emotional intelligence is essential for students' ability to regulate emotions, maintain motivation, and establish positive social relationships, which contribute to academic success and well-being. This research employed a quantitative approach using the Rasch measurement model to evaluate the instrument's psychometric properties. The instrument consisted of 30 items representing five dimensions: self-awareness, self-regulation, self-motivation, empathy, and social skills, and was administered to 75 students. The results showed high item reliability (0.98) and acceptable person reliability (0.75), with a strong item separation index of 7.08. Most items fit the Rasch model, although a few items indicated misfit and require refinement. The rating scale functioned effectively, and the instrument successfully differentiated students' emotional intelligence levels. In conclusion, the developed scale is a valid and reliable tool for assessing emotional intelligence and can support educational interventions aimed at enhancing students' holistic development.

1. Introduction

The contemporary educational paradigm increasingly emphasizes the importance of students' holistic development, including cognitive, emotional, and social competencies. In this context, emotional intelligence has become a central construct in educational psychology because it supports students' ability to understand emotions, regulate emotional responses, sustain motivation, and build constructive social relationships. Emotional intelligence is particularly important for secondary school students because adolescence is a developmental stage marked by academic pressure, emotional fluctuation, peer influence, and identity formation. Recent

studies show that emotional intelligence contributes to students' academic motivation, learning engagement, resilience, and psychological well-being (Taibolatov et al., 2024). Emotional intelligence also supports students' ability to manage academic stress and adapt to changing educational demands, especially in increasingly complex learning environments (Mohamed et al., 2025). Therefore, emotional intelligence should be considered an essential component of student development rather than a supplementary psychological attribute. Emotional intelligence also contributes to personality development, self-regulation, empathy, and students' social adjustment, reinforcing its importance as a fundamental component of holistic education (Kalsum & Sirozi, 2025).

Recent empirical evidence indicates that emotional intelligence is closely related to academic achievement and students' learning processes. Emotional intelligence influences academic achievement through emotional well-being, motivation, and learning strategies, suggesting that students' emotional competence contributes indirectly to their academic success (Nieto-Carracedo et al., 2024). Previous studies also indicate that emotional intelligence positively influences students' readiness and learning outcomes, highlighting its importance in educational settings (Febriyana et al., 2023). Students with higher emotional intelligence tend to demonstrate stronger achievement motivation, better planning ability, more positive attitudes toward learning, and greater academic engagement (Bereded et al., 2025). Emotional intelligence is also associated with resilience, social support, and mental health, indicating its broader role in supporting students' psychological adjustment in educational settings (Wang et al., 2025). These findings strengthen the argument that emotional intelligence is a multidimensional construct that affects both academic and psychosocial outcomes.

The urgency of emotional intelligence assessment has also increased due to changes in contemporary learning environments. Students are now required to manage not only classroom-based academic demands but also digital interactions, social comparison, and technology-mediated communication. These conditions may increase emotional pressure and require stronger emotional regulation, empathy, and social awareness. Recent research shows that emotional intelligence contributes to holistic student development by improving resilience, motivation, and adaptive learning behavior (Xuan, 2025). Emotional intelligence is also linked to subjective well-being and academic achievement, showing that emotionally competent students are more likely to maintain psychological balance while pursuing academic goals (Caballero-García & Ruiz, 2025). Thus, schools need valid instruments to identify students' emotional strengths and difficulties accurately.

Although emotional intelligence has been widely studied, the availability of valid and reliable instruments for measuring emotional intelligence among secondary school students remains a methodological concern. Many existing instruments were developed in different cultural and educational contexts, which may limit their relevance for Indonesian students. Moreover, emotional intelligence measurement requires strong psychometric evidence to ensure that each item represents the intended construct accurately. Recent Rasch-based studies demonstrate that emotional intelligence instruments can be evaluated more rigorously by examining

item difficulty, item fit, reliability, rating scale functioning, and respondent classification (Apriliana & Masi, 2025). Rasch modeling is also useful for identifying misfitting items and ensuring that measurement results are more objective and interpretable (Setyawarno et al., 2025). Therefore, the development of an emotional intelligence scale should be supported by modern psychometric analysis.

The present study conceptualizes emotional intelligence through five dimensions: self-awareness, self-regulation, self-motivation, empathy, and social skills. These dimensions represent students' ability to recognize their own emotions, control emotional reactions, maintain learning motivation, understand others' feelings, and communicate effectively in social contexts. These dimensions are relevant to school life because students must continuously manage academic challenges, peer relationships, classroom interactions, and personal expectations. Research on teacher emotional intelligence also shows that self-awareness, self-regulation, motivation, empathy, and social skills are closely related to students' academic motivation, confirming the educational relevance of these dimensions (Rahman et al., 2024). Therefore, these five dimensions provide a strong conceptual basis for constructing an emotional intelligence scale for secondary school students.

Based on these considerations, this study aims to develop and validate an emotional intelligence scale for secondary school students using the Rasch model. The study seeks to produce a psychometrically sound instrument that can measure students' emotional intelligence accurately across the dimensions of self-awareness, self-regulation, self-motivation, empathy, and social skills. The instrument is expected to support educators, school counselors, and researchers in identifying students' emotional competence and designing appropriate educational interventions. Ultimately, this research contributes to the development of valid psychological assessment tools that support students' academic success, emotional well-being, and social adaptability.

2. Methodology

This study employed a quantitative instrument-development design to construct and validate an emotional intelligence scale for secondary school students. The study used a cross-sectional survey approach because the data were collected once from respondents to evaluate the psychometric quality of the instrument. This design is appropriate for scale development studies that aim to examine item functioning, reliability, response category performance, and measurement fairness. Recent validation studies have shown that cross-sectional survey designs are commonly used in educational and psychological instrument development because they allow researchers to obtain empirical evidence regarding the structure and quality of measurement tools within a defined respondent group (Noorkholisoh et al., 2024). In addition, Rasch-based validation has been widely recommended in recent educational measurement studies because it enables researchers to evaluate both item quality and respondent characteristics on the same measurement scale (Darman et al., 2024). Therefore, this methodological approach was selected to

ensure that the developed scale could be evaluated rigorously and replicated in similar educational contexts.

Instrument

The instrument used in this study was an emotional intelligence scale developed for secondary school students. The scale was constructed based on five dimensions: self-awareness, self-regulation, self-motivation, empathy, and social skills. These dimensions were translated into behavioral indicators and then developed into questionnaire items. The initial instrument consisted of 30 items using a five-point Likert response format, ranging from 1 = strongly disagree to 5 = strongly agree. A Likert-type format was selected because it allows respondents to express the degree of agreement with each statement and is widely used in psychological and educational scale development. Recent studies on emotional intelligence measurement have emphasized that clear item wording, appropriate response categories, and strong psychometric validation are essential to ensure that emotional constructs are measured accurately (Fernández-Berrocal et al., 2025).

The instrument development process involved item construction, expert judgment, readability testing, and empirical validation. Item construction was conducted by aligning each item with the operational indicators of the five emotional intelligence dimensions. Expert judgment was used to evaluate the relevance, clarity, and representativeness of the items before field administration. This step is important because expert review provides initial evidence of content validity and helps reduce ambiguity in item wording. Recent instrument validation studies have emphasized that expert review and readability testing are essential stages before statistical analysis because poorly written items may produce misfit or biased responses during Rasch analysis (Guo et al., 2024). Readability testing was conducted to ensure that the items were understandable for secondary school students and appropriate for their developmental level.

Data Collection

Data were collected from 75 secondary school students who completed the emotional intelligence questionnaire. The respondents were selected based on their relevance to the purpose of the study, namely students at the secondary school level who were able to understand and respond to self-report emotional statements. Before completing the questionnaire, respondents received instructions regarding the purpose of the study, the response format, and the importance of answering honestly according to their actual condition. The questionnaire was administered in a structured setting to minimize distractions and ensure that all respondents received the same instructions. This procedure is consistent with recent scale validation studies that emphasize the importance of standardized administration procedures to reduce response error and improve data quality (Setyawarno et al., 2025).

Ethical considerations were also applied during the data collection process. Participation was voluntary, and respondents were informed that their responses would be used only for research purposes. Confidentiality was maintained by

ensuring that individual responses were not publicly identified. These procedures are important in self-report psychological measurement because respondents may provide inaccurate answers if they feel pressured or unsafe. Recent studies on emotional and psychological scale validation emphasize that ethical and standardized data collection procedures are necessary to improve the trustworthiness of responses and reduce socially desirable answering patterns (Rogers et al., 2025).

Data Analysis

The data were analyzed using the Rasch measurement model with the assistance of WINSTEPS version 4.3.2. Rasch analysis was selected because it provides detailed information about item difficulty, respondent ability, item fit, person fit, reliability, separation index, rating scale functioning, and potential item bias. Unlike raw-score analysis, the Rasch model transforms ordinal responses into interval-level logit measures, allowing item and person estimates to be interpreted on the same scale. Recent studies have demonstrated that Rasch analysis is effective for validating educational and psychological instruments because it provides more precise evidence of measurement quality than conventional descriptive statistics alone(Apriyani & Aini, 2025); Pertiwi & Yudiarso, 2024). Rasch analysis has also been used in recent emotional intelligence validation studies to examine whether items function consistently in measuring emotional constructs (Apriliana & Masi, 2025).

Several criteria were used to evaluate the quality of the instrument. Item fit and person fit were examined using Infit and Outfit Mean Square values to determine whether items and respondents conformed to model expectations. Reliability and separation indices were analyzed to assess the stability of item calibration and the ability of the instrument to differentiate respondent ability levels. Rating scale functioning was evaluated by examining category use, average measures, outfit values, threshold ordering, and probability curves. Recent methodological studies have emphasized that rating scale analysis is essential in Rasch validation because response categories must function clearly and progressively to support valid interpretation of Likert-type instruments (Peeters & Augustine, 2023). Differential Item Functioning analysis was also conducted to examine whether items functioned differently across gender, age, and class groups. This analysis was necessary to ensure that the instrument measured emotional intelligence fairly across respondent characteristics. Recent Rasch studies have highlighted the importance of Differential Item Functioning analysis in identifying potential measurement bias and improving the fairness of psychological and educational instruments (Hermansyah et al., 2024).

3. Results and Discussion

The instrument development process began with item generation based on the five theoretical dimensions of emotional intelligence. The initial version consisted of 30 items, which were subsequently evaluated through expert judgment to establish

content validity and refined according to the experts' recommendations. The revised instrument then underwent readability testing to ensure that the questionnaire was appropriate for the linguistic and cognitive characteristics of secondary school students before empirical administration. Following these preliminary development stages, the final version of the instrument was administered to 75 secondary school students for psychometric validation using the Rasch measurement model. The following sections present the results of each stage, beginning with expert judgment and readability testing, followed by the empirical Rasch validation.

Expert Judgment

Prior to empirical validation, the instrument underwent content validation through expert judgment involving three experts with expertise in educational psychology and guidance and counseling. The experts evaluated the relevance, clarity, representativeness, and language appropriateness of all 30 items based on the five dimensions of emotional intelligence. Overall, the experts agreed that the instrument adequately represented the intended construct, although several suggestions were provided to improve wording clarity, sentence effectiveness, and the alignment of several items with their corresponding indicators. All recommendations were incorporated before the readability testing and empirical validation stages. A summary of the expert judgment results is presented in Table 1.

Table 1. Summary of Expert Judgment Results

Aspect Evaluated	Expert Feedback	Follow-up Revision
Relevance of indicators	Indicators were considered appropriate for measuring emotional intelligence.	Retained.
Clarity of wording	Several statements required clearer and more effective wording.	Item wording was revised.
Suitability of items with indicators	Items generally represented the intended dimensions.	Minor revisions were made to improve alignment.
Language use	Several expressions were simplified to improve comprehensibility.	Language was refined before field testing.

As shown in Table 1, the expert judgment indicated that the instrument possessed satisfactory content validity after minor revisions were made to improve the clarity and representativeness of several items. These revisions ensured that the instrument was conceptually appropriate before proceeding to the readability testing and empirical validation stages. The expert judgment process also demonstrated that the revisions primarily strengthened the instrument's content validity rather than altering its conceptual structure. Most suggestions focused on improving sentence clarity, reducing ambiguous wording, and ensuring that each statement accurately represented its intended behavioral indicator. For example, several items were revised to use more concrete expressions that are appropriate for secondary school students, while maintaining consistency with the theoretical dimensions of emotional intelligence. These refinements were intended to minimize potential

misunderstandings during empirical testing and to ensure that respondents interpreted each item as representing the targeted construct.

Readability Test

Before the empirical validation, the revised instrument was subjected to a readability test involving three junior high school students. The purpose of this stage was to evaluate whether the questionnaire instructions and item statements were easily understood by the target respondents. Overall, the students indicated that the questionnaire was understandable and that the statements were generally clear. Nevertheless, several suggestions were provided regarding the use of simpler vocabulary and clearer sentence structures to avoid multiple interpretations. Based on these comments, minor revisions were made to improve the readability of several items before the instrument was administered in the field study. A summary of the readability test results is presented in Table 2.

Table 2. Summary of Readability Test Results

Aspect Evaluated	**Findings**	**Revision**
Instructions	Easily understood by students	No revision required
Clarity of statements	Most items were understandable	Several sentences were simplified
Vocabulary	A few words were considered less familiar	Vocabulary was adjusted using simpler expressions
Overall readability	The questionnaire was considered appropriate for junior high school students	Minor language revisions were made before empirical testing

As presented in Table 2, the readability test demonstrated that the questionnaire was generally easy for students to understand. Minor revisions were made to simplify several words and improve sentence clarity, thereby enhancing the comprehensibility of the instrument before empirical validation. This test further confirmed that the instrument was appropriate for the cognitive and linguistic characteristics of secondary school students. Students generally interpreted the questionnaire instructions and response categories as intended, although several statements required simpler vocabulary and shorter sentence structures to reduce potential multiple interpretations. These revisions were implemented before field administration to improve response accuracy and minimize measurement error associated with language comprehension rather than the emotional intelligence construct itself.

Instrument Specification

The emotional intelligence scale was developed based on five theoretical dimensions: self-awareness, self-regulation, self-motivation, empathy, and social skills. Each dimension was operationalized into three behavioral indicators, with two items representing each indicator, resulting in a total of 30 items. This specification ensured that all dimensions of emotional intelligence were

proportionally represented in the instrument prior to empirical validation. This balanced distribution of items across dimensions was intended to ensure that each component of emotional intelligence was adequately represented, allowing the subsequent Rasch analysis to evaluate the psychometric performance of each construct within a comprehensive measurement framework. The complete specification of the instrument is presented in Table 3.

Table 3. Instrument Specification of the Emotional Intelligence Scale

Dimension	**Behavioral Indicator**	**Item Number**	**Number of Items**
Self-awareness	Recognizing one's own emotions	1–2	2
	Identifying the causes of emotional changes	3–4	2
	Recognizing personal strengths and weaknesses	5–6	2
Subtotal			6
Self-regulation	Controlling negative emotions	7–8	2
	Remaining calm under pressure	9–10	2
	Thinking before acting	11–12	2
Subtotal			6
Self-motivation	Striving to achieve goals	13–14	2
	Perseverance	15–16	2
	Maintaining learning motivation	17–18	2
Subtotal			6
Empathy	Understanding others' feelings	19–20	2
	Taking another person's perspective	21–22	2
	Providing emotional support	23–24	2
Subtotal			6
Social skills	Building positive relationships	25–26	2
	Working collaboratively	27–28	2
	Communicating effectively and politely	29–30	2
Subtotal			6
Total			30

As presented in Table 3, each dimension of emotional intelligence is represented by six items distributed across three behavioral indicators. This balanced distribution ensured comprehensive coverage of the emotional intelligence construct and provided a theoretical foundation for the subsequent empirical validation using the Rasch measurement model. Overall, the five dimensions demonstrated satisfactory representation in the developed instrument. Self-awareness, self-regulation, and self-motivation showed stable item performance, indicating that these dimensions were operationalized consistently. In contrast, the empathy and social skills dimensions each contained one misfitting item (Items 19 and 25), suggesting that these constructs require further refinement to improve conceptual precision and wording. Nevertheless, the balanced distribution of items across all dimensions provides comprehensive coverage of emotional intelligence while identifying specific areas for future instrument improvement.

Analisis Rasch

Item Discrimination and Reliability

The analysis of item discrimination and reliability is a crucial step in evaluating the overall quality of the measurement instrument. This analysis aims to determine whether the items are capable of distinguishing respondents based on their levels of emotional intelligence and whether the instrument produces consistent and stable measurements. By examining indicators such as mean measure, standard deviation, fit statistics, item separation, and reliability indices, researchers can assess the extent to which the instrument functions effectively within the Rasch measurement framework. The following figure presents the results of the item discrimination and reliability analysis.

TABLE 3.1 C:\Users\indii\OneDrive\Desktop\ANALIS ZOU569WS.TXT Dec 15 2025 14:59
INPUT: 75 Person 25 Item REPORTED: 75 Person 25 Item 5 CATS MINISTEP 4.3.2

SUMMARY OF 75 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	95.9	25.0	1.42	.28	1.02	-.16	.99	-.22
SEM	.9	.0	.07	.00	.06	.20	.06	.19
P.SD	7.9	.0	.61	.02	.51	1.74	.52	1.67
S.SD	7.9	.0	.61	.02	.52	1.75	.53	1.68
MAX.	115.0	25.0	3.20	.38	2.90	4.70	3.01	4.80
MIN.	73.0	25.0	-.19	.26	.27	-3.69	.30	-3.35

REAL RMSE .31 TRUE SD .53 SEPARATION 1.71 Person RELIABILITY .75
MODEL RMSE .28 TRUE SD .54 SEPARATION 1.94 Person RELIABILITY .79
S.E. OF Person MEAN = .07

Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .78 SEM = 3.70

SUMMARY OF 25 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	287.6	75.0	.00	.16	.99	-.16	.99	-.16
SEM	10.6	.0	.25	.01	.06	.34	.06	.34
P.SD	51.8	.0	1.23	.03	.30	1.67	.31	1.69
S.SD	52.9	.0	1.26	.03	.30	1.70	.32	1.72
MAX.	353.0	75.0	3.62	.23	1.87	4.28	1.88	4.30
MIN.	125.0	75.0	-1.81	.15	.62	-2.71	.62	-2.71

REAL RMSE .17 TRUE SD 1.22 SEPARATION 7.08 Item RELIABILITY .98
MODEL RMSE .17 TRUE SD 1.22 SEPARATION 7.35 Item RELIABILITY .98
S.E. OF Item MEAN = .25

Item RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000

Figure 1. The Results of The Item Discrimination and Reliability Analysis

The results of the Rasch model analysis indicate that the developed emotional intelligence instrument demonstrates strong psychometric properties in terms of item discrimination and reliability. The mean item measure was recorded at 0.00 logits, which suggests that the overall level of item difficulty is well-centered and appropriately aligned with the average ability of the respondents. The standard deviation of 1.23 logits reflects a substantial spread of item difficulty, indicating that the instrument successfully captures a wide range of emotional intelligence levels among students.

Furthermore, the average Infit and Outfit Mean Square (MNSQ) values were approximately 0.99, which fall within the acceptable range (0.5–1.5), demonstrating that the majority of items fit the Rasch model expectations. This indicates that the items function consistently in measuring the intended construct. The item separation index was 7.08, suggesting that the items can be categorized

into multiple levels of difficulty, which enhances the instrument's ability to differentiate between varying levels of emotional intelligence. Additionally, the item reliability value of 0.98 indicates a very high level of consistency in item calibration. These findings confirm that the instrument possesses excellent internal structure and measurement precision.

Item Difficulty and Respondent Ability

The examination of item difficulty and respondent ability provides important insights into the alignment between the instrument and the characteristics of the respondents. This analysis evaluates whether the difficulty level of the items is appropriate for the respondents' ability levels and whether the instrument can effectively differentiate among varying levels of emotional intelligence. Indicators such as person separation, item separation, and reliability values are used to assess the measurement precision and targeting of the instrument. The following figure illustrates the distribution of item difficulty and respondent ability.

TABLE 10.1 C:\Users\indii\OneDrive\Desktop\ANALI ZOU569WS.TXT Dec 15 2025 14:59
INPUT: 75 Person 25 Item REPORTED: 75 Person 25 Item 5 CATS MINISTEP 4.3.2

Person: REAL SEP.: 1.71 REL.: .75 ... Item: REAL SEP.: 7.08 REL.: .98

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
19	254	75	.82	.15	1.87	4.28	1.88	4.30	A .27	.43	37.3	49.5	A19
25	204	75	1.91	.15	1.81	4.06	1.81	4.03	B-.49	.43	41.3	49.9	A25
1	125	75	3.62	.16	1.48	2.77	1.58	2.98	C-.08	.41	33.3	45.5	A1
6	252	75	.87	.15	1.15	.92	1.15	.92	D .51	.43	56.0	50.0	A6
2	286	75	.12	.15	1.02	.21	1.01	.14	E .53	.42	42.7	42.8	A2
3	299	75	-.17	.15	1.01	.14	1.01	.12	F .43	.41	36.0	43.6	A3
8	286	75	.12	.15	1.01	.08	1.01	.10	G .49	.42	48.0	42.8	A8
22	331	75	-.98	.17	.94	-.32	1.01	.10	H .34	.36	56.0	50.8	A22
7	261	75	.67	.15	1.00	.03	.98	-.05	I .54	.43	52.0	47.5	A7
11	349	75	-1.62	.21	.99	.03	.92	-.25	J .41	.30	76.0	69.4	A11
16	329	75	-.92	.17	.93	-.42	.98	-.08	K .43	.37	56.0	49.0	A16
12	351	75	-1.71	.22	.96	-.13	.84	-.58	L .46	.30	76.0	71.9	A12
5	244	75	1.04	.15	.93	-.34	.93	-.34	M .40	.43	49.3	51.5	A5
10	353	75	-1.81	.23	.93	-.23	.89	-.34	l .43	.29	78.7	73.4	A10
4	308	75	-.38	.15	.89	-.74	.90	-.63	k .55	.40	41.3	44.7	A4
13	342	75	-1.34	.19	.90	-.53	.87	-.57	j .36	.33	62.7	61.2	A13
23	317	75	-.60	.16	.87	-.87	.87	-.83	i .36	.39	56.0	45.8	A23
21	246	75	1.00	.15	.85	-.88	.86	-.83	h .49	.43	50.7	51.3	A21
24	246	75	1.00	.15	.85	-.88	.86	-.83	g .49	.43	50.7	51.3	A24
9	350	75	-1.67	.21	.84	-.77	.77	-.96	f .50	.30	78.7	70.1	A9
14	293	75	-.03	.15	.81	-1.36	.80	-1.47	e .54	.42	48.0	42.6	A14
17	307	75	-.35	.15	.77	-1.77	.76	-1.72	d .52	.40	52.0	44.7	A17
18	298	75	-.14	.15	.72	-2.20	.70	-2.27	c .60	.41	54.7	43.5	A18
15	296	75	-.10	.15	.71	-2.30	.70	-2.29	b .51	.42	53.3	42.6	A15
20	262	75	.65	.15	.62	-2.71	.62	-2.71	a .48	.43	50.7	47.6	A20
MEAN	287.6	75.0	.00	.16	.99	-.2	.99	-.2			53.5	51.3	
P.SD	51.8	.0	1.23	.03	.30	1.7	.31	1.7			12.5	9.6	

Figure 2. The Distribution of Item Difficulty and Respondent Ability

The analysis of item difficulty and respondent ability further supports the quality of the instrument. The person separation index was 1.71 with a reliability value of 0.75, indicating that respondents can be grouped into approximately two distinct levels of emotional intelligence. This level of separation suggests that while the instrument is effective in distinguishing general ability differences, further refinement may enhance its sensitivity in capturing more nuanced variations among students. In contrast, the item separation index remained high at 7.08 with a reliability of 0.98, confirming that the items are well distributed across a broad

range of difficulty levels. This balance between item difficulty and respondent ability indicates that the instrument is well-targeted, meaning that the items are neither too easy nor too difficult for the sample population. Consequently, the instrument is capable of providing meaningful measurements across different levels of emotional intelligence.

The high item reliability and strong item separation found in this study indicate that the emotional intelligence scale has strong capacity to distinguish item difficulty levels. This result suggests that the items are sufficiently varied and stable in measuring emotional intelligence among secondary school students. In Rasch-based instrument validation, high item reliability is commonly interpreted as evidence that item calibration is consistent and that the instrument can be reused in similar populations. This finding is consistent with recent Rasch studies showing that well-calibrated instruments can provide more objective measurement than raw-score-based analysis because item difficulty and respondent ability are placed on the same logit scale (Darman et al., 2024). Similar findings were also reported in a recent emotional intelligence instrument study using Rasch analysis, which emphasized that strong item reliability supports the validity of emotional intelligence measurement in student populations (Tessa & Suryana, 2023). Therefore, the reliability results in this study confirm that the developed scale has adequate psychometric strength for educational assessment purposes.

Item Difficulty and Respondent Ability

The Wright map, also known as the item-person map, is used to visualize the distribution of item difficulty and respondent ability on a common logit scale. This analysis is essential to evaluate the targeting of the instrument, ensuring that the items are appropriately matched with the respondents' ability levels. By examining the spread and overlap between items and respondents, researchers can identify whether the instrument adequately covers the range of emotional intelligence levels present in the sample. The following figure presents the Wright map of the instrument.

The Wright map analysis illustrates the distribution of both item difficulty and respondent ability on the same logit scale. The results show that most respondents are located within the range of -1 to +1 logits, indicating that the majority of students possess moderate levels of emotional intelligence. A smaller proportion of respondents are positioned above +2 logits, representing individuals with higher emotional competence, while very few fall below -1 logits.

On the item side, the distribution ranges from approximately -2 to +3 logits, demonstrating that the instrument includes items with varying levels of difficulty. Item A1 appears at the highest logit level, indicating that it is the most difficult item and requires a higher level of emotional intelligence to endorse. Conversely, items A13, A11, A9, A10, and A12 are located at the lower end of the scale, suggesting that they are relatively easier for respondents.

The alignment between the mean person ability and mean item difficulty suggests good targeting of the instrument. This balance indicates that the scale is appropriate for the population being measured and can accurately assess students' emotional intelligence levels.

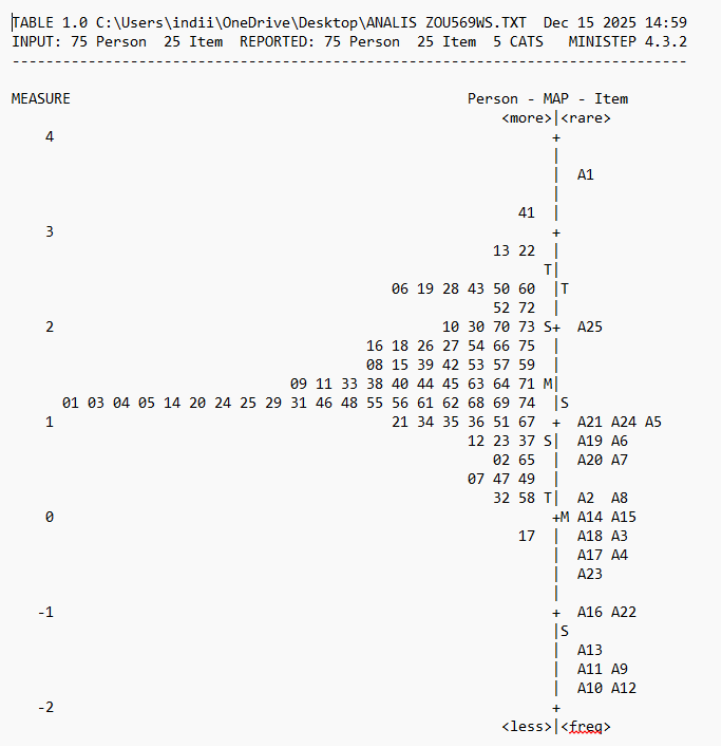


Figure 3. The Wright Map of The Instrument

The results show that item difficulty was distributed across a relatively broad range, while respondent ability was mostly concentrated at a moderate level. This pattern indicates that the instrument is generally well targeted to the sample, although further refinement may improve its sensitivity in distinguishing students with very high or very low emotional intelligence. Recent research on emotional intelligence and academic achievement also suggests that students' emotional competencies often vary in relation to motivation, emotional well-being, and learning strategies (Nieto-Carracedo et al., 2024). In this context, the presence of items with different difficulty levels is important because emotional intelligence is not a single observable behavior but a multidimensional psychological construct. A broad item difficulty range enables the instrument to capture different levels of self-awareness, regulation, motivation, empathy, and social skills. Thus, the item difficulty findings support the usefulness of the scale for identifying diverse emotional competence profiles among students.

The Wright map results show that most respondents were located within the moderate ability range, while the items were distributed from easy to difficult levels. This indicates that the scale is sufficiently aligned with the emotional intelligence characteristics of the respondents. In Rasch measurement, good alignment between item difficulty and respondent ability is important because it

shows that the instrument is neither too easy nor too difficult for the target population. Recent educational measurement studies using Rasch analysis have emphasized that item-person maps are useful for evaluating whether a test or scale adequately covers the range of respondent ability (Setyawarno et al., 2025). The finding that several items were easier while one item appeared more difficult suggests that the scale captures both basic and more complex emotional competencies. Therefore, the Wright map supports the conclusion that the instrument is appropriately targeted for secondary school students.

Item Fit and Person Fit Analysis

Item fit analysis is conducted to evaluate the extent to which each item functions consistently in measuring the intended construct of emotional intelligence. This analysis is essential to determine whether each item aligns with the expectations of the Rasch model and contributes meaningfully to the measurement process. Fit statistics, particularly Infit and Outfit Mean Square (MNSQ), are used to identify items that may deviate from the model due to randomness, ambiguity, or misinterpretation by respondents. Items with acceptable fit values indicate that they are valid indicators of the construct, while misfitting items require further evaluation or revision. The following figure presents the results of the item fit analysis.

TABLE 10.1 C:\Users\indii\OneDrive\Desktop\ANALI ZOU569WS.TXT Dec 15 2025 14:59
INPUT: 75 Person 25 Item REPORTED: 75 Person 25 Item 5 CATS MINISTEP 4.3.2

Person: REAL SEP.: 1.71 REL.: .75 ... Item: REAL SEP.: 7.08 REL.: .98

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
19	254	75	.82	.15	1.87	4.28	1.88	4.30	A .27	.43	37.3	49.5	A19
25	204	75	1.91	.15	1.81	4.06	1.81	4.03	B-.49	.43	41.3	49.9	A25
1	125	75	3.62	.16	1.48	2.77	1.58	2.98	C-.08	.41	33.3	45.5	A1
6	252	75	.87	.15	1.15	.92	1.15	.92	D .51	.43	56.0	50.0	A6
2	286	75	.12	.15	1.02	.21	1.01	.14	E .53	.42	42.7	42.8	A2
3	299	75	-.17	.15	1.01	.14	1.01	.12	F .43	.41	36.0	43.6	A3
8	286	75	.12	.15	1.01	.08	1.01	.10	G .49	.42	48.0	42.8	A8
22	331	75	-.98	.17	.94	-.32	1.01	.10	H .34	.36	56.0	50.8	A22
7	261	75	.67	.15	1.00	.03	.98	-.05	I .54	.43	52.0	47.5	A7
11	349	75	-1.62	.21	.99	.03	.92	-.25	J .41	.30	76.0	69.4	A11
16	329	75	-.92	.17	.93	-.42	.98	-.08	K .43	.37	56.0	49.0	A16
12	351	75	-1.71	.22	.96	-.13	.84	-.58	L .46	.30	76.0	71.9	A12
5	244	75	1.04	.15	.93	-.34	.93	-.34	M .40	.43	49.3	51.5	A5
10	353	75	-1.81	.23	.93	-.23	.89	-.34	l .43	.29	78.7	73.4	A10
4	308	75	-.38	.15	.89	-.74	.90	-.63	k .55	.40	41.3	44.7	A4
13	342	75	-1.34	.19	.90	-.53	.87	-.57	j .36	.33	62.7	61.2	A13
23	317	75	-.60	.16	.87	-.87	.87	-.83	i .36	.39	56.0	45.8	A23
21	246	75	1.00	.15	.85	-.88	.86	-.83	h .49	.43	50.7	51.3	A21
24	246	75	1.00	.15	.85	-.88	.86	-.83	g .49	.43	50.7	51.3	A24
9	350	75	-1.67	.21	.84	-.77	.77	-.96	f .50	.30	78.7	70.1	A9
14	293	75	-.03	.15	.81	-1.36	.80	-1.47	e .54	.42	48.0	42.6	A14
17	307	75	-.35	.15	.77	-1.77	.76	-1.72	d .52	.40	52.0	44.7	A17
18	298	75	-.14	.15	.72	-2.20	.70	-2.27	c .60	.41	54.7	43.5	A18
15	296	75	-.10	.15	.71	-2.30	.70	-2.29	b .51	.42	53.3	42.6	A15
20	262	75	.65	.15	.62	-2.71	.62	-2.71	a .48	.43	50.7	47.6	A20
MEAN	287.6	75.0	.00	.16	.99	-.2	.99	-.2			53.5	51.3	
P.SD	51.8	.0	1.23	.03	.30	1.7	.31	1.7			12.5	9.6	

Figure 4. The Results of Item Fit Analysis

The item fit analysis indicates that the majority of items demonstrate acceptable fit statistics, with Infit and Outfit MNSQ values close to 1.0 and standardized Z-values near zero. This suggests that most items function consistently and align well with the Rasch model expectations. However, Item 19 and Item 25 exhibit relatively high MNSQ values (approximately 1.8) along with elevated ZSTD values (around 4.0),

indicating potential misfit. These results suggest that the response patterns for these items may be inconsistent, possibly due to ambiguity or varying interpretations among respondents. Nevertheless, all items show positive point-measure correlations ranging from 0.27 to 0.60, indicating that they still contribute positively to distinguishing respondents based on their ability levels.

The misfit observed in Item 19 and Item 25 warrants further conceptual consideration beyond their statistical indicators because both items may capture constructs extending beyond the intended dimensions of emotional intelligence. According to Goleman's emotional intelligence framework, each item should represent a specific emotional competency while minimizing contamination from unrelated psychological characteristics. Consequently, item misfit may indicate not only measurement inconsistency but also imperfect alignment between item wording and the underlying construct.

Item 19 ("I know when my friend is feeling sad even if they do not tell me") was designed to assess the empathy dimension. Conceptually, empathy refers to recognizing, understanding, and responding appropriately to other people's emotions through observable emotional cues and interpersonal awareness. However, this item requires respondents to infer another person's emotional state without explicit behavioral or verbal information. Such judgments may depend on friendship closeness, frequency of interaction, familiarity with the friend, or situational knowledge rather than empathy itself. As a result, respondents with comparable levels of emotional intelligence may interpret the statement differently, producing inconsistent response patterns that contribute to Rasch misfit. To improve construct validity, the item could be revised to focus on observable empathic behavior rather than implicit emotional prediction, for example: *"I pay attention to changes in my friend's emotions and try to understand how they feel."*

Similarly, Item 25 ("I easily make friends with new people") was intended to measure social skills. Within emotional intelligence theory, however, social skills emphasize effective communication, cooperation, conflict management, and maintaining constructive interpersonal relationships rather than the ease of initiating friendships. The current wording may instead reflect stable personality characteristics such as extraversion, sociability, or self-confidence, which are theoretically distinct from emotional intelligence. Consequently, students with strong emotional regulation and interpersonal competence but introverted personalities may provide lower ratings, thereby reducing item fit. A more construct-focused alternative would emphasize interpersonal effectiveness rather than sociability, for example: *"I communicate respectfully and cooperate effectively with people I have just met."*

These findings demonstrate that Rasch misfit should not automatically be interpreted as evidence that an item must be discarded. Instead, misfitting items provide valuable methodological information for improving construct representation and item wording. In the present study, Item 19 and Item 25 continue to show positive point-measure correlations, indicating that they still contribute to measuring emotional intelligence. Nevertheless, refining their wording to better

reflect the theoretical definitions of empathy and social skills is expected to improve construct validity and strengthen the psychometric quality of the instrument in future validation studies.

While item fit analysis focuses on evaluating the quality and consistency of each item in measuring the intended construct, it is equally important to examine how respondents interact with these items. A well-functioning instrument not only requires valid and reliable items but also depends on consistent response patterns from participants. Therefore, after ensuring that the items conform to the Rasch model expectations, the next step is to analyze person fit to determine whether respondents provide answers that align with their estimated ability levels. This complementary analysis allows for a more comprehensive evaluation of the measurement process by considering both the instrument structure and respondent behavior.

Person fit analysis is performed to assess the consistency of respondents' answer patterns in relation to the Rasch model. This analysis aims to identify whether respondents provide responses that are logically consistent with their estimated ability levels. Fit statistics such as Infit and Outfit Mean Square (MNSQ) are used to detect unusual or inconsistent response patterns, which may arise from factors such as misunderstanding of items, lack of attention, or random answering behavior. Identifying person misfit is important to ensure the accuracy and validity of the measurement results. The following figure illustrates the results of the person fit analysis.

BETTER FITTING NOT SHOWN													
28	107	25	2.30	.30	.81	-.63	.66	-.95	.82	.66	52.0	55.4	28PAX
45	97	25	1.48	.28	.79	-.73	.79	-.67	.73	.69	56.0	49.5	45PCZ
33	95	25	1.33	.27	.78	-.79	.74	-.89	.78	.70	64.0	50.5	33LAX
57	99	25	1.63	.28	.74	-.94	.75	-.82	.74	.69	56.0	51.3	57PBY
73	103	25	1.95	.29	.75	-.90	.68	-1.02	.80	.68	44.0	51.7	73PAX
72	105	25	2.12	.30	.74	-.94	.63	-1.17	z .76	.67	76.0	52.4	72PBY
10	104	25	2.04	.29	.73	-1.01	.62	-1.23	y .88	.67	40.0	52.1	10LBY
46	94	25	1.25	.27	.73	-.99	.71	-1.04	x .79	.70	52.0	50.4	46LCZ
25	92	25	1.11	.27	.70	-1.11	.68	-1.16	w .73	.70	60.0	50.1	25PBY
32	80	25	.28	.26	.60	-1.59	.66	-1.27	v .61	.69	72.0	50.0	32PBY
48	94	25	1.25	.27	.65	-1.34	.66	-1.27	u .75	.70	64.0	50.4	48PBY
34	91	25	1.04	.27	.65	-1.36	.63	-1.45	t .82	.70	56.0	49.2	34PCZ
11	97	25	1.48	.28	.64	-1.41	.64	-1.32	s .83	.69	52.0	49.5	11PCZ
26	100	25	1.71	.28	.58	-1.75	.63	-1.31	r .79	.69	64.0	51.1	26PAX
3	93	25	1.18	.27	.60	-1.59	.59	-1.61	q .83	.70	60.0	50.1	03LAX
61	94	25	1.25	.27	.60	-1.63	.59	-1.61	p .77	.70	80.0	50.4	61PAX
64	97	25	1.48	.28	.58	-1.73	.56	-1.74	o .80	.69	60.0	49.5	64PCZ
1	92	25	1.11	.27	.54	-1.94	.56	-1.76	n .76	.70	56.0	50.1	01PAX
44	96	25	1.40	.27	.56	-1.83	.54	-1.86	m .87	.70	64.0	50.5	44PBY
53	99	25	1.63	.28	.56	-1.85	.52	-1.88	l .84	.69	72.0	51.3	53PBY
54	102	25	1.87	.29	.56	-1.85	.50	-1.85	k .88	.68	68.0	50.1	54PAX
66	100	25	1.71	.28	.52	-2.07	.48	-2.04	j .89	.69	64.0	51.1	66PBY
74	92	25	1.11	.27	.49	-2.20	.50	-2.10	i .79	.70	64.0	50.1	74LCZ
2	86	25	.68	.26	.49	-2.17	.48	-2.21	h .83	.70	60.0	46.7	02PAX
51	91	25	1.04	.27	.44	-2.50	.46	-2.34	g .80	.70	72.0	49.2	51PCZ
12	88	25	.82	.26	.42	-2.62	.42	-2.58	f .85	.70	68.0	47.6	12LCZ
68	93	25	1.18	.27	.40	-2.74	.41	-2.61	e .92	.70	76.0	50.1	68PCZ
67	91	25	1.04	.27	.37	-2.96	.38	-2.85	d .87	.70	72.0	49.2	67LBY
62	94	25	1.25	.27	.35	-3.10	.35	-3.00	c .86	.70	80.0	50.4	62PBY
35	91	25	1.04	.27	.27	-3.69	.31	-3.30	b .85	.70	64.0	49.2	35PAX
63	97	25	1.48	.28	.28	-3.69	.30	-3.35	a .93	.69	84.0	49.5	63PBY
MEAN	95.9	25.0	1.42	.28	1.02	-.2	.99	-.2			53.5	51.3	
P.SD	7.9	.0	.61	.02	.51	1.7	.52	1.7			14.7	3.8	

Figure 5. The Results of Person Fit Analysis

The person fit analysis shows that the majority of respondents demonstrate acceptable fit statistics, indicating consistent response patterns in accordance with the Rasch model. Most respondents have Infit and Outfit MNSQ values within the acceptable range, suggesting that their responses reflect their actual levels of

emotional intelligence. However, a small number of respondents exhibit extreme fit values, which may indicate inconsistent answering behavior or unusual response patterns. These inconsistencies could be caused by factors such as misunderstanding of the items or lack of attention during the response process. Overall, the results suggest that the respondent data are generally reliable and suitable for further analysis.

The person fit results indicate that most respondents provided consistent answers, although a small number showed unusual response patterns. This finding suggests that the majority of students understood the items and responded in accordance with their emotional intelligence levels. In psychological and educational measurement, person misfit may arise from careless responses, misunderstanding of item content, fatigue, or inconsistent self-perception. Recent studies using Rasch analysis in educational contexts have emphasized that person fit is important because instrument validity depends not only on item quality but also on the consistency of respondent behavior (Apriliana & Masi, 2025). Therefore, the acceptable person fit in this study strengthens the credibility of the data. Nevertheless, future studies should consider improving administration procedures, providing clearer instructions, and checking response consistency to minimize person misfit.

Reliability of Item, Person, and Instrument

Reliability analysis is performed to assess the consistency and stability of the measurement instrument. This includes evaluating both item reliability and person reliability, as well as the separation indices that indicate the instrument's ability to distinguish between different levels of item difficulty and respondent ability. High reliability values reflect consistent measurement results, while separation indices provide insight into the precision of classification. Because these statistics were obtained from the Rasch summary analysis presented in Figure 1, the following discussion focuses on their interpretation rather than reproducing the same output.

The overall reliability analysis confirms that the instrument is highly reliable. The item reliability value of 0.98 indicates that the items are consistently calibrated and stable across the measurement process. The person reliability value of 0.75 suggests a moderate level of consistency in respondents' answers, which is acceptable for psychological measurement. The separation indices further support these findings, with item separation at 7.08 and person separation at 1.71. These values indicate that the instrument is highly effective in distinguishing item difficulty levels and reasonably effective in differentiating respondent abilities. The overall results demonstrate that the instrument provides stable and reliable measurements of emotional intelligence among secondary school students.

The reliability findings indicate that the instrument has very strong item reliability and acceptable person reliability. This means that the items are stable and well calibrated, while respondents can be classified into general levels of emotional intelligence. Similar results have been reported in recent Rasch validation studies, where high item reliability was interpreted as evidence of strong item quality, while moderate person reliability was considered acceptable when the respondent group

was relatively homogeneous (Dewi et al., 2025). In the present study, the moderate person separation may reflect the fact that most respondents were secondary school students with relatively similar developmental and educational backgrounds. This pattern is reasonable because emotional intelligence among students may cluster around moderate levels due to shared school experiences. Overall, the reliability results confirm that the instrument is sufficiently stable for educational assessment and can be improved further by testing it with a larger and more diverse sample. Similar findings have also been reported by Nurani et al. (2025), who demonstrated that Rasch analysis provides strong evidence for evaluating the psychometric quality of educational assessment instruments.

Differential Item Functioning (DIF)

DIF based on gender

Differential Item Functioning (DIF) analysis based on gender is conducted to examine whether the instrument functions equivalently for male and female respondents. This analysis is important to ensure that no item unfairly advantages or disadvantages a particular gender group. By comparing item performance across gender categories, researchers can identify potential bias that may affect the validity of the measurement results. The absence of substantial DIF indicates that the instrument measures emotional intelligence fairly across gender, whereas significant DIF suggests that certain items require further review to improve measurement fairness. The following figure presents the DIF analysis based on gender.

P	.03	3.56	.19	L	-.12	3.85	.34	-.30	.39	-.76	27	.4511	.6691	.4134	-.64	9	1 A1
P	.00	.12	.17	L	-.01	.12	.30	.00	.35	.00	28	1.000	.2830	.5947	-.31	9	2 A2
P	.06	-.27	.17	L	-.18	.13	.30	-.40	.35	-1.15	29	.2613	.0721	.7884	-.19	9	3 A3
P	-.01	-.38	.18	L	.04	-.44	.32	.07	.37	.18	28	.8569	.0224	.8811	-.11	9	4 A4
P	.07	.92	.17	L	-.23	1.42	.30	-.49	.35	-1.42	28	.1663	1.3426	.2466	-.86	9	5 A5
P	.05	.78	.17	L	-.16	1.14	.30	-.36	.35	-1.04	28	.3074	.1009	.7507	-.20	9	6 A6
P	-.01	.69	.17	L	.05	.59	.30	.10	.35	.30	28	.7673	.0054	.9416	.05	9	7 A7
P	-.01	.15	.17	L	.05	.04	.30	.10	.35	.30	28	.7655	2.3394	.1261	-1.33	9	8 A8
P	-.02	-1.61	.24	L	.06	-1.93	.52	.32	.57	.57	24	.5731	2.1105	.1463	1.15	9	9 A9
P	-.02	-1.73	.25	L	.08	-2.24	.59	.51	.64	.80	23	.4322	2.8905	.0891	2.59	9	10 A10
P	.03	-1.73	.25	L	-.09	-1.32	.40	-.41	.47	-.86	30	.3961	.0830	.7732	-.24	9	11 A11
P	-.07	-1.50	.23	L	.22	-3.36	1.01	1.86	1.03	1.80	18	.0883	2.4823	.1151	1.65	9	12 A12
P	.04	-1.45	.22	L	-.12	-1.02	.37	-.43	.43	-1.00	30	.3266	.2397	.6244	-.31	9	13 A13
P	.00	-.03	.17	L	.01	-.05	.31	.02	.35	.06	28	.9527	.8803	.3481	-.71	9	14 A14
P	-.06	.00	.17	L	.19	-.44	.32	.44	.36	1.22	27	.2324	.2519	.6158	.32	9	15 A15
P	-.05	-.82	.19	L	.15	-1.32	.40	.50	.45	1.11	24	.2778	.8477	.3572	.68	9	16 A16
P	-.03	-.30	.17	L	.11	-.55	.33	.25	.37	.68	27	.5053	.4694	.4932	.41	9	17 A17
P	-.03	-.09	.17	L	.11	-.34	.32	.25	.36	.71	27	.4855	.0920	.7617	-.20	9	18 A18
P	.06	.72	.17	L	-.19	1.14	.30	-.42	.35	-1.20	28	.2390	1.2045	.2724	-.70	9	19 A19
P	.01	.65	.17	L	-.02	.68	.30	-.03	.35	-.09	28	.9264	.9436	.3313	.69	9	20 A20
P	.03	.95	.17	L	-.09	1.14	.30	-.19	.35	-.54	28	.5930	.0126	.9108	-.07	9	21 A21
P	-.02	-.93	.19	L	.07	-1.16	.38	.23	.43	.53	26	.5977	.5008	.4792	.62	9	22 A22
P	-.08	-.45	.18	L	.25	-1.16	.38	.71	.42	1.68	24	.1060	3.2610	.0709	1.31	9	23 A23
P	.03	.95	.17	L	-.09	1.14	.30	-.19	.35	-.54	28	.5930	.0126	.9108	-.07	9	24 A24
P	.03	1.87	.17	L	-.09	2.05	.30	-.18	.34	-.54	28	.5950	.0126	.9108	.07	9	25 A25

Figure 6. DIF Based on Gender

The results of the DIF analysis based on gender indicate that most items do not show significant differences between male and female respondents. This is evidenced by p-values that are generally greater than 0.05, suggesting that the instrument functions fairly across gender groups. In other words, both male and female students have an equal probability of responding to the items when controlling for their ability levels. However, several items exhibit significant DIF, as indicated by high t-values and p-values below 0.05. These items may contain potential gender bias, possibly due to differences in interpretation or contextual

relevance. Therefore, such items require further review to ensure fairness and equivalence across gender groups. Overall, the instrument demonstrates good measurement invariance in terms of gender.

The DIF analysis based on gender shows that most items did not function differently for male and female respondents. This indicates that the instrument is generally fair across gender groups. However, several items showed potential gender-related bias and therefore require further review. Recent research emphasizes that DIF analysis is important because an item may appear valid overall but still function differently for specific demographic groups (Chen et al., 2025). In emotional intelligence measurement, gender-related DIF may occur because boys and girls may differ in emotional expression, communication style, or social expectations. Therefore, items that show potential gender bias should be reviewed in terms of wording, context, and emotional content. Overall, the gender DIF results suggest that the scale is largely fair, but several items should be refined to strengthen measurement invariance.

DIF based on age

DIF analysis based on age is conducted to evaluate whether the instrument performs consistently across different age groups. This analysis aims to determine whether items are interpreted similarly by respondents of varying developmental stages. Ensuring that items function equivalently across age groups is essential to maintain the validity and generalizability of the instrument. Furthermore, DIF analysis helps identify items that may unintentionally favor or disadvantage particular age groups despite respondents having comparable levels of emotional intelligence. The following figure/table illustrates the DIF analysis based on age.

The results of the DIF analysis based on age show that the majority of items do not exhibit significant differences across age groups. Most p-values are above 0.05, indicating that the instrument is generally fair and consistent for respondents of different ages. This suggests that the items are appropriately understood and interpreted regardless of age differences. However, a small number of items display significant DIF, as indicated by p-values below 0.05. These items may be influenced by differences in cognitive maturity, experience, or social context among respondents of different age groups. Consequently, these items should be carefully reviewed to ensure that they do not introduce bias. Overall, the findings indicate that the instrument has good cross-age validity.

The DIF analysis based on age indicates that most items functioned consistently across age groups, although a few items showed potential age-related bias. This result suggests that the instrument is generally appropriate for students of different ages within the secondary school range. Age-related DIF may occur because emotional understanding, self-regulation, and social awareness develop gradually during adolescence. Recent emotional intelligence studies have shown that emotional intelligence is associated with well-being, academic achievement, and developmental adaptation, making age sensitivity an important consideration in scale development (Caballero-García & Ruiz, 2025). Therefore, items showing

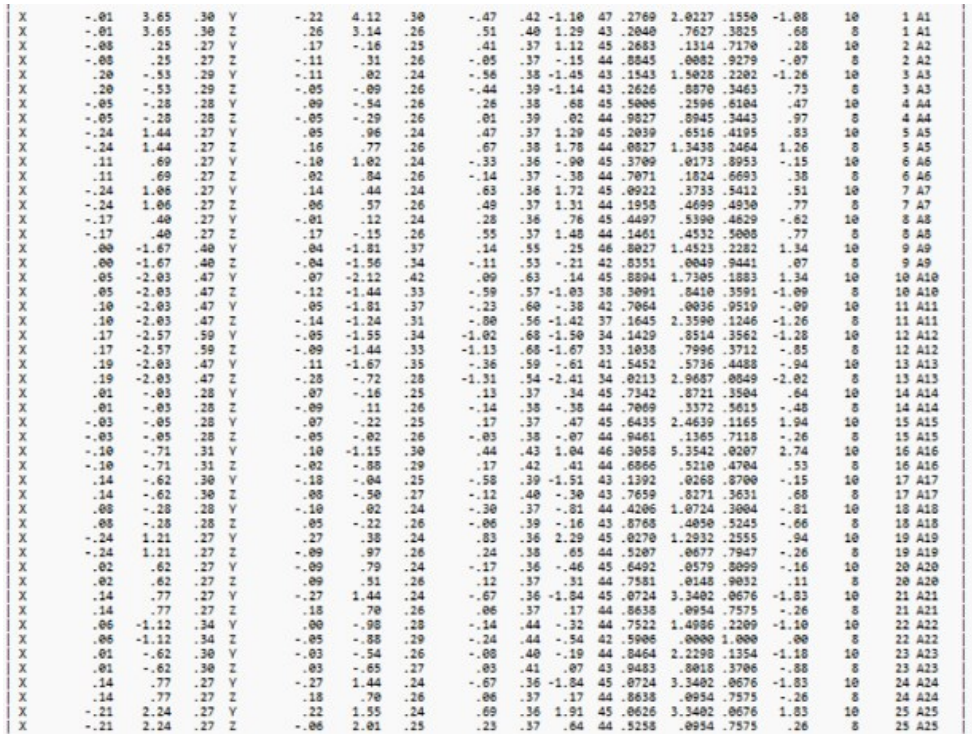
potential age bias should be reviewed to ensure that they are equally understandable for younger and older students. The overall results support the use of the scale across age groups, with minor refinement needed for specific items.

Person CLASS/	Obs-Exp Average	DIF MEASURE	DIF S.E.	Person CLASS/	Obs-Exp Average	DIF MEASURE	DIF S.E.	DIF CONTRAST	DIF S.E.	JOINT t	Rasch-Welch Prob.		Mantel Chi-sqw	Prob.	Size CUMOR	Active Slices	Item Number	Name
A	-.01	3.65	.30	B	-.22	4.12	.30	-.47	.42	-1.10	47	.2769	2.8227	.1558	-1.08	10	1	A1
A	-.06	3.65	.30	C	.26	3.14	.26	.51	.40	1.29	43	.2040	.7627	.3825	.68	10	1	A1
A	-.08	.25	.27	B	.17	-.16	.26	.41	.37	1.12	45	.2683	.1314	.7170	.28	10	2	A2
A	-.08	.25	.27	C	-.11	.31	.26	-.05	.37	-1.15	44	.3845	.8082	.9279	-.07	8	2	A2
A	.20	-.53	.29	B	-.11	.02	.24	-.56	.38	-1.45	43	.1543	1.5020	.2202	-1.26	8	3	A3
A	.20	-.53	.29	C	-.05	-.09	.26	-.44	.39	-1.14	43	.2626	.8870	.3463	.73	10	3	A3
A	-.05	-.28	.28	B	.05	-.54	.26	.01	.39	.62	45	.5006	.2945	.6104	.47	10	4	A4
A	-.05	-.28	.28	C	-.05	-.29	.26	.01	.39	.02	44	.9827	.8945	.3443	.97	8	4	A4
A	-.24	1.44	.27	B	.05	.96	.24	.47	.37	1.29	45	.2039	.6561	.4195	.83	10	5	A5
A	-.24	1.44	.27	C	.16	.77	.24	.67	.38	.78	44	.0827	1.3438	.2464	1.26	8	6	A5
A	.11	.69	.27	B	-.10	1.02	.24	-.33	.36	-.90	45	.3790	.0173	.8953	.15	8	6	A6
A	.11	.69	.27	C	.02	.44	.26	-.14	.37	-.38	44	.7071	.1824	.6693	.38	8	6	A6
A	-.24	1.86	.27	B	-.14	.44	.24	.63	.36	-.72	45	.0922	.3739	.5412	.51	10	7	A7
A	-.24	1.86	.27	C	.06	.57	.26	.49	.37	1.31	44	.1958	.4699	.4933	.77	8	7	A7
A	-.17	.40	.27	B	-.01	.12	.24	.28	.36	.76	45	.4497	.5390	.4629	-.62	10	8	A8
A	-.17	.40	.27	C	.17	-.15	.26	.55	.37	1.48	44	.1461	.4532	.5008	.77	8	8	A8
A	.00	-1.67	.40	B	.04	-1.81	.37	.14	.55	.25	46	.8027	1.4523	.2282	1.34	10	9	A9
A	.00	-1.67	.40	C	-.04	-1.56	.34	-.11	.53	-.21	42	.8351	.0489	.9441	.07	8	9	A9
A	.05	-2.03	.47	B	.07	-2.12	.42	.09	.57	.14	45	.8894	1.7365	.1683	1.34	10	10	A10
A	.05	-2.03	.47	C	-.12	-1.44	.33	-.59	.57	-1.03	33	.3091	.4410	.3591	-1.09	8	10	A10
A	.10	-2.03	.47	B	.05	-1.81	.37	-.23	.60	-.38	42	.7064	.8936	.9519	-.09	10	11	A11
A	.10	-2.03	.47	C	-.14	-1.24	.31	-.80	.56	-1.42	37	.1645	2.3594	.1246	-1.26	8	11	A11
A	.17	-2.57	.59	B	-.05	-1.55	.34	-1.02	.68	-1.50	34	.1429	.8514	.3562	-1.26	10	12	A12
A	.17	-2.57	.59	C	-.09	-1.44	.33	-1.13	.66	-.67	33	.1038	2.7996	.3712	-.85	8	12	A12
A	.19	-2.03	.47	B	.11	-1.67	.35	-.36	.59	-.61	41	.5452	2.5736	.4483	-.94	10	13	A13
A	.19	-2.03	.47	C	-.28	-.72	.28	-.38	.54	-2.41	34	.0213	.9833	.0849	.01	8	13	A13
A	.01	-.03	.28	B	.07	-.16	.25	.13	.37	.34	45	.7342	.8721	.3504	.64	10	14	A14
A	.01	-.03	.28	C	-.09	.11	.26	-.14	.38	-.38	44	.7069	.3372	.5615	-.48	8	14	A14
A	-.03	-.05	.28	B	.07	-.22	.25	.17	.37	.47	45	.6435	2.4639	.1165	1.94	10	15	A15
A	-.03	-.05	.28	C	-.05	-.02	.26	.03	.38	-.07	44	.9461	.1365	.7118	3.26	8	15	A15
A	-.10	-.71	.31	B	.10	-1.15	.30	.44	.43	1.04	46	.3956	5.3542	.0207	.24	10	16	A16
A	-.10	-.71	.31	C	-.02	.88	.29	-.55	.42	-1.41	44	.6866	.5210	.4704	.53	8	16	A16
A	.14	-.62	.30	B	.18	-.04	.25	-.12	.39	-1.51	43	.1192	.0268	.8700	-.15	10	17	A17
A	.14	-.62	.28	C	.08	-.50	.27	-.12	.49	-.30	43	.7659	1.6721	.3631	-.81	8	18	A17
A	.08	-.28	.28	B	-.10	.02	.24	-.30	.37	.81	44	.4296	.4950	.3004	-.66	10	18	A18
A	.08	-.28	.28	C	.05	-.22	.26	.06	.39	.16	43	.8768	1.9821	.5245	.94	8	18	A18
A	-.24	1.21	.27	B	.27	.35	.24	.83	.36	2.29	45	.0270	.9821	.2555	.94	10	19	A19
A	-.24	1.21	.27	C	-.09	.97	.26	.24	.38	.65	44	.5207	.8057	.7947	-.26	8	19	A19
A	.02	.62	.27	B	-.09	.79	.24	-.17	.36	-.46	45	.6492	.9579	.8099	-.16	10	20	A20
A	.02	.62	.27	C	.09	.51	.26	.12	.37	.31	44	.7581	.9148	.9032	.11	8	20	A20
A	.14	.77	.27	B	-.27	1.44	.24	-.67	.36	-1.84	45	.0724	3.3402	.0676	-1.83	10	21	A21
A	.14	.77	.27	C	.18	.70	.26	.06	.37	.17	44	.8638	.9054	.7575	-.26	8	22	A21
A	.06	-1.12	.34	B	.00	-.98	.28	-.14	.44	-.32	44	.7522	1.4986	.2209	-1.10	10	22	A22
A	.06	-1.12	.34	C	-.05	-.65	.29	-.24	.44	-.54	42	.5986	.4008	1.000	.00	8	23	A22
A	.11	-.77	.30	B	-.27	1.44	.24	-.67	.10	1.84	43	.9468	2.2610	.3754	1.88	10	23	A23
A	.14	.77	.27	C	.18	.70	.26	.06	.37	.17	44	.0238	.9054	.8376	-.26	8	24	A24
A	-.21	2.74	.27	B	.22	1.55	.24	.69	.36	1.91	46	.0656	3.3402	.0676	1.83	10	25	A25
A	-.21	2.24	.27	C	-.06	2.01	.25	.23	.37	.64	44	.5258	.9054	.7575	.26	8	25	A25

Figure 7 DIF Based on Age

DIF based on class

Differential Item Functioning (DIF) analysis based on class level was conducted to evaluate whether the emotional intelligence scale functioned equivalently across students from different grade levels. This analysis is important because variations in academic experience, developmental maturity, and school-related social interactions may influence how students interpret and respond to questionnaire items. By comparing item performance across class groups while controlling for respondents' ability levels, DIF analysis identifies items that may provide an unfair advantage or disadvantage to particular groups. The absence of substantial DIF indicates that the instrument measures emotional intelligence consistently regardless of students' class level, whereas significant DIF suggests that certain items may require further review to improve measurement fairness. Therefore, this analysis provides important evidence regarding the measurement invariance and fairness of the developed instrument. The following figure presents the results of the DIF analysis based on class level.



X	-.01	3.65	.30	Y	-.22	4.12	.30	-.47	.42	-1.10	47	.2769	2.0227	.1550	-1.08	10	1	A1
X	-.01	3.65	.30	Z	.26	3.14	.26	.51	.40	1.29	43	.2040	.7627	.3825	.68	8	1	A1
X	-.08	.25	.27	Y	.17	-.16	.25	.41	.37	1.12	45	.2683	.1314	.7170	.28	10	2	A2
X	-.08	.25	.27	Z	-.11	.31	.26	-.05	.37	-.15	44	.8845	.0082	.9279	-.07	8	2	A2
X	.20	-.53	.29	Y	-.11	.02	.24	-.56	.38	-1.45	43	.1543	1.5028	.2202	-1.26	10	3	A3
X	.20	-.53	.29	Z	-.05	-.09	.26	-.44	.39	-1.14	43	.2626	.8870	.3463	.73	8	3	A3
X	-.05	-.28	.28	Y	.09	-.54	.26	.26	.38	.68	45	.5006	.2596	.6104	.47	10	4	A4
X	-.05	-.28	.28	Z	-.05	-.29	.26	.01	.39	.02	44	.9827	.8945	.3443	.97	8	4	A4
X	-.24	1.44	.27	Y	.05	.96	.24	.47	.37	1.29	45	.2039	.6516	.4195	.83	10	5	A5
X	-.24	1.44	.27	Z	.16	.77	.26	.67	.38	1.78	44	.0827	1.3438	.2464	1.26	8	5	A5
X	.11	.69	.27	Y	-.10	1.02	.24	-.33	.36	-.90	45	.3709	.0173	.8953	-.15	10	6	A6
X	.11	.69	.27	Z	.02	.84	.26	-.14	.37	-.38	44	.7071	.1824	.6693	.38	8	6	A6
X	-.24	1.06	.27	Y	.14	.44	.24	.63	.36	1.72	45	.0922	.3733	.5412	.51	10	7	A7
X	-.24	1.06	.27	Z	.06	.57	.26	.49	.37	1.31	44	.1958	.4699	.4830	.77	8	7	A7
X	-.17	.40	.27	Y	-.01	.12	.24	.28	.36	.76	45	.4497	.5390	.4629	-.62	10	8	A8
X	-.17	.40	.27	Z	.17	-.15	.26	.55	.37	1.48	44	.1461	.4532	.5008	.77	8	8	A8
X	.00	-1.67	.40	Y	.04	-1.81	.37	.14	.55	.25	46	.8027	1.4523	.2282	1.34	10	9	A9
X	.00	-1.67	.40	Z	-.04	-1.56	.34	-.11	.53	-.21	42	.8351	.0049	.9441	.07	8	9	A9
X	.05	-2.03	.47	Y	.07	-2.12	.42	.09	.63	.14	45	.8894	1.7305	.1883	1.34	10	10	A10
X	.05	-2.03	.47	Z	-.12	-1.44	.33	-.59	.57	-1.03	38	.3091	.8410	.3591	-1.09	8	10	A10
X	.10	-2.03	.47	Y	.05	-1.81	.37	-.23	.60	-.38	42	.7064	.0036	.9519	-.09	10	11	A11
X	.10	-2.03	.47	Z	-.14	-1.24	.31	-.80	.56	-1.42	37	.1645	2.3590	.1246	-1.26	8	11	A11
X	.17	-2.57	.59	Y	-.05	-1.55	.34	-1.02	.68	-1.50	34	.1429	.8514	.3562	-1.28	10	12	A12
X	.17	-2.57	.59	Z	-.09	-1.44	.33	-1.13	.68	-1.67	33	.1038	.7996	.3712	-.85	8	12	A12
X	.19	-2.03	.47	Y	.11	-1.67	.35	-.36	.59	-.61	41	.5452	.5736	.4488	-.94	10	13	A13
X	.19	-2.03	.47	Z	-.28	-.72	.28	-1.31	.54	-2.41	34	.0213	2.9687	.0849	-2.02	8	13	A13
X	.01	-.03	.28	Y	.07	-.16	.25	.13	.37	.34	45	.7342	.8721	.3504	.64	10	14	A14
X	.01	-.03	.28	Z	-.09	.11	.26	-.14	.38	-.38	44	.7069	.3372	.5615	-.48	8	14	A14
X	-.03	-.05	.28	Y	.07	-.22	.25	.17	.37	.47	45	.6435	2.4639	.1165	1.94	10	15	A15
X	-.03	-.05	.28	Z	-.05	-.02	.26	-.03	.38	-.07	44	.9461	.1365	.7118	-.26	8	15	A15
X	-.10	-.71	.31	Y	.10	-1.15	.30	.44	.43	1.04	46	.3058	5.3542	.0207	2.74	10	16	A16
X	-.10	-.71	.31	Z	-.02	-.88	.29	.17	.42	.41	44	.6866	.5210	.4704	.53	8	16	A16
X	.14	-.62	.30	Y	-.18	-.04	.25	-.58	.39	-1.51	43	.1392	.0268	.8700	-.15	10	17	A17
X	.14	-.62	.30	Z	.08	-.50	.27	-.12	.40	-.30	43	.7659	.8171	.3631	.68	8	17	A17
X	.08	-.28	.28	Y	-.10	.02	.24	-.30	.37	-.81	44	.4206	1.0724	.3004	-.81	10	18	A18
X	.08	-.28	.28	Z	.05	-.22	.26	-.06	.39	-.16	43	.8768	.4050	.5245	-.66	8	18	A18
X	-.24	1.21	.27	Y	.27	.38	.24	.83	.36	2.29	45	.0270	1.2932	.2555	.94	10	19	A19
X	-.24	1.21	.27	Z	-.09	.97	.26	.24	.38	.65	44	.5207	.0677	.7947	-.26	8	19	A19
X	.02	.62	.27	Y	-.09	.79	.24	-.17	.36	-.46	45	.6492	.0579	.8099	-.16	10	20	A20
X	.02	.62	.27	Z	.09	.51	.26	.12	.37	.31	44	.7581	.0148	.9032	.11	8	20	A20
X	.14	.77	.27	Y	-.27	1.44	.24	-.67	.36	-1.84	45	.0724	3.3402	.0676	-1.83	10	21	A21
X	.14	.77	.27	Z	.18	.70	.26	.06	.37	.17	44	.8638	.0954	.7575	-.26	8	21	A21
X	.06	-1.12	.34	Y	.00	-.98	.28	-.14	.44	-.32	44	.7522	1.4986	.2209	-1.10	10	22	A22
X	.06	-1.12	.34	Z	-.05	-.88	.29	-.24	.44	-.54	42	.5906	.0000	1.000	.00	8	22	A22
X	.01	-.62	.30	Y	-.03	-.54	.26	-.00	.40	-.19	44	.8464	2.2298	.1354	-1.18	10	23	A23
X	.01	-.62	.30	Z	.03	-.65	.27	.03	.41	.07	43	.9483	.8018	.3706	-.88	8	23	A23
X	.14	.77	.27	Y	-.27	1.44	.24	-.67	.36	-1.84	45	.0724	3.3402	.0676	-1.83	10	24	A24
X	.14	.77	.27	Z	.18	.70	.26	.06	.37	.17	44	.8638	.0954	.7575	-.26	8	24	A24
X	-.21	2.24	.27	Y	.22	1.55	.24	.69	.36	1.91	45	.0626	3.3402	.0676	1.83	10	25	A25
X	-.21	2.24	.27	Z	-.06	2.01	.25	.23	.37	.64	44	.5258	.0954	.7575	.26	8	25	A25

Figure 8. DIF Based on Class

The results of the DIF analysis based on class level indicate that respondents can be grouped into several categories based on their response patterns and measurement scores. Each respondent is classified into specific groups (e.g., A1, A2, A3, and so on) based on their position in the measurement model. The numerical values presented in the table represent factor scores or component values that reflect respondents' relative positions on the measured dimensions. These scores help explain differences between groups, where respondents with higher positive values tend to exhibit stronger characteristics related to emotional intelligence. Overall, the classification results demonstrate that the instrument is capable of differentiating respondents based on class-related characteristics. However, variations across groups should be interpreted carefully, as they may reflect genuine differences in ability rather than measurement bias.

The DIF analysis based on class indicates variation across class levels, which may reflect differences in students' academic experience, developmental maturity, and exposure to social learning situations. This does not necessarily indicate serious item bias, but it suggests that class level should be considered when interpreting emotional intelligence scores. Recent research on emotional intelligence and student adjustment shows that educational transitions and academic demands can influence students' emotional regulation, motivation, and social adaptation (Fernanda et al., 2024). Therefore, differences across class levels may represent meaningful developmental variation rather than measurement error. However, future studies should examine whether specific items consistently favor certain grade levels. This would help ensure that the instrument remains fair and applicable across different secondary school classes.

Effectiveness of Rating Scale

Rating scale analysis was conducted to evaluate whether the five response categories of the emotional intelligence scale functioned as intended and were interpreted consistently by respondents. This analysis is essential because the quality of category functioning directly influences the accuracy and validity of Rasch measurement. Several indicators, including average measures, Outfit Mean Square (MNSQ), threshold ordering, and category probability curves, were examined to determine whether each response category represented a distinct level of the underlying construct. Properly functioning response categories indicate that respondents can meaningfully differentiate between levels of agreement and use the rating scale in a consistent manner. Consequently, the results of this analysis provide evidence regarding the appropriateness of the five-point Likert scale for measuring emotional intelligence among secondary school students. The following figure presents the results of the rating scale analysis.

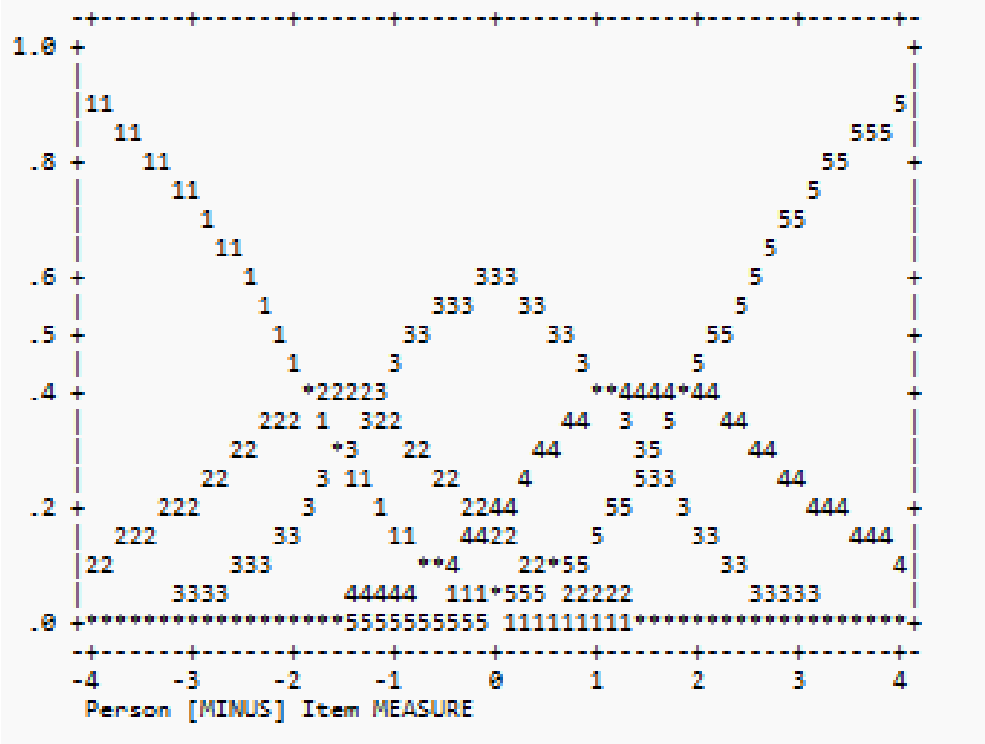


Figure 9. The Effectiveness of The Rating Scale Used in The Instrument

The evaluation of the five-point Likert scale reveals that the rating categories function effectively. The average measures increase monotonically from category 1 to category 5, indicating that respondents with higher levels of emotional intelligence tend to select higher response categories. The Outfit MNSQ values for each category fall within the acceptable range (0.5–1.5), confirming that no category deviates significantly from the model expectations. Additionally, the Andrich thresholds show a consistent and ordered progression, indicating that each response category is clearly distinguishable from the others. The category probability curves also demonstrate distinct peaks for each category, suggesting that

respondents are able to differentiate between response options effectively. These results confirm that the rating scale is appropriate and contributes to the validity of the instrument.

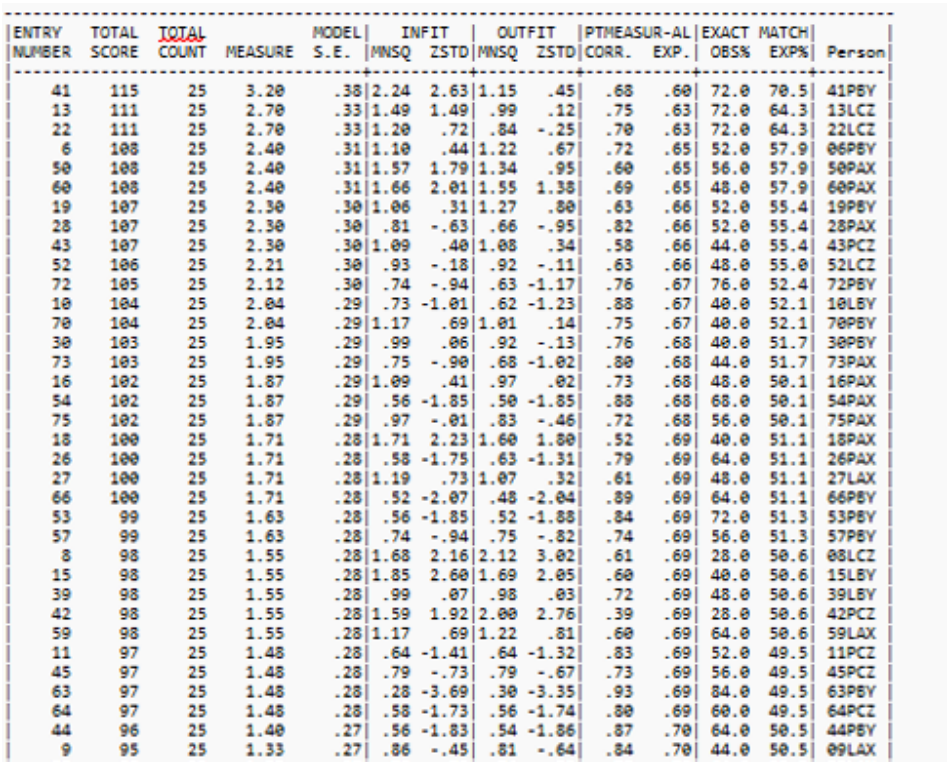
The five-point rating scale functioned effectively, as shown by the ordered category structure and clear progression of response categories. This indicates that students were able to distinguish between the available response options. In Rasch analysis, effective rating scale functioning is important because disordered categories may indicate confusion or overlap between response options. Recent instrument validation studies using Rasch modeling have emphasized that rating scale analysis helps determine whether Likert categories are meaningful and interpretable for respondents (Darman et al., 2024). The results of this study suggest that the five-category response format is appropriate for measuring emotional intelligence among secondary school students. Therefore, the rating scale can be retained in future use, although further testing with larger samples may provide stronger evidence of category stability.

Respondent Classification

Respondent classification analysis was conducted to examine how respondents were distributed across different levels of emotional intelligence and to evaluate the effectiveness of the instrument in distinguishing individual ability levels. Unlike item analysis, which focuses on the quality of the measurement instrument, respondent classification emphasizes the consistency and accuracy of participants' response patterns. Person measures indicate each respondent's estimated emotional intelligence level, while person fit statistics evaluate whether the observed responses conform to the expectations of the Rasch measurement model. In addition, exact match percentages provide evidence of the agreement between observed responses and model predictions, reflecting the overall quality of respondent classification. Collectively, these indicators provide important evidence regarding the precision of person measurement and the effectiveness of the instrument in differentiating students with varying levels of emotional intelligence. The following figure presents the respondent classification results based on the Rasch model analysis.

The classification of respondents based on Rasch analysis indicates a wide range of ability levels, with logit values ranging from approximately -3.00 to +3.00. This distribution indicates that the instrument successfully captured varying levels of emotional intelligence among secondary school students rather than clustering respondents within a narrow ability range. Respondents with higher logit values demonstrated stronger emotional competencies across the measured dimensions, whereas those with lower logit values exhibited comparatively weaker emotional abilities. Nevertheless, the majority of respondents were concentrated within the moderate logit range, suggesting that most students possessed average levels of emotional intelligence. This pattern is consistent with the characteristics of a relatively homogeneous school population, where respondents share similar educational experiences and developmental backgrounds. The observed spread of person measures also indicates that the instrument is capable of distinguishing

students across different ability levels. Consequently, the respondent classification supports the adequacy of the instrument for assessing emotional intelligence within the target population.



ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Person
41	115	25	3.20	.38	2.24	2.63	1.15	.45	.68	.60	72.0	70.5	41PBY
13	111	25	2.70	.33	1.49	1.49	.99	.12	.75	.63	72.0	64.3	13LCZ
22	111	25	2.70	.33	1.20	.72	.84	-.25	.70	.63	72.0	64.3	22LCZ
6	108	25	2.40	.31	1.10	.44	1.22	.67	.72	.65	52.0	57.9	06PBY
50	108	25	2.40	.31	1.57	1.79	1.34	.95	.60	.65	56.0	57.9	50PAX
60	108	25	2.40	.31	1.66	2.01	1.55	1.38	.69	.65	48.0	57.9	60PAX
19	107	25	2.30	.30	1.06	.31	1.27	.80	.63	.66	52.0	55.4	19PBY
28	107	25	2.30	.30	.81	-.63	.66	-.95	.82	.66	52.0	55.4	28PAX
43	107	25	2.30	.30	1.09	.40	1.08	.34	.58	.66	44.0	55.4	43PCZ
52	106	25	2.21	.30	.93	-.18	.92	-.11	.63	.66	48.0	55.0	52LCZ
72	105	25	2.12	.30	.74	-.94	.63	-1.17	.76	.67	76.0	52.4	72PBY
10	104	25	2.04	.29	.73	-1.01	.62	-1.23	.88	.67	40.0	52.1	10LBY
70	104	25	2.04	.29	1.17	.69	1.01	.14	.75	.67	40.0	52.1	70PBY
30	103	25	1.95	.29	.99	.06	.92	-.13	.76	.68	40.0	51.7	30PBY
73	103	25	1.95	.29	.75	-.90	.68	-1.02	.80	.68	44.0	51.7	73PAX
16	102	25	1.87	.29	1.09	.41	.97	.02	.73	.68	48.0	50.1	16PAX
54	102	25	1.87	.29	.56	-1.85	.50	-1.85	.88	.68	68.0	50.1	54PAX
75	102	25	1.87	.29	.97	-.01	.83	-.46	.72	.68	56.0	50.1	75PAX
18	100	25	1.71	.28	1.71	2.23	1.60	1.80	.52	.69	40.0	51.1	18PAX
26	100	25	1.71	.28	.58	-1.75	.63	-1.31	.79	.69	64.0	51.1	26PAX
27	100	25	1.71	.28	1.19	.73	1.07	.32	.61	.69	48.0	51.1	27LAX
66	100	25	1.71	.28	.52	-2.07	.48	-2.04	.89	.69	64.0	51.1	66PBY
53	99	25	1.63	.28	.56	-1.85	.52	-1.88	.84	.69	72.0	51.3	53PBY
57	99	25	1.63	.28	.74	-.94	.75	-.82	.74	.69	56.0	51.3	57PBY
8	98	25	1.55	.28	1.68	2.16	2.12	3.02	.61	.69	28.0	50.6	08LCZ
15	98	25	1.55	.28	1.85	2.60	1.69	2.05	.60	.69	40.0	50.6	15LBY
39	98	25	1.55	.28	.99	.07	.98	.03	.72	.69	48.0	50.6	39LBY
42	98	25	1.55	.28	1.59	1.92	2.00	2.76	.39	.69	28.0	50.6	42PCZ
59	98	25	1.55	.28	1.17	.69	1.22	.81	.60	.69	64.0	50.6	59LAX
11	97	25	1.48	.28	.64	-1.41	.64	-1.32	.83	.69	52.0	49.5	11PCZ
45	97	25	1.48	.28	.79	-.73	.79	-.67	.73	.69	56.0	49.5	45PCZ
63	97	25	1.48	.28	.28	-3.69	.30	-3.35	.93	.69	84.0	49.5	63PBY
64	97	25	1.48	.28	.58	-1.73	.56	-1.74	.80	.69	60.0	49.5	64PCZ
44	96	25	1.40	.27	.56	-1.83	.54	-1.86	.87	.70	64.0	50.5	44PBY
9	95	25	1.33	.27	.86	-.45	.81	-.64	.84	.70	44.0	50.5	09LAX

Figure 10. The Classification Of Respondents Based On The Rasch Model Analysis

The Infit and Outfit MNSQ values for most respondents fall within the acceptable range, indicating consistent response patterns. Furthermore, the predominantly positive point-measure correlations indicate that respondents' answers were positively associated with the underlying construct measured by the instrument. The exact match percentages, which generally ranged from 60% to 75%, also demonstrate satisfactory agreement between the observed responses and the responses predicted by the Rasch model. These findings suggest that respondents understood the questionnaire items appropriately and answered them in a logically consistent manner.

The respondent classification results show that students varied in their emotional intelligence levels, with most respondents located in the moderate category. This finding is consistent with recent studies indicating that emotional intelligence is associated with academic motivation, engagement, psychological well-being, and social support among students (Tessa & Suryana, 2023). The ability of the instrument to classify respondents into different levels is practically important because it allows teachers, counselors, and researchers to identify students who may need emotional development support. This finding is consistent with Fitriana et al. (2026), who reported that emotional intelligence is closely associated with students' peer interaction and social development in school settings. Recent studies also show

that emotional intelligence contributes to resilience and mental health, which strengthens the relevance of using such instruments in school-based assessment (Wang et al., 2025). Therefore, the classification results confirm that the developed scale is useful not only for research but also for educational intervention planning.

Despite the satisfactory psychometric properties demonstrated by the emotional intelligence scale, several limitations should be acknowledged. First, the empirical validation involved only 75 secondary school students. Although the Rasch model can be applied to relatively small samples, the present study involved only 75 respondents from a relatively homogeneous educational setting. This sample size is sufficient for preliminary instrument validation but may limit the stability of item calibration and person estimates. Consequently, the findings should be interpreted as initial psychometric evidence rather than definitive validation.

Second, the respondents were drawn from a relatively homogeneous educational context, which may limit the instrument's applicability to students from different schools, regions, or demographic backgrounds. Variations in educational environments and socio-cultural characteristics may influence students' responses to emotional intelligence items. Therefore, future studies are recommended to validate the instrument using larger and more diverse samples to strengthen the stability of Rasch estimates, examine the consistency of item functioning across populations, and improve the generalizability of the instrument.

4. Conclusion

This study aimed to develop and validate an emotional intelligence scale for secondary school students using a Rasch model approach. The findings demonstrate that the instrument has successfully achieved strong psychometric quality, as reflected in its overall reliability, item functioning, and measurement consistency. The items were well distributed across different levels of difficulty, enabling the instrument to capture a wide range of students' emotional intelligence. Most items functioned appropriately within the measurement model, indicating that they effectively represent the intended construct. In addition, the rating scale operated effectively, and the majority of respondents provided consistent responses, suggesting that the instrument is both stable and practical for use in educational settings.

Although a small number of items showed indications of misfit and potential bias, these issues do not undermine the overall quality of the instrument but rather provide opportunities for refinement and improvement. The instrument also demonstrated the ability to differentiate respondents based on their levels of emotional intelligence, confirming its usefulness as a diagnostic tool for educational and psychological purposes. In conclusion, this research has successfully developed a valid and reliable emotional intelligence scale that can be utilized by educators, school counselors, and researchers to assess students' emotional competencies. Future research is recommended to expand the sample size, refine misfitting items,

and test the instrument across diverse educational contexts to further strengthen its generalizability and applicability.

References

Apriliana, I. P. A., & Masi, L. M. (2025). Rasch Analysis of Students' Emotional Intelligence; Implications for Vocational Education. *KnE Social Sciences*, *10*(9), 15–25. https://doi.org/10.18502/kss.v10i9.18461

Apriyani, D., & Aini, F. Q. (2025). Design of Science Process Skills Instruments Reaction Rate Material: Rasch Model Approach. *Journal of Educational Sciences*, *9*(2), 876–885. https://doi.org/10.31258/jes.6.3.p.444-458

Bereded, D. G., Abebe, A. S., & Negasi, R. D. (2025). Emotional Intelligence and Academic Achievement Among First-Year Undergraduate University Students: The Mediating Role of Academic Engagement. *Frontiers in Education*, *10*, 1567418. https://doi.org/10.3389/feduc.2025.1567418

Caballero-García, P. Á., & Ruiz, S. S. (2025). Emotional Intelligence and Its Relationship with Subjective Well-Being and Academic Achievement in University Students. *Journal of Intelligence*, *13*(4), 42. https://doi.org/10.3390/jintelligence13040042

Chen, X., Aryadoust, V., & Zhang, W. (2025). A systematic review of differential item functioning in second language assessment. *Language Testing*, *42*(2), 193–222. https://doi.org/10.1177/02655322241290188

Darman, D. R., Suhandi, A., Kaniawati, I., Samsudin, A., & Wibowo, F. C. (2024). Development and Validation of Scientific Inquiry Literacy Instrument (SILI) Using Rasch Measurement Model. *Education Sciences*, *14*(3), 322. https://doi.org/10.3390/educsci14030322

Dewi, W. R., Maharani, Z., Alim, J. A., & Anggriani, M. D. (2025). Rasch Model Analysis of a Learning Autonomy Instrument in Digital Higher Education. *Saneskara: Journal of Social Studies*, *2*(2), 65–73. https://doi.org/10.62491/sjss.v2i2.2025.53

Febriyana, N., Henny Indrawati, & Makhdalena. (2023). The Influence of Emotional Intelligence, Industrial Work Practices, Soft Skills, and Self-Efficacy on Students Work Readiness. *Journal of Educational Sciences*, *7*(3), 444–458. https://doi.org/10.31258/jes.6.3.p.444-458

Fernanda, S., N. Rinita, Nisa Amelia, & Alviana Chintya. (2024). Emotional Intelligence and Adjustment Among First-Year University Students. *International Journal of Studies in International Education*, *1*(4), 97–106. https://doi.org/10.62951/ijsie.v1i4.356

Fernández-Berrocal, P., Vilca-Pareja, V., Rojas Zegarra, M. E., Hillpa-Zuñiga, M. E., Yana-Calla, V. R., & Cabello, R. (2025). Psychometric properties of the TMMS-24 emotional intelligence scale in Peruvian university students. *Frontiers in Psychology*, *16*, 1611923. https://doi.org/10.3389/fpsyg.2025.1611923

Fitriana, D., Herlina, Asriani, Rizal, & Purbarani, D. A. (2026). The Relationship Between Peer Interaction and Emotional Intelligence of Grade V Students at Elementary School Inpres 1 Birobuli. *Journal of Educational Sciences*, *10*(3). https://doi.org/10.31258/jes.10.3.p.1141-1149

Guo, S., Shi, L., & Zhai, X. (2024). Validating an Instrument for Teachers' Acceptance of Artificial Intelligence in Education. *Educational and Psychological Measurement*.

Kalsum, U., & Sirozi, M. (2025). The Role of Emotional Intelligence (Emotional Quotient/EQ) in Shaping Students' Personality From the Perspective of Educational Psychology. *Journal of Educational Sciences*, *9*(6), 876–885. https://doi.org/10.31258/jes.6.3.p.444-458

Hermansyah, E., Rizkyanti, C. A., & Hakim, L. N. (2024). Rasch Calibration and Differential Item Functioning (DIF) Analysis of the School Well-Being Scale for Students. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran*, *10*(1), 1. https://doi.org/10.33394/jk.v10i1.10759

Mohamed, M. G., Al-Yafeai, T. M., Adam, S., Hossain, M. M., Ravi, R. K., Jalo, F. M., & Osman, A. E. (2025). The significance of emotional intelligence in academic stress, resilience, and safe transition from high school to university: An SEM analysis among Northern Emirati university students. *Global Transitions*, *7*, 109–117. https://doi.org/10.1016/j.glt.2025.02.003

Nieto-Carracedo, A., Gómez-Iñiguez, C., Tamayo, L. A., & Igartua, J.-J. (2024). Emotional Intelligence and Academic Achievement Relationship: Emotional Well-Being, Motivation, and Learning Strategies as Mediating Factors. *Psicología Educativa*, *30*(2), 67–74. https://doi.org/10.5093/psed2024a7

Noorkholisoh, L., Yustiana, Y. R., Budiman, N., & Suryana, D. (2024). Validity and Reliability Analysis Using the Rasch Model in Developing Creativity Tests Instruments for Elementary School Students. *Jurnal Ilmiah Bimbingan Konseling Undiksha*, *15*(1), 128–135. https://doi.org/10.23887/jibk.v15i1.81249

Nurani, N. S., Rokhman, F., & Sarwi. (2025). Measuring The Unseen: Digitalization of Physics Diagnostic Assessment Based on Rasch Model. *Journal of Educational Sciences*, *9*(6), 4950–4963. https://doi.org/10.31258/jes.9.6.p.4950-4963

Peeters, M. J., & Augustine, J. M. (2023). Using Rasch measurement for instrument rating scale refinement. *Currents in Pharmacy Teaching and Learning*, *15*(1), 110–118. https://doi.org/10.1016/j.cptl.2023.02.015

Pertiwi, L. C., & Yudiarso, A. (2024). Validation of the Indonesian Version of the Partner's Emotional Dependency Scale (SED): A Rasch Model Approach. *Jurnal Psikologi; Vol 22, No 2 (2023): October 2023DO - 10.14710/Jp.22.2.102-117*. https://ejournal.undip.ac.id/index.php/psikologi/article/view/56227

Rahman, Md. H., Bin Amin, M., Yusof, M. F., Islam, Md. A., & Afrin, S. (2024). Influence of teachers' emotional intelligence on students' motivation for academic learning: An empirical study on university students of Bangladesh. *Cogent Education*, *11*(1), 2327752. https://doi.org/10.1080/2331186X.2024.2327752

Rogers, S. L., Brown, N., Goulding, M., Campbell, K., Mills, B., Hollett, R., Cruickshank, T., & Nosaka, K. (2025). Reliability and Validity of the Brief Emotional Experience Scale (BEES) as a Measure of Emotional Well-

Being. *Behavioral Sciences*, *15*(5), 643. https://doi.org/10.3390/bs15050643

Setyawarno, D., Maryati, & Natadiwijaya, I. F. (2025). Promoting a valid question model for measuring computational thinking skills based on confirmatory factor analysis and Rasch model. *Cogent Education*, *12*(1), 2505339. https://doi.org/10.1080/2331186X.2025.2505339

Taibolatov, K. M., Pfeyfer, N. E., Burdina, E. I., Kudysheva, A. A., & Bolatov, A. K. (2024). The Role of Emotional Intelligence on Academic Motivation of Schoolchildren. *Frontiers in Education*, *9*, 1265946. https://doi.org/10.3389/feduc.2024.1265946

Tessa, T., & Suryana, D. (2023). A Rasch Model Analysis: Validity of the College Student's Emotional Intelligence Instrument. *OPTIMA: Journal of Guidance and Counseling*, *3*(1).

Wang, T., Zhu, P., Fu, Y., & Choi, S. (2025). Exploring the Links Between Emotional Intelligence Resilience Support and Mental Health in College Students. *Scientific Reports*, *16*(1), 3609. https://doi.org/10.1038/s41598-025-33601-7

Xuan, Q. (2025). Emotional intelligence and holistic student development: An assessment of psychological and social efficacy in vocational university English education. *Frontiers in Psychology*, *16*, 1664645. https://doi.org/10.3389/fpsyg.2025.1664645

How to cite this article:

Nurihsani, I. S., Agustin, M., & Muqodas, I. (2026). Development and Rasch Validation of An Emotional Intelligence Scale for Secondary School Students. *Journal of Educational Sciences*, *10(7)*, 1195-1221.