IJOLAE Indonesian Journal on Learning and Advanced Education

http://journals.ums.ac.id/index.php/ijolae

Evaluating Students' Academic Resilience in Chemistry Learning: Insights from a Rasch Model Analysis

Desfi Annisa[1]✉, Hari Sutrisno[2], Endang Widjajanti Laksono[3], Sri Novita Yanda[4]
[1-3]Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta, Indonesia
[4]Department of Educational Research, Evaluation, and Assessment, Northern Illinois University, United States

DOI: 10.23917/ijolae.v6i3.23522
Received: February 21st, 2024. Revised: May 28th, 2024. Accepted: June 21st, 2024
Available Online: July 5th, 2024. Published Regularly: September, 2024

Abstract

Teaching chemistry is a complex subject that requires a certain level of knowledge and skills to understand. Students must overcome this challenge because teaching chemistry involves abstract ideas (atoms, molecules, and electrons) and principles, laws, reaction equations, and mathematical operations. Increasing academic resilience is important in enhancing students' understanding and well-being in learning. This research aimed to test the validity and reliability of students' chemistry academic resilience tests using the Rasch model. Data collection was conducted using Google Forms. Data analysis utilized the Rasch model, assisted by the Winstep application, to reveal various aspects of the assessment. Based on the research findings, a Cronbach's alpha coefficient of 0.78 indicates strong internal consistency. In addition, item reliability reached a significant value of 0.99, while person reliability of 0.78 confirmed the consistency of respondents in providing accurate answers in the assessed categories. Furthermore, the OUTFIT and INFIT MNSQ (persons) have average values of +1.02 and +1.01, respectively (the closer to 1, the better). In contrast, the INFIT and OUTFIT ZSTD values are 0.1 and 0.3, respectively (the closer to 0, the better). The most difficult question is coded ra11 with a logit score of 0.61, and the easiest is coded ra4 with a logit score of -0.05. Therefore, the student academic resilience instrument is an effective measuring tool. In the future, chemistry educators and researchers can benefit from the potential impact of this research on Indonesian education.

Keywords: academic resilience, chemistry learning, rasch model

✉***Corresponding Author:***
Desfi Annisa, Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta, Indonesia
Email: desfiiannisaa@gmail.com

1. Introduction

As students gain knowledge, their behaviour will improve, fostering positive changes in the teaching and learning process. However, it is challenging to disentangle learning from the educational process. Students must persevere in their academic endeavours, particularly in subjects like chemistry, which are often perceived as formidable. Despite its reputation for difficulty, chemistry holds significant importance in both middle and high school curricula and in many college degree programs (Senocak & Baloglu, 2014). Moreover, (Moreno et al., 2021) emphasize that mastering chemistry demands considerable effort and persistence. Additionally, (Zahro' & Ismono, 2021) highlight the sequential nature of chemistry and its reliance on computations as contributing factors to students' struggles with the subject. Weak achievement is often associated with a reliance on rote memorization rather than meaningful comprehension. According to (Oladejo et al., 2023), mastery of the subject leads to perceiving its concepts as straightforward.

In secondary schools, the issue of low achievement among chemistry students is crucial and potentially contributes to the challenges associated with grasping chemical concepts (Timilsena et al., 2022). Some students may exhibit overt signs of the effects of a stressful life, such as dread and unwillingness to complete tasks, psychological conflict due to the fear of stopping their studies, feelings of unreliability, and fear of the future, among others. These experiences create a lasting impression on the student's psyche, emphasizing the need for a high degree of conformity to appropriately adapt and deal with difficult life transitions by encouraging academic resilience (Ayasrah & Albalawi, 2022). The capacity to overcome obstacles in the classroom, increase motivation and concentration, enhance learning abilities, and adjust to shifting circumstances are all aspects of academic resilience (Suud et al., 2023). Students may have varying personal viewpoints during the learning process, which might alter their behavioural, affective, and cognitive reactions (Alkan & Yucel, 2019).

This goal can be achieved by improving students' learning efficiency and implementing logical processes in chemistry teaching. However, some students can overcome obstacles and difficult circumstances and achieve academic success because they believe learning comes from hard work and commitment, not luck. These students are known as academically resilient students. It is common for students to face difficulties and setbacks in class; if they can adapt and overcome these challenges, it means they have the capacity for academic resilience (Ulya & Gumiandari, 2023). In the educational context, academic resilience is the ability to overcome challenges to excel academically (Cassidy, 2016). Academic resilience is the capacity to deal with academic problems such as academic stress, failure, and boredom (Mahato et al., 2023). On the other hand, resilience describes a person's capacity to overcome obstacles and recover from setbacks (Lakhan et al., 2020). After failed attempts, resilience plays an important role in education and research in the teaching and learning process (Weidlich & Kalz, 2021), and academic resilience is thought to sustain student learning success (Annisa et al., 2023).

Academically gifted children who can maintain their motivation to succeed and excel in class, even when facing difficulties and other situations that could negatively impact their performance, are described as having academic resilience (Mahmudah et al., 2022). Academic success depends on student motivation. Students may only gain academic progress if they can handle challenges, study pressure, and school-related stress (Honra, 2022). If students have good academic resilience, threats can be turned into opportunities for them to grow and learn, fostering positive change (Masten et al., 2021). According to (Rao & Krishnamurthy, 2018), children can be taught academic resilience in the classroom. This is consistent with the explanation by (Rojas, 2015), which states that academic resilience is not merely a quality that a person inherently possesses but rather a dynamic process of adapting to difficulties. It requires actions and behaviours that each individual can adopt and shape.

Low academic resilience in students will hurt their social life, relationships with parents, instructors, and peers, and academic success (Beri & Kumar, 2018). Several elements, such as self-efficacy, self-control, low anxiety, persistence, and planning skills, influence academic resilience (Beale, 2020). According to (Romano et al., 2021), resilient adolescents appear to overcome stressful situations at school, maintain high motivation levels, and achieve extraordinary performance despite facing difficulties. Resilient students re-engage and persevere through challenging assignments. According to (Reeve et al., 2020), these characteristics predict many positive outcomes, including classroom engagement, enjoyment of school, and overall self-esteem. According to further research, adolescents who demonstrate academic resilience outperform their classmates in protection against severe maladaptation, academic achievement, and school context engagement (Allan et al., 2014; Fiorilli et al., 2020).

Finding appropriate methods to assess academic resilience in chemistry studies is a challenge. Research on academic resilience shows a strong correlation with individual academic achievement. Currently, there is a significant increase in global research investigating academic resilience (Kumalasari & Akmal, 2020). Research on mathematics resilience is still in its early stages in Indonesia. Although some studies have been conducted in academic institutions, primary and secondary educational settings, and other contexts, the number is limited, and additional studies are still needed to advance understanding in this area.

To provide accurate and objective values, evaluation tools used in educational research need improvement to become better-measuring tools (Muntazhimah, 2019). Research by (Ramdani et al., 2020) developed an Indonesian academic resilience scale with excellent validity and reliability to assess the academic resilience of junior high school students. The scale created can be applied as an evaluation tool when implementing interventions (Nicoll, 2014). To produce more comprehensive instrument information that better meets measurement standards, the junior high school student academic resilience scale will be validated using the Rasch Model (Briggs, 2019; Rachman & Napitupulu, 2017; Winarti & Mubarak, 2019). This measure of academic resilience is intended to be applicable and based on conditions in Indonesia. The Rasch model, popularized by Danish mathematician Georg Rasch in 1960, is a statistical modelling technique that creates equal interval measurement scales from raw data (Sumintono & Widhiarso, 2015). According to (Bond & Fox, 2007a), the Rasch model is a probabilistic response model where the probability of success is based on individual aptitude and item difficulty. By improving the analysis and informing psychometric scale assessments, Rasch's modelling draws on item response theory (Van Zile-Tamsen, 2017). The advantage of the Rasch Model is that it can identify model inaccuracies and produce more accurate estimates (Taufiq et al., 2021). Rasch Model analysis is not deterministic; it can identify the items being measured more precisely by employing a probabilistic approach to examine the measurement object (Indihadi et al., 2022). The interaction between respondents and question items can be described and measured using the Rasch model (Sumintono & Widhiarso, 2014). The validity of the instrument was tested using Rasch model analysis. Metrics such as unidimensionality, Wright map analysis, item analysis, participant ability analysis, and instrument analysis are used to assess the quality of the instrument in Rasch model analysis (Muslihin et al., 2022).

To enhance the quality of chemistry teaching and learning for students, the Rasch model is used to measure and validate students' cognitive abilities and the effectiveness of assessment techniques (Chow et al., 2018; Noben et al., 2021). Meanwhile, some teachers rarely conduct a needs analysis before designing learning tests. However, needs analysis plays a crucial role in learning planning, from designing the syllabus, teaching materials, and class activities to student assessment (Asti Ramadhani). The Rasch model analysis is the best choice for teachers in this situation. Besides interpreting students' knowledge through various Rasch data presentations, studying the structure of questions with varying levels of difficulty can assess students' overall understanding of chemistry. Through Rasch model analysis, students' characteristics and ability levels can be accurately mapped and used to enhance the quality of learning, particularly in the theoretical aspects of the subject matter (Ngadi, 2023). Rasch model analysis provides valuable information about the suitability of an item in measuring students' academic resilience in chemistry learning. By closely examining the results of Rasch model analysis, you can determine not only the reliability of the instrument created but also the patterns of students' responses (Sumintono, 2018; Chiang, 2015). The development of this academic resilience instrument will provide a more comprehensive and concrete picture of the measurement aspect of the test because the Rasch model involves two basic parameters: (a) student ability and (b) the level of difficulty of questions or the ability to ask questions (Khairani & Razak, 2015). Therefore, Rasch modelling is important in chemistry learning because it can be applied across various scientific disciplines and student grade levels. This research tests measurements of academic resilience based on Indonesian students' conditions, with few studies specifically focusing on chemistry learning. This encourages researchers to create measuring tools, such as instruments for assessing students' academic resilience in chemistry while maximizing their utilization during the learning process. Evaluation instruments in educational research must be developed into robust and high-quality measuring tools to yield objective and accurate results (Muntazhimah, 2019). As a result, teachers gain valuable information to enhance the quality of teaching and evaluate chemistry learning. This is supported by research (Cheung et al., 2011), which emphasizes the importance of conducting student diagnosis and learning processes as strategies for designing effective teaching and improving students' understanding of chemistry. Instrument quality measurements in this research include unidimensionality, Wright map analysis, item analysis, participant ability analysis, and instrument analysis (Planinic et al., 2019). According to (van de Grift et al., 2019), the Rasch model can be a useful tool for classroom assessment. Rasch is a technique for assessing student aptitude, and it can also be used to evaluate the quality of questions presented in a statistical data format (Almubarak et al., 2023).

The conceptual framework developed for this study, depicted in Figure 1, aims to provide more precise direction and guidance. To better align with the culture and characteristics of Indonesian students, the academic resilience questionnaire has been modified. This study intends to examine the validity and reliability of a new instrument designed to measure students' academic resilience in chemistry learning.

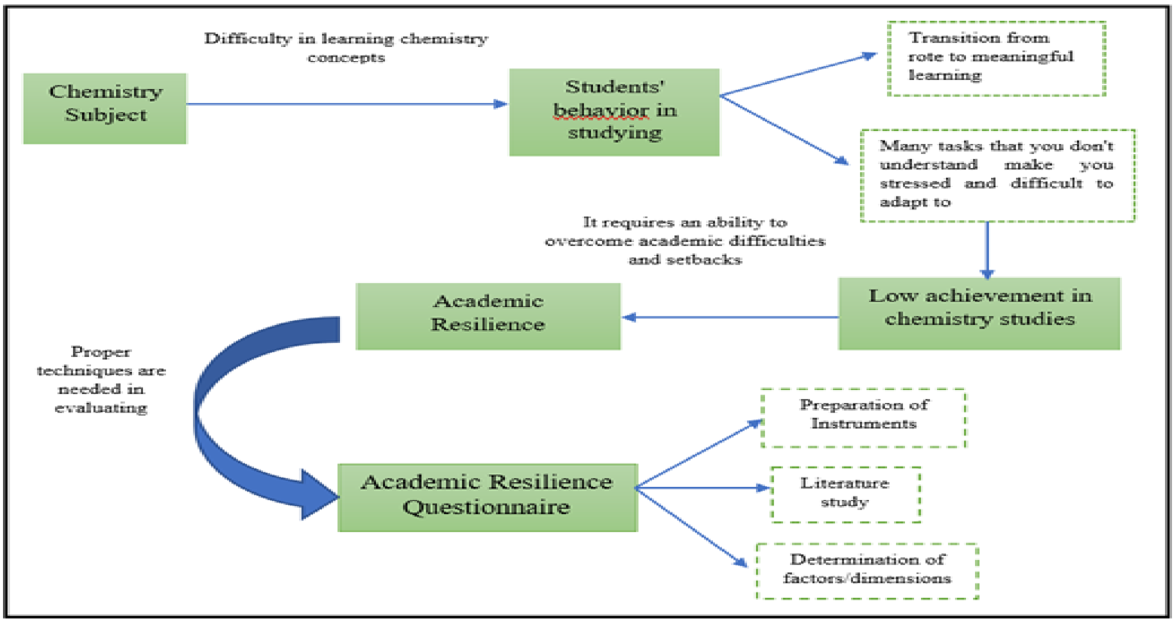

Figure 1. Academic Resilience Research Conceptual Framework

2. Method

This study employed a quantitative technique and a survey design. Through survey research, information from a sample of people is obtained, allowing the generalization of knowledge about the population (Gul, 2023). This study aims to examine the academic resilience instrument, which involves consultation with three experts (Cassidy, 2016; Idris et al., 2019; Ramirez-Granizo et al., 2020). The validation procedure for this academic resilience tool uses the Rasch Model. The Rasch Model can be used to (a) facilitate the development of instruments that produce valuable data, (b) produce data that can be used confidently for parametric and descriptive statistics, and (c) generate outcome measures that provide practitioners and researchers in school psychology with clinically relevant guidance (Boone & Noltemeyer, 2017).

The study's population consisted of tenth and eleventh-grade students from 13 different schools. The survey had 817 participants from Pekanbaru, Indonesia, selected using simple random sampling techniques. This method was employed to minimize data bias, as random sampling involves randomly selecting members of the population, ensuring every individual has an equal opportunity to be chosen. Random sampling is considered the most straightforward approach for data collection.

A Google Forms-based academic resilience questionnaire was used to collect data for this study. Google Forms provides more than just an assessment tool; it is a tried-and-true method for gathering data. Additionally, teachers and students can communicate via Google Forms (Rodriguez, 2018). Creating an online survey with Google Forms is quick and straightforward, and the results are stored in an online spreadsheet (Sivakumar, 2019). The questionnaire consists of three components comprising twelve statement items, specifically designed to cater to the requirements of Indonesian students enrolled in chemistry programs. The research instruments are accessible online, facilitating the data collection process. This grid of academic resilience instruments is intended for students gathering data related to chemistry. Table 1 illustrates it as follows.

Table 1. Academic Resilience Instrument

Variable	Dimension	No	Item
Academic Resilience	Spirituality	1	If I dedicate myself to God, I believe I will succeed and accomplish great things.
		2	I say a prayer before and after studying chemistry.
		3	I successfully turned down a friend's request to skip chemistry class.
		4	I don't thank God when I get high grades.
		5	I try not to freak out when I have a lot of challenging chemistry homework.
		6	I didn't grasp the lecture but completed the chemistry test independently.
	Emotion control	7	I accepted that I would receive low marks.
		8	When I earn good marks, I don't feel grateful.
		9	Instead of attending the chemistry lectures, I would instead visit the canteen.
	Self-reflection	10	I was confused when I received bad grades and was told to stay in my room.
		11	If there are too many deadlines for chemistry assignments, I get anxious.
		12	I came upon a friend's solution when a chemistry exam proved challenging.

Rasch model data analysis was performed using Winstep software, designed by (Linacre, 2011). Respondents and items interact simultaneously, according to the Rasch model. Instead of a raw score, the Rasch model displays a logit value representing the likelihood of selecting an item among respondents. This logit value is derived after applying the logarithmic function to the odds ratio value of the item, transforming the raw score into the item's logit value. According to (Sumintono & Widhiarso, 2014), The odds ratio value is a probability number that indicates the degree of respondents' agreement with an item compared to respondents who disagree. Our evaluation of the item becomes more objective when we use its logit value, as it converts the ordinal raw score into ratio data that meets all integer requirements.

The data used in the Rasch model is the probability score (P), or the comparison between the number of questions and the correct answers. The following formula is then used to translate the probability score into an odds ratio value:

$$Odds\ Ratio = \frac{P}{(1-P)}$$

Next, we use the following formula to obtain the logit value by applying the logarithm function (Sumintono & Widhiarso, 2014).

$$Logit = Log(\frac{P}{(1-P)})$$

This value is called the logit W-score or measurement value. The logit value has been scaled and can be used for various analyses.

Survey questions with ordinal or polytomous responses can also be calibrated using Rasch analysis. The Polychotomous Rasch Model is used in this type of problem. The Rating Scale Model (RSM) and the Partial Credit Model (PCM) constitute the Polychotomous Rasch Model. PCM, a refinement of Rasch analysis, is employed to analyze questions with polytomous or ordinal-type responses. Because PCM can handle questions with multiple categories of reactions, this method is highly relevant for assessing questions in education and other fields (Bond & Fox, 2007). Research using PCM for question analysis was conducted by (Wahyuningsih, 2021).

The PCM model is used to analyze cases involving polytomous responses based on Rasch's analysis. The following formula represents the fundamental Rasch analysis model:

$$p = P(X = 1)\frac{\exp(\theta - \delta)}{1 + \exp(\theta - \delta)}$$

The elements in the formula above can be explained as follows. For the PCM model, it has the following formula.

$$\Pr(X_{ni} = x = \frac{exp \sum_{k=0}^{x}(\theta_n - \delta_{ik})}{\sum_{h=0}^{m_i} \exp \sum_{k=0}^{h}(\theta_n - \delta_{ik})}$$

An explanation of each component in the PCM formula is as follows.

Pr (Xni=x) : probability of the nth participant with ability θ to answer to get a score x on the ith question
x : participant's score
θ_n : ability of the nth participant
δ_ik : level of difficulty for stage k in question item i (Wu et al., 2016).

A rating scale validity analysis test aims to determine whether respondents find the rating alternatives confusing. Using Likert ratings as an example, five options are provided ranging from STS (Strongly to Disagree Strongly), TS (Tend to Disagree), N (Neutral), S (Tend to Agree), to SS (Strongly Agree), all representing polytomous data. The instrument's ranking assumptions can be verified through Rasch model analysis. The Rating (Half Credit) scale test is utilized in the Winsteps program to measure rating scales. The observed average and Andrich thresholds, which indicate the precision of the options provided to respondents, are displayed as results.

Unidimensionality is a crucial metric for assessing whether a new instrument can capture the essence of the target, in this case, the academic resilience of chemistry students. Principal Component Analysis (PCA) of residuals is used in Rasch model analysis to determine how effectively various items assess what needs to be measured (Misbach & Sumintono, 2014).

Using various measurement models, including Infit, Outfit, Mean-Square Fit Statistics, and Standard Fit Statistics (ZSTD), the Rasch model is employed to examine academic resilience tests. According to Linacre (2002), two statistics determine whether data are suitable for the Rasch model: Outfit (outlier-sensitive or information-weighted fit) and Infit (inlier-sensitive or information-weighted fit). These statistics are typically presented in mean squared (MNSQ) and z-standardized (ZSTD) formats. According to Bond and Fox (2007), ZSTD is a standardized form of MNSQ that adjusts for sample size. MNSQ, on the other hand, represents the average of the squared residuals for an item.

Therefore, the Winsteps Rasch program output in the form of Outfit Mean Square (MNSQ) data needs to be examined to determine whether an item or respondent (person) fits or misfits the Rasch model in this research. Because it is not influenced by sample size, the MNSQ statistic was selected.

The model suitability for measurement can be evaluated using the following general guidelines, according to Linacre (2002): MNSQ > 2.0 indicates harm to the measurement system; 1.5 < MNSQ ≤ 2.0 indicates it has no significance for measurement; 0.5 ≤ MNSQ ≤ 1.5 indicates useful for measurement, and MNSQ < 0.5 means it is not helpful for measurement even though it does not harm the measurement system. Misfit items in a measurement tool should be avoided since they don't add much to the test score's dependability (Zubairi & Kassim, 2006). If an item is found to misfit the Rasch model, it may indicate several potential issues with its construction or validity. According to (Zubairi & Kassim, 2006), misfitting items could have problematic or flawed construction, such as low discriminatory power, meaning they fail to distinguish between different levels of the trait being measured effectively. Alternatively, misfitting items may have questionable validity, suggesting that they measure an ability or attribute unrelated to the assessed intended construct. Identifying and addressing misfitting items is crucial in ensuring the accuracy and reliability of measurement instruments, particularly in academic resilience assessments where precise and valid measurement is essential for meaningful educational research and practice (Reise, 1990).

ZSTD is also a t-test for a hypothesis, which inquires whether the data matches the model. As a result, the predicted value of z is almost equal to zero. The z value has a mean practically equal to 0 and a standard deviation of 1 when the observed data agrees with the model. Giant (z > +2) or low (z < -2) ZSTD values suggest that the item does not fit into the predicted model. The standardized z value (ZSTD) of outfit and infit can be negative or positive. ZSTD values that are negative signify less variation in the model. All participants could answer the question correctly, but all subjects with low ability answered it erroneously. This approach to the responses was close to the Guttman-style response string model. A positive value, on the other hand, denotes that answer variations exceed model variations. Reactions are erratic and unforeseen (Bond & Fox, 2015).

According to (Boone et al., 2014), the criteria applied are the same ones used to confirm that the question items are appropriate.

1. The value of the accepted outfit means square (MNSQ) is 0.5 < MNSQ < 1.5.
2. The Z-standard (ZSTD) outfit value that is accepted is -2.0 < ZSTD < +2.0.

This suggests that question items of poor quality should be reconsidered or replaced if they fail to meet these two criteria. The sample size significantly influences the applicability of these criteria rather than the consistency of item difficulty alone. Errors in answer keys, careless responses from a large number of participants, and questions with low discriminatory power can all contribute to an item's inadequacy. It's essential to recognize that the ZSTD value is heavily influenced by sample size; it tends to exceed three with larger samples (>500). Therefore, some experts caution against relying solely on the ZSTD criterion when the sample size is substantial (Sumintono & Widhiarso, 2015).

3. Result and Discussion

a. Unidimensional

The unidimensionality of a measurement determines its suitability for the intended purpose. A measuring tool is considered unidimensional when assessing a single ability or construct effectively (Baghaei, 2013; Tabatabaee-Yazdi et al., 2018). The Rasch model requires several critical assumptions for optimal analysis of a scale: additivity (ensuring consistent measurement intervals across the continuum), local independence (each item's response is independent of others), invariance (functioning consistently across different factors influencing the sample), and unidimensionality (focused on a single construct or latent trait). These assumptions are essential for accurately interpreting the results and ensuring the reliability and validity of the measurement tool in educational and psychological research (Bond & Fox, 2015).

In a unidimensional analysis, we examined Table 23 of the output from Winsteps version 3.73. According to (Boone et al., 2014), Unidimensionality can be confirmed if the variance explained by the measures is less than 20%, which is categorised as follows: adequate (20–40%), good (40–60%), and perfect (greater than 60%). Table 2 below illustrates the values that indicate unidimensionality.

Table 2. Unidimensional analysis

Table of standardized residual variance (in eigenvalue units)				
		Empirical		Modeled
Total raw variance in observations	17.4	100.0		100.0
Raw variance explained by measures	5.4	31.1		31.8
Raw variance explained by persons	1.5	8.6		8.8
Raw variance explained by items	3.9	22.5		23.0
Raw unexplained variance (total)	12.0	68.9	100.0	68.2
Unexplained variance in 1st contrast	2.2	12.5	18.2	
Unexplained variance in 2nd contrast	1.7	9.8	14.2	
Unexplained variance in 3rd contrast	1.4	8.1	11.7	
Unexplained variance in 4th contrast	1.1	6.6	9.6	
Unexplained variance in 5th contrast	1.0	6.0	8.7	

According to Table 2, the raw variance explained by the measurements is 31.1%, meeting the specified categorization criteria. Additionally, unidimensionality can be confirmed if there is less than 15% unexplained variance in the first to fifth residual contrasts. According to (Baghaei & Aryadoust, 2015), the unexplained variance values in the first remaining PCA construct are categorized as weak (>15%), moderate (10-15%), strong (5-10%), solid (3-5%), and extraordinary (less than 3%). In Table 2, the unexplained variance was found to be 12.5% in the first, 9.8% in the second, 8.1% in the third, 6.6% in the fourth, and 6.0% in the fifth contrast, all below 15%. Therefore, the instrument construct effectively assesses the overall academic resilience of students.

b. Wright Map Analysis (Person-Item Map)

Next, consider the difficulty level of each item. Rasch analysis with Winsteps offers several unique features, including a map that displays the distribution of both the subject's abilities and the difficulty of items on the same scale. This map, known as the Wright Map, sorts values according to size, typically ranging between -3 and +3 logits. However, a logit value above two or below -2 may be considered extreme. Guidelines classify these items into four groups based on their difficulty levels (Sumintono & Widhiarso, 2015), namely:

1. Measure value <-1 = very easy item
2. Measure value -1 to.d. 0 = easy items
3. Measure value 0 to 1 = difficult item
4. Measure value> 1 = very difficult item

The variable map table output indicates that students' academic resilience ranges from -2 to +3 logits. Question number 11 poses the most significant challenge for most students, whereas item number 4 is the easiest for all students. On average, students' abilities are at +1.2 logits, which falls between logits -2 and +3. This average suggests that students' abilities exceed the difficulty level of item 0.

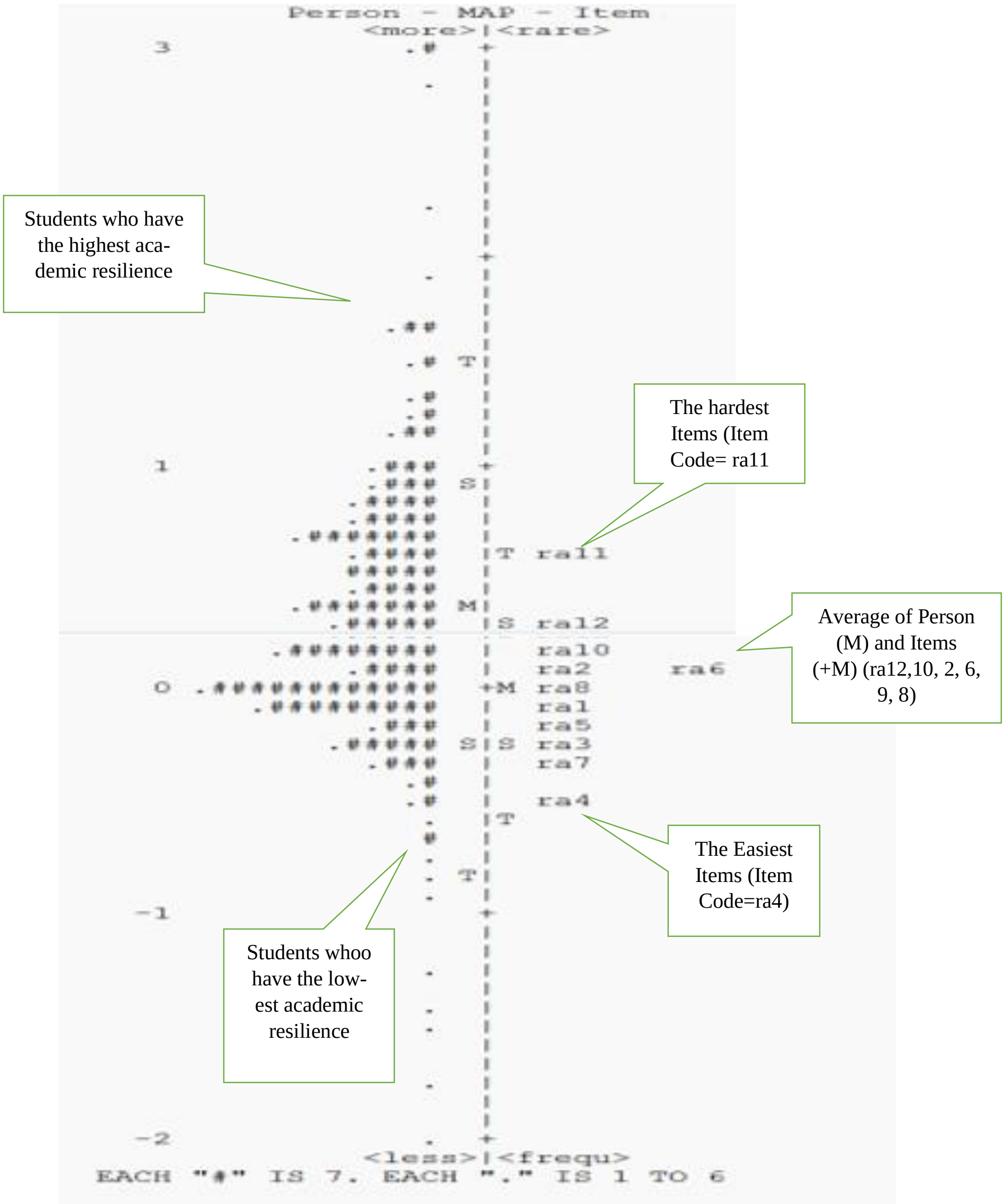

Figure 2. Wright-Map of Items and Persons on Academic Resilience; Rasch Measurement (N=817, I=12)

c. Analysis of Item

The items' difficulty appropriateness and bias were examined.

1) Item Difficulty Level

The Winstep application's 13-item measure order table can be used to determine an item's difficulty level. It is evident from this table that the standard deviation is 0.28. The item difficulty level can be divided into multiple categories based on the combination of the standard deviation and average logit values. These levels are extremely difficult (more significant than + 1 SD), challenging (0.0 logit + 1 SD), easy (0.0 logit – 1 SD), and very easy (less than -1 SD). For the highly challenging level, the limit value is more than 0.28. Level of difficulty: 0.0–0.28. Less than -0.28 is the elementary level, while the easy level is 0.0 - (-0.28). The degree of appropriateness of the items is determined by the logit value of every item in Table 13, ranging from the most challenging to the easiest. Item 11 is the only one categorized as extremely difficult. Three items, 4, 15, and 14, are demanding. It is known that items 12, 10, 2, 9, 6, and

8 are among the six that are classified as easy. Three items, 1, 5, and 3, are categorized as simple. Items 7 and 4 are the two that are categorized as really easy. Table 3 displays the findings for the items' difficulty level.

Table 3. Difficulty Category

Entry Number	Total Score	Measure	INFIT		OUTFIT		Pt. Measure		Exact Obs%	Match Exp%	Item
			MNSQ	ZSTD	MNSQ	ZSTD	Corr	Exp			
11	2159	0.61	1.20	4.2	1.21	4.3	0.40	0.54	25.8	30.1	ra11
12	2537	0.24	0.96	-0.9	1.01	0.2	0.44	0.52	39.5	31.5	ra12
10	2593	0.18	0.76	-6.0	0.79	-4.9	0.51	0.51	45.5	32.3	ra10
2	2674	0.10	0.80	-4.9	0.81	-4.4	0.57	0.51	40.6	32.4	ra2
9	2708	0.07	0.81	-4.7	0.81	-4.4	0.55	0.51	40.3	31.8	ra9
6	2733	0.05	0.95	-1.1	1.03	0.6	0.51	0.50	36.5	32.6	ra6
8	2741	0.04	1.18	4.0	1.20	4.0	0.43	0.50	35.6	32.6	ra8
1	2824	-0.05	1.16	3.5	1.16	3.1	0.55	0.50	28.1	32.5	ra1
5	2959	-0.18	1.16	3.5	1.10	1.9	0.53	0.49	31.6	32.1	ra5
3	3026	-0.26	1.19	3.9	1.14	2.7	0.49	0.48	34.0	32.5	ra3
7	3072	-0.31	0.92	-1.7	1.00	-0.1	0.51	0.48	37.3	33.4	ra7
4	3235	-0.50	1.08	1.6	1.00	0.1	0.49	0.46	32.4	34.3	ra4
Mean	2771.8	0.00	1.01	0.1	1.02	0.3			35.6	32.3	
SD	271.7	0.28	0.16	3.7	0.14	3.1			5.4	1.0	

2) Item Conformity Level

The Item Fit Order (Table 10 in Winstep) provides columns for various metrics such as measurement point correlation (PT MEASURE CORR), OUTFIT mean square (MNSQ), and OUTFIT Z standard (ZSTD), which are used to evaluate item fit. These metrics help educators and researchers avoid misconceptions about the subject being measured. According to guidelines (Ariffin et al., 2010), a PT MEASURE CORR value above 3.8 or an index ranging from 0.20 to 0.79 (Linacre, 2009), indicates that an item effectively differentiates abilities among respondents. A negative or zero value suggests that the responses are not aligned with the measured variable or construct (Linacre, 2009). Find out which items are not suitable it can be seen from the value of means- square (MNSQ). According to (Linacre, 2009), the MNSQ value indicates how accurate an item is when compared to the ideal item.

According to (Boone et al., 2014), the criteria used to assess the appropriateness of items include the values of mean-square fit, z-standard fit, and point-measure correlation. If an item does not meet these criteria, it is advisable to repair or replace it.

According to (Linacre, 2009), the MNSQ value was found to be between 0.5 and 1.5. Additionally, it can be determined by summing the mean value of the squared infit (mean MNSQ Infit). For instruments using a Likert scale, MNSQ values between 0.4 and 1.4 indicate an appropriate fit (Sumintono & Widhiarso, 2015).

Table 4. The Level of Suitability Items

Entry Number	Total Score	Measure	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	Pt. Measure Corr	Exp	Exact Obs%	Match Exp%
11	2159	0.61	1.20	4.2	1.21	4.3	0.40	0.54	25.8	30.1
8	2741	0.04	1.18	4.0	1.20	4.0	0.43	0.50	35.6	32.6
3	3026	-0.26	1.19	3.9	1.14	2.7	0.49	0.48	34.0	32.5
5	2959	-0.18	1.16	3.5	1.10	1.9	0.53	0.49	31.6	32.1
1	2824	-0.05	1.16	3.5	1.16	3.1	0.55	0.50	28.1	32.5
4	3235	-0.50	1.08	1.6	1.00	0.1	0.49	0.46	32.4	34.3
6	2733	0.05	0.95	-1.1	1.03	0.6	0.51	0.50	36.5	32.6
12	2537	0.24	0.96	-0.9	1.01	0.2	0.44	0.52	39.5	31.5
7	3071	-0.31	0.92	-1.7	1.00	-0.1	0.51	0.48	37.3	33.4
9	2708	0.07	0.81	-4.7	0.81	-4.4	0.55	0.51	40.3	31.8
2	2674	0.10	0.80	-4.9	0,81	-4.4	0.57	0.51	40.6	32.4
10	2593	0.18	0.76	-6.0	0.79	-4.9	0.51	0.51	45.5	32.3
Mean	2771.8	0.00	1.01	0.1	1.02	0.3			35.6	32.3
SD	271.7	0.28	0.16	3.7	0.14	3.1			5.4	1.0

The first requirement is that there should be no inconsistencies in any item. However, among the second criterion's seven entries, 11, 8, 3, 1, 9, and 10 do not meet the second threshold. Furthermore, the outcome is less than 0.85 even if the twelve components in the third criterion, as determined by point measure correlation, are above 0.4.

3) Diagnostic Rating Scale

Table 5 in the Winsteps application offers interpretations for diagnostic rating scales. This diagnostic aims to determine respondents' understanding of variations in response options related to academic resilience characteristics 1, 2, 3, 4, and 5. If the observed average and Andrich threshold values increase proportionally with the respondents' level, then respondents comprehend the discrepancies in answers. The table provides detailed Andrich threshold values for each level.

Table 5. Rating Scale Diagnostic

Category Label	Observed Count	%	Observed Sample Average	Expect	INFIT MNSQ	OUTFIT MNSQ	Andrich Threshold	Category Measure
1	344	14	-0.15	-0.26	1.22	1.28	NONE	(-1.81)
2	332	14	-0.14	-0.06	0.84	0.79	-0.16	-0.67
3	574	24	0.12	0.16	0.86	0.77	-0.50	0.01
4	411	17	0.36	0.41	1.00	0.93	0.62	0.68
5	763	31	0.77	0.72	0.94	0.97	0.04	(1.77)

Table 4 displays the appropriateness and differentiation among levels 1, 2, 3, 4, and 5. The findings of the analysis outline the levels of the academic resilience instrument in terms of actual student behavioural contexts. Therefore, the academic resilience assessment recommends a five-level answer option scale.

d. Item Bias Detection

Differential Item Functioning (DIF) aims to determine whether the provided items show bias toward respondents based on their gender. Bias should not influence the measurement process of a good item. This bias is assessed through DIF analysis using the Rasch model.

Items with a probability value less than 5% (0.05) can be identified as biased (Sumintono & Widhiarso, 2015). Several authors also employed a chi-square test to find DIF (Le, 2009). In this investigation, only gender bias is apparent. The study on gender bias revealed that several questions ($p < 0.05$) exhibited prejudice, precisely questions 4, 7, and 11.

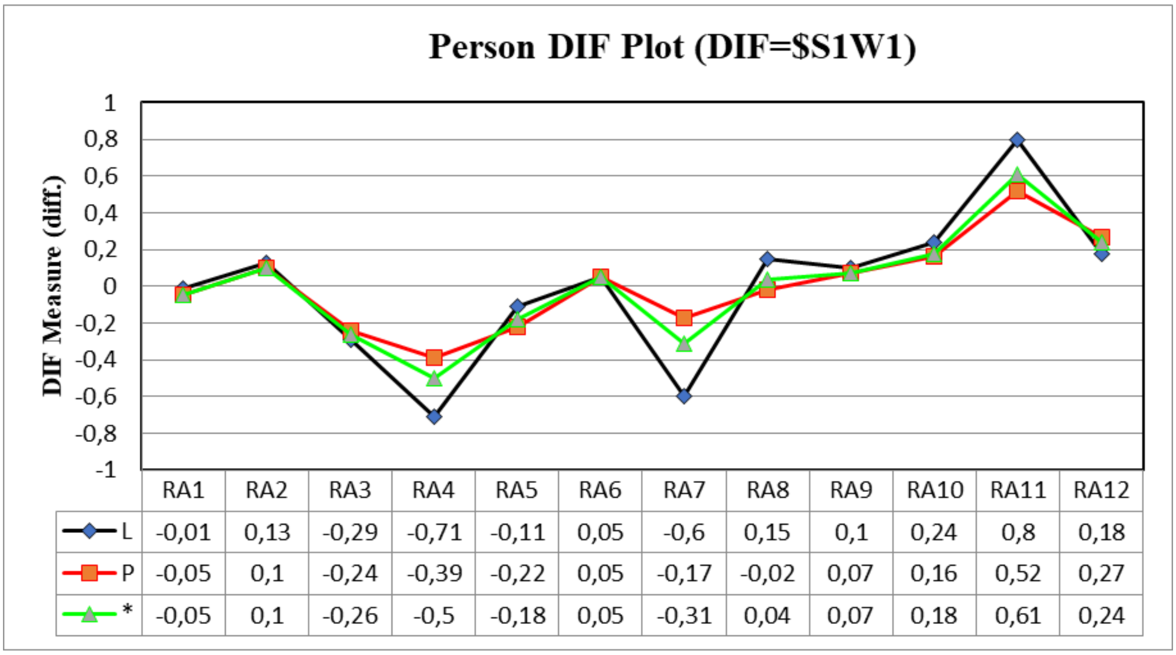

Figure 3. Item Logit Position Based on Gender

Based on this image, male students find question number four easier to answer than their female counterparts, suggesting a more pronounced positive impact on male student.

e. Analysis Instruments

Instrument analysis often utilizes statistical summary tables. Table 5 presents the instrument analysis.

Table 5. Instruments Analysis

	Total Score	Count	Measure	Model S. E	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
Mean	40.7	12.0	0.36	0.30	1.02	-3	1.02	-3
SD	8.5	0.0	0.79	0.22	0.65	2.0	0.66	2.0
Max	60.0	12.0	3.91	1.78	3.16	4.4	4.77	4.5
Min	12.0	12.0	-4.0	0.24	0.06	-4.7	0.07	-4.7
Real RMSE	0.40							
Model RMSEA	0.38	True SD	0.68	Separation	1.71	Person reliability		0.71
S.E. of Person Mean = 0.07	0.03	True S	0.70	Separation	1.86	Person reliability		0.78

Person raw score-to-measure correlation = 0.92
Cronbach's alpha = 20 Person Raw Score (test) reliability = 0.78

Summary Item	Total Score	Count	Measure	Model S. E	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
Mean	2771.8	817.0	0.00	0.03	1.01	0.1	1.02	0.3

Summary Item	Total Score	Count	Measure	Model S.E	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
SD	271.7	0.0	0.28	0.00	0.16	3.7	0.14	3.1
Max	3235.0	817.0	0.61	0.03	1.20	4.2	1.21	4.3
Min	2159.0	817.0	-0.50	0.03	0.76	-6.0	0.79	-4.9
Real RMSE	0.03							
Model RMSEA	0.03	True SD	0.28	Separation	8.25	Item reliability		0.99
S.E. of Person Mean = 0.07	0.08	True SD	0.28	Separation	8.59	Item reliability		0.99

The person measure in the student academic resilience data disclosure tool displays the average score for each participant. A participant's abilities exceed the instrument item difficulties when their average score exceeds the item mean (0.00 logits in this example).

The Cronbach's Alpha for this interaction stands at 0.78, falling within the desirable range. *Cronbach's Alpha* is a measure used to assess the consistency between items and individuals. Additionally, among respondents in the exceptional group, the Person Reliability score of 0.71 indicates consistent responses. Item reliability, measured at 0.99, evaluates the quality of the instrument's items in the interim.

Data from the Person and Item tables include MNSQ INFIT and MNSQ OUTFIT. According to the Person Table, the average MNSQ INFIT and MNSQ OUTFIT values are 1.02. The MNSQ OUTFIT is recorded at 1.02, while the MNSQ INFIT is at 1.01. Since the ideal value is 1, approaching this number is preferred, indicating that an average person or object fits reasonably well. Additionally, both the person's OUTFIT ZSTD and INFIT ZSTD are at -3. Furthermore, the items' ZSTD INFIT value is 0.1, and the ZSTD OUTFIT value is 0.3. Ideally, ZSTD should approach zero, indicating a good fit for the person or item.

The last aspect concerns the separation of people and items. Through their academic resilience tool, students can evaluate their ability to demonstrate capabilities across various abilities based on individual separation using various items. The instrument is most effective when it provides various individual abilities. Item separation illustrates how well a large sample is spread along a linear interval scale. Additionally, this index defines the precision of the measured construct; higher separation indicates greater precision.

According to Table 5, the separation value for people is 1.71, and for items, it is 8.25. A more considerable separation value indicates a better distinction between individuals and items. The separation value is calculated using the following formula:

$$H = [\frac{(4\ x\ separation) + 1}{3}]$$

$$Hstudents = \left[\frac{(4\ x\ 1.71) + 1}{3}\right] = 2.613$$

$$Hinstrumen = \left[\frac{(4\ x\ 8.25) + 1}{3}\right] = 11.33$$

The separation value for items is 11, while for individuals, it is 2.61, rounded to three. Participants in the study were grouped into three categories based on their varying skill levels. These categories were formed according to the difficulty level of the items, ranging from easiest to most challenging.

This study aims to apply the Rasch model to examine the validity and reliability of students' academic resilience measures. Using the Rasch model, academic resilience measures completed by students in their chemistry classes were evaluated according to several criteria, including WrightMap analysis, unidimensionality, item analysis (including item difficulty, item appropriateness, and item bias), and instrument analysis.

According to the results of the unidimensionality analysis, the raw variance measurement explained by the measurements is 31.1%, which meets the satisfactory criteria. This demonstrates that the study meets the minimum criterion of unidimensionality of 20% (Sumintono & Widhiarso, 2015).

The unexplained variance values for all first through fifth contrasts in the results are less than 15%, indicating that every instrument item can effectively measure every aspect of academic resilience. This demonstrates that the academic resilience instrument can assess its intended target's ability to demonstrate academic resilience as expected.

The item analysis results reveal that complicated, challenging, easy, and easy categories correspond to varying difficulty levels for items within the academic resilience instrument.

Table 5 presents the results, showing that item reliability is at 0.99, categorized as excellent, and person reliability is at 0.78, categorized as good. This indicates that the questions were well-crafted and respondents' responses were consistent. According to (Hinton et al., 2014), a reliability coefficient is consistently in the high category when reliability scores exceed 0.90. This indicates that every item currently used is reliable for measuring the construct (Herwin & Nurhayati, 2021). According to (Fitri, 2017), Consistency in this context refers to administering the same inquiry to the same individual at different points in time to yield consistent results. Therefore, this academic resilience tool provides reliable measurements for students.

According to Table 2, the item that respondents find most challenging to agree with in the instrument is statement item 11, which has a measured value of +0.61. Statement 6 indicates signs that the Chemistry assignment deadline is approaching too quickly. Naturally, in this instance, the teacher's deadline for completing chemical assignments is an indicator, as students are expected to meet it. However, students may experience stress due to the overwhelming workload and time constraints of learning assignments (Livana et al., 2020). This could be one of the causes of students' anxiety as an assignment deadline approaches. Another reason is that students struggle to understand the concepts in chemistry. Many students find chemistry difficult and uninteresting, which contributes to it being one of the scientific subjects they dislike (Sariati et al., 2020). Students often struggle with learning, have difficulty making connections between ideas, and find it challenging to apply their linguistic, mathematical, and reasoning skills when studying chemistry (Zakiyah et al., 2018; Hermita et al, 2021). Students generally find it challenging to agree with statement number 11, especially those who struggle with answering chemistry questions, indicated by a rating of -0.50. Statement number 4, with a rating of +0.50, is the easiest for respondents to agree with. The fourth statement, "When I score well in chemistry, I do not give thanks to God," implies that students believe they always appreciate their success. Being grateful enhances emotional well-being, increasing happiness and contentment with life. Additionally, gratitude can empower individuals, helping them feel more in control of themselves and their surroundings.

Resilience is characterized by cognitive flexibility, self-control, and adaptive coping. Grateful individuals can adapt to their circumstances, manage their environments, leverage positive relationships, and transform negative emotions into positive ones. This ability often helps them overcome challenges they encounter (Lin et al., 2015).

The activities listed above are arranged from the highest to lowest item level. A measurement value <-1 categorizes items in the easy group, following the categorization suggested by (Sumintono & Widhiarso, 2015), which is used to calculate the applicable measurement value. It can be stated that measurement values between -1 and 0 indicate easy items, values between 0 and 1 are challenging, and values greater than 1 are complex. Numerous data support this, confirming that the optimal instrument difficulty level ranges between 2 and 2 (Tjabolo & Otaya, 2019). Although some questions are deemed problematic, overall, the items' difficulty level satisfies the requirements (Angriani et al., 2018).

Each item within the academic resilience instrument meets the criteria for at least one of the three appropriateness standards, ensuring that all items effectively measure academic resilience. Moreover, the diagnostic rating scale analysis results indicate that participants understand the distinctions between alternative responses 1, 2, 3, 4, and 5, as evidenced by the increasing Andrich threshold values and observed averages from smallest to largest. The analysis of the instrument's results yielded a Cronbach's alpha value of 0.78, indicating good reliability. Additionally, item reliability falls within the excellent category with a value of 0.99, while the analysis of person reliability demonstrates a similarly high level of reliability. Accordingly, (Linacre, 2009) endorses that the reliability value of items from 0.67 to 0.80 is categorized as simple, 0.81 to 0.90 is classified as good, and> 0.91 is very good. Meanwhile, the reliability index of items and respondents received was ≥ 0.8 (Bond & Fox, 2007).

The statistical findings indicate that Rasch model analysis provides abundant information about an item's ability to measure students' overall academic resilience. By examining the results of the Rasch model analysis more closely, one can uncover the instrument's precision and patterns in students' responses. This insight is crucial for initiatives to adapt and refine educational activities by considering students' knowledge, experiences, strengths, and weaknesses. It also helps teachers organize study groups to improve academic outcomes in the teaching process.

Evaluation results are considered valuable information for teachers because they enable accurate conceptions about the state and conditions of the educational system. Studies (Widana et al., 2018) have affirmed that assessments that enhance teaching quality provide precise and reliable data, analysis, and interpretations. This allows an accurate understanding of the educational system's status and circumstances. According to (Angel et al., 2019), Diagnostic evaluation aims to investigate students' actual circumstances about their academic performance. It is a process that involves making relevant judgments to achieve efficiency in the educational field. Aside from that, (Ngadi, 2023) said that the Rasch model analysis of knowledge test results might effectively map students' traits and skill levels, which can then be utilized to enhance the quality of learning, particularly in general knowledge/theory topics.

4. Conclusion

Validity and reliability are critical considerations when developing new research tools. Using the Rasch model to analyze academic resilience instruments ensures accurate, impartial, and consistent data by capturing interactions between respondents and statement items. The study's findings confirm the validity and reliability of the developed academic resilience instrument. The instrument demonstrates robust performance with an item reliability value of

0.99, indicating exceptional quality, and a person reliability value of 0.78, reflecting good consistency among respondents' responses. Specifically designed to assess students' academic resilience in chemistry education, this novel resilience scale shows promising utility. However, further research is recommended, significantly if extending its application beyond the Indonesian context or into higher education.

5. References

Alkan, F., & Yucel, A. S. (2019). An analysis of achievement goal orientations and chemistry motivations of students within the structural equation modeling and an evaluation of the influencing factors. *New Trends and Issues Proceedings on Humanities and Social Sciences*, *6*(2), 01–08. https://doi.org/10.18844/prosoc.v6i2.4276.

Allan, J. F., McKenna, J., & Dominey, S. (2014). Degrees of resilience: profiling psychological resilience and prospective academic achievement in university inductees. *British Journal of Guidance and Counselling*, *42*(1), 9–25. https://doi.org/10.1080/03069885.2013.793784

Almubarak, A., Saadi, P., Prayogi, R., & Maldini, P. P. (2023). Assessing Students' Understanding of Chemical Bonds Material by Rasch Modeling. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *5*(3), 217–232. https://doi.org/10.23917/ijolae.v5i3.22242

Angel, M., Cevallos, S., Alexander, C., Rosado, Z., Viviana, O., & Terán, T. (2019). *The Procedure Used on Diagnostic Evaluation Process*. *3*, 1–10.

Angriani, A. D., Nursalam, N., Fuadah, N., & Baharuddin, B. (2018). Pengembangan Instrumen Tes Untuk Mengukur Kemampuan Pemecahan Masalah Matematika Siswa. *AULADUNA: Jurnal Pendidikan Dasar Islam*, *5*(2), 211. https://doi.org/10.24252/auladuna.v5i2a9.2018

Annisa, D., Sutrisno, H., & Widjajanti, E. (2023). Academic Resilience in Chemistry during Covid-19: A Systematic Review. *Jurnal Penelitian Pendidikan IPA*, *9*(11), 1111–1119. https://doi.org/10.29303/jppipa.v9i11.4354

Ariffin, S. R., Omar, B., Isa, A., & Sharif, S. (2010). Validity and reliability multiple intelligent item using rasch measurement model. Procedia Social and Behavioral Sciences, 9, 729–733. https://doi.org/10.1016/j.sbspro.2010.12.225

Baghaei, P. (2013). Development and psychometric evaluation of a multidimensional scale of willingness to communicate in a foreign language. *European Journal of Psychology of Education*, *28*(3), 1087–1103. https://doi.org/10.1007/s10212-012-0157-y

Baghaei, P., & Aryadoust, V. (2015). Modeling Local Item Dependence Due to Common Test Format With a Multidimensional Rasch Model. *International Journal of Testing*, *15*(1), 71–87. https://doi.org/10.1080/15305058.2014.941108

Beale, J. (2020). Academic Resilience and its Importance in Education after Covid-19. *Eton Journal for Innovation and Research in Education*, *4*, 1–6.

Beri, N., & Kumar, D. (2018). Predictors of Academic Resilience Among Students: a Meta-Analysis. *I-Manager's Journal on Educational Psychology*, *11*(4), 37. https://doi.org/10.26634/jpsy.11.4.14220

Bond, T. G, Fox, C. M. (2015). Applying the rasch model: Fundamental measurement in the human sciences:Third Edition. In *Andrew's Disease of the Skin Clinical Dermatology*. Routledge Taylor & Francis Group.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model : Fundamental*

Measurement. Lawrence Erlbaum Associates

Boone, W. J., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education*, *4*(1), 1–13. https://doi.org/10.1080/2331186X.2017.1416898

Boone, W. J., Yale, M. S., & Staver, J. R. (2014). Rasch analysis in the human sciences. In *Rasch Analysis in the Human Sciences*. https://doi.org/10.1007/978-94-007-6857-4

Briggs, D. C. (2019). Interpreting and visualizing the unit of measurement in the Rasch Model. *Measurement: Journal of the International Measurement Confederation*, *146*(July), 961–971. https://doi.org/10.1016/j.measurement.2019.07.035

Cassidy, S. (2016). The Academic Resilience Scale (ARS-30): A new multidimensional construct measure. *Frontiers in Psychology*, *7*(NOV), 1–11. https://doi.org/10.3389/fpsyg.2016.01787

Cheung, F. M., van de Vijver, F. J. R., & Leong, F. T. L. (2011). Toward a new approach to the study of personality in culture. *American Psychologist*, *66*(7), 593–603. https://doi.org/10.1037/a0022389

Chiang, W.-W. (2015). Ninth Grade Student' Self-assessment in Science: A Rasch Analysis Approach. *Procedia - Social and Behavioral Sciences*, *176*, 200–210. https://doi.org/10.1016/j.sbspro.2015.01.462

Chow, J., Tse, A., & Armatas, C. (2018). Comparing trained and untrained teachers on their use of LMS tools using the Rasch analysis. *Computers and Education*, *123*, 124–137. https://doi.org/10.1016/j.compedu.2018.04.009

Fiorilli, C., Farina, E., Buonomo, I., Costa, S., Romano, L., Larcan, R., & Petrides, K. V. (2020). Trait emotional intelligence and school burnout: The mediating role of resilience and academic anxiety in high school. *International Journal of Environmental Research and Public Health*, *17*(9). https://doi.org/10.3390/ijerph17093058

Fitri. (2017). Analisis Validitas dan Reliabilitas Instrumen Kinerja Akuntan Menggunakan Pendekatan Rasch Model. *Jurnal Ilmiah Akuntansi Peradaban*, *3*(1), 34–45.

Gul, Y. E. (2023). A Theoretical Perspective On Survey Method From Quantitative Research Methods. *Universum*, *4*(106),64-68. https://doi.org/10.32743/UniPsy.2023.106.4.15254.

Hermita, N., Putra, Z., Alim, J., Wijaya, T., Anggoro, S., & Diniya, D. (2021). Elementary Teachers' Perceptions on Genially Learning Media Using Item Response Theory (IRT). *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *4*(1), 1-20. doi:https://doi.org/10.23917/ijolae.v4i1.14757

Herwin, & Nurhayati, R. (2021). Measuring students' curiosity character using confirmatory factor analysis. *European Journal of Educational Research*, *10*(2), 773–783. https://doi.org/10.12973/EU-JER.10.2.773

Hinton, P., McMurray, I., & Brownlow, C. (2014). SPSS Explained. In *SPSS Explained*. https://doi.org/10.4324/9781315797298

Honra, J. R. (2022). Development and Validation of Filipino Learners' Academic Resilience Scale (FLARS). *American Journal of Education and Technology*, *1*(2), 107–113. https://doi.org/10.54536/ajet.v1i2.664

Ibrahem Ayasrah, J., & Nawaf Albalawi, K. (2022). Academic resilience and its relationship with academic achievement of the first-year students of university. *Journal of Positive School Psychology*, *6*(11), 2647–2666.

Idris, I., Khairani, A. Z., & Shamsuddin, H. (2019). The influence of resilience on psychological well-being of Malaysian University undergraduates. *International Journal of Higher Education*, *8*(4), 153–163. https://doi.org/10.5430/ijhe.v8n4p153

Indihadi, D., Suryana, D., & Ahmad, A. B. (2022). the Analysis of Construct Validity of Indonesian Creativity Scale Using Rasch Model. *Creativity Studies*, *15*(2), 560–576. https://doi.org/10.3846/cs.2022.15182

Kumalasari, D., & Akmal, S. Z. (2020). Resiliensi akademik dan kepuasan belajar daring di masa pandemi COVID-19: Peran mediasi kesiapan belajar daring. *Persona:Jurnal Psikologi Indonesia*, *9*(2), 353–368. https://doi.org/10.30996/persona.v9i2.4139

Lakhan, G. R., Ullah, M., Channa, A., Siddique, M., & Gul, S. (2020). *The Effect Of Academic Resilience And Attitude On Managerial Performance*. *19*(3), 3326–3340. http://ilkogretim-online.org

Le, L. T. (2009). Investigating Gender Differential Item Functioning Across Countries and Test Languages for PISA Science Items. *International Journal of Testing*, *9*(2), 122–133. https://doi.org/10.1080/15305050902880769

Lin, C., Education, T., & District, G. S. (2015). Impact Of Gratitude On Resource Development And Emotional Well-Being. *Social Behavior And Personality*, *43*(3), 493–504. http://dx.doi.org/10.2224/sbp.2015.43.3.493

Linacre, J.M. (2002). What do infit and outfit mean-square and standardized mean? Rasch Measurement Transaction, 16, 878.

Linacre, J. M. (2009). Local independence and residual covariance: A study of olympic figure skating ratings. *Journal of Applied Measurement*, *10*(2), 157–169.

Linacre, J. M. (2011). *Winsteps Help for Rasch Analysis*. http://homes.jcu.edu.au/~edtgb/%5Cnpapers3://publication/uuid/D56B724A-62FF-4D00-84E1-ECC888298B70

Livana, Mubin, & Basthomi, Y. (2020). Penyebab Stres Mahasiswa Selama Pandemi Covid-19. *Jurnsl Ilmu Keperawatan Jiwa*, *3*(2), 203–208.

Mahato, D., Gayen, P., Mahato, R. C. (2023). Relationship Between Academic Resilience and Internet Addiction Of Undergraduate Students of Purulia District of West Bengal: A Study. *EPRA International Journal of Multidisciplinary Research (IJMR)-Peer Reviewed Journal*, *9*(3), 103–106. https://doi.org/10.36713/epra2013

Mahmudah, U., Lola, M. S., Fatimah, S., & Suryandari, K. C. (2022). Academic Resilience and Science Academic Emotion in Numeration Under Online Learning: Predictive Capacity of an Artificial Neural Network. *Jurnal Pendidikan IPA Indonesia*, *11*(4), 542–551. https://doi.org/10.15294/jpii.v11i4.39091

Masten, A. S., Lucke, C. M., Nelson, K. M., & Stallworthy, I. C. (2021). Resilience in Development and Psychopathology: Multisystem Perspectives. *Annual Review of Clinical Psychology*, *17*, 521–549. https://doi.org/10.1146/annurev-clinpsy-081219-120307

Ngadi, I. (2023). Analisis Model Rasch Untuk Mengukur Kompetensi Pengetahuan Siswa Smkn 1 Kalianget Pada Mata Pelajaran Perawatan Sistem Kelistrikan Sepeda Motor. *Jurnal Pendidikan Vokasi Otomotif*, *6*(1).

Misbach, I. H., & Sumintono, B. (2014). Pengembangan dan Validasi Instrumen “Persepsi Siswa tehadap Karakter Moral Guru” di Indonesia dengan Model Rasch. *PROCEEDING Seminar Nasional Psikometri*, *May*, 148–162.

Muntazhimah, M. (2019). Pengembangan Instrumen Kemampuan Berpikir Reflektif Matematis siswa Kelas 8 SMP. *Imajiner: Jurnal Matematika Dan Pendidikan Matematika*, *1*(5), 237–242.

https://doi.org/10.26877/imajiner.v1i5.4551

Muslihin, H. Y., Suryana, D., Ahman, Suherman, U., & Dahlan, T. H. (2022). Analysis of the Reliability and Validity of the Self-Determination Questionnaire Using Rasch Model. *International Journal of Instruction*, *15*(2), 207–222. https://doi.org/10.29333/iji.2022.15212a

Nicoll, W. G. (2014). Developing Transformative Schools: A Resilience-Focused Paradigm for Education. *International Journal of Emotional Education*, *6*(1), 47–65. http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1085706&site=ehost-live

Noben, I., Maulana, R., Deinum, J. F., & Hofman, W. H. A. (2021). Measuring university teachers' teaching quality: a Rasch modelling approach. *Learning Environments Research*, *24*(1), 87–107. https://doi.org/10.1007/s10984-020-09319-w

Oladejo, A. I., Ademola, I. A., Ayanwale, M. A., & Tobih, D. (2023). Concept Difficulty in Secondary School Chemistry – an Intra-Play of Gender, School Location and School Type. *Journal of Technology and Science Education*, *13*(1), 255–275. https://doi.org/10.3926/jotse.1902

Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, *15*(2), 20111. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111

Rachman, T., & Napitupulu, D. B. (2017). Rasch Model for Validation a User Acceptance Instrument for Evaluating E-learning System. *CommIT (Communication and Information Technology) Journal*, *11*(1), 9. https://doi.org/10.21512/commit.v11i1.2042

Ramdani, R., Hanurawan, F., Ramli, M., Lasan, B. B., & Afdal, A. (2020). Development and Validation of Indonesian Academic Resilience Scale Using Rasch Models. *International Journal of Instruction*, *14*(1), 105–120. https://doi.org/10.29333/IJI.2021.1417A

Ramirez-Granizo, I. A., Sánchez-Zafra, M., Zurita-Ortega, F., Puertas-Molero, P., González-Valero, G., & Ubago-Jiménez, J. L. (2020). Multidimensional self-concept depending on levels of resilience and the motivational climate directed towards sport in schoolchildren. *International Journal of Environmental Research and Public Health*, *17*(2). https://doi.org/10.3390/ijerph17020534

Reeve, J., Cheon, S. H., & Yu, T. H. (2020). An autonomy-supportive intervention to develop students' resilience by boosting agentic engagement. *International Journal of Behavioral Development*, *44*(4), 325–338. https://doi.org/10.1177/0165025420911103

Rao, P. S., & Krishnamurthy, A. R. (2018). Impact of Academic Resilience on the Scholastic Performance of High School Students. *Indian Journal of Mental Health*, *5*(July), 453–462.

Reise, S. P. (1990). A Comparison of Item- and Person-Fit Methods of Assessing Model-Data Fit in IRT. *Applied Psychological Measurement*, *14*(2), 127–137. https://doi.org/10.1177/014662169001400202

Rodriguez, E. (2018). Google Forms in Library Instruction: Creating an Active Learning Space and Communicating with Students. *Innovative Pedagogy*, *1*(7), 69–82.

Rojas F., L. F. (2015). Factors Affecting Academic Resilience in Middle School Students: A Case Study. *GiST Education and Learning Research Journal*, *11*(11), 63–78. https://doi.org/10.26817/16925777.286

Romano, L., Angelini, G., Consiglio, P., & Fiorilli, C. (2021). Academic resilience and engagement in high school students: The mediating role of perceived teacher emotional support. *European Journal of Investigation in Health, Psychology and Education*, *11*(2), 334–344. https://doi.org/10.3390/ejihpe11020025

Sariati, Kadek, N., Suardana, Nyoman, I., Wiratini, & Made, N. (2020). Analisis Kesulitan Belajar Kimia Siswa Kelas Xi Pada Materi Larutan Penyangga. *Jurnal Imiah Pendidikan Dan Pembelajaran P-ISSN : 1858-4543 e-ISSN : 2615-6091*, *4*(1), 86–97. https://doi.org/10.23887/jipp.v4i1.15469.

Senocak, E., & Baloglu, M. (2014). The adaptation and preliminary psychometric properties of the Derived Chemistry Anxiety Rating Scale. *Chemistry Education Research and Practice*, *15*(4), 800–806. https://doi.org/10.1039/c4rp00073k

Sivakumar Ramaraj. (2019). Google Forms in Education. *Journal of Contemporary Educational Research and Innovations*, *9*(1), 35–39.

Sumintono, B. (2018). *Rasch Model Measurements as Tools in Assesment for Learning*. *173*(Icei 2017), 38–42. https://doi.org/10.2991/icei-17.2018.11

Sumintono, B., Islam, U., Indonesia, I., Widhiarso, W., & Mada, U. G. (2014). *untuk Penelitian Ilmu-Ilmu Sosial*. *November*.

Sumintono, B., Islam, U., Indonesia, I., Widhiarso, W., & Mada, U. G. (2015). *RascH*. *September*.

Sumintono, B., & Widhiarso, W. (2015). Penilaian Pendidikan dan Ujian. *AplikAsi RascH PemodelAn Pada Assessment Pendidikan*, *September*, 1–24.

Suud, F. M, Uyun, M, Na'imah, T. (2023). Development of achievement motivation training module to improve Islamic student academic resilience in disaster areas. *Psikis: Jurnal Psikologi Islami*, *9*(2), 312–323.

Tabatabaee-Yazdi, M., Motallebzadeh, K., Ashraf, H., & Baghaei, P. (2018). Development and validation of a teacher success questionnaire using the rasch model. *International Journal of Instruction*, *11*(2), 129–144. https://doi.org/10.12973/iji.2018.11210a

Taufiq, A., Yudha, E. S., Md, Y. H., & Suryana, D. (2021). Examining the Supervision Work Alliance Scale: A Rasch Model Approach. *The Open Psychology Journal*, *14*(1), 179–184. https://doi.org/10.2174/1874350102114010179

Timilsena, N. P., Krishna, ;, Maharjan, B., & Devkota, M. (2022). Teachers' And Students' Experiences In Chemistry Learning Difficulties. *Journal of Positive School Psychology*, *2022*(10), 2856–2867. http://journalppw.com

Tjabolo, S. A., & Otaya, L. G. (2019). Quality of school exam tests based on item response theory. *Universal Journal of Educational Research*, *7*(10), 2156–2164. https://doi.org/10.13189/ujer.2019.071013

Ulya, H., & Gumiandari, S. (2023). Efforts to Increase Academic Resilience of IAIN Syekh Nurjati Cirebon Students Through Strengthening Self-Regulation Capabilities. *Darussalam: Journal of Psychology and Educational*, *2*(1) 55-71. https://doi.org/10.55849/djpe.v2i1.43

van de Grift, W. J. C. M., Houtveen, T. A. M., van den Hurk, H. T. G., & Terpstra, O. (2019). Measuring teaching skills in elementary education using the Rasch model. *School Effectiveness and School Improvement*, *30*(4), 455–486. https://doi.org/10.1080/09243453.2019.1577743

Van Zile-Tamsen, C. (2017). Using Rasch Analysis to Inform Rating Scale Development. *Research in Higher Education*, *58*(8), 922–933. https://doi.org/10.1007/s11162-017-9448-0

Wahyuningsih, S. (2021). Using the Rasch's Partial Credit Model to Analyze the

Quality of an Essay Math Test. *Proceedings of the 1st International Conference on Mathematics and Mathematics Education (ICMMEd 2020)*, *550*(Icmmed 2020), 257–265. https://doi.org/10.2991/assehr.k.210508.073

Weidlich, J., & Kalz, M. (2021). Exploring predictors of instructional resilience during emergency remote teaching in higher education. *International Journal of Educational Technology in Higher Education*, *18*(1). https://doi.org/10.1186/s41239-021-00278-7

Widana, I. W. (2018). Higher Order Thinking Skills Assessment towards Critical Thinking on Mathematics Lesson. *International Journal of Social Sciences and Humanities (IJSSH)*, *2*(1), 24–32. https://doi.org/10.29332/ijssh.v2n1.74

Winarti, A., & Mubarak, A. (2019). Rasch Modeling: A Multiple Choice Chemistry Test. *Indonesian Journal on Learning and Advanced Education (IJOLAE), 2*(1), 1-9. doi:https://doi.org/10.23917/ijolae.v2i1.8985

Wu, M., Tam, H. P., & Jen, T.-H. (2016). Educational Measurement for Applied Researchers. In *Educational Measurement for Applied Researchers*. https://doi.org/10.1007/978-981-10-3302-5

Zahro', S. F., & Ismono, I. (2021). Analisis Kemampuan Multirepresentasi Siswa Pada Materi Kesetimbangan Kimia di Masa Pandemi Covid-19. *Chemistry Education Practice*, *4*(1), 30. https://doi.org/10.29303/cep.v4i1.2338

Zakiyah, Z., Ibnu, S., & Subandi, S. (2018). Analisis Dampak Kesulitan Siswa pada Materi Stoikiometri terhadap Hasil Belajar Termokimia dan Upaya Menguranginya dengan Metode Pemecahan Masalah. *EduChemia (Jurnal Kimia Dan Pendidikan)*, *3*(1), 119. https://doi.org/10.30870/educhemia.v3i1.1784

Zamri bin Khairani, A., & Bin Abd. Razak, N. (2015). Modeling a Multiple Choice Mathematics Test with the Rasch Model. *Indian Journal of Science and Technology*, *8*(12). https://doi.org/10.17485/ijst/2015/v8i12/70650

Zubairi, A., & Kassim, A. N. L. (2006). Classical and Rasch analyses of dichotomously scored reading comprehension test items. *Malaysian Journal of ELT Research*, *2*(March), 1–20.