IJOLAE Indonesian Journal on Learning and Advanced Education

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

# Assessing Students Understanding of Chemical Bonds Material by Rasch Modeling

**Almubarak[1], Parham Saadi[2], Restu Prayogi[3], Pamela Paula Maldini[4]**
[1-3]Faculty of Teacher Training and Education, Universitas Lambung Mangkurat, Indonesia
[4]Faculty of Management and Business, Tampere University of Applied Sciences, Finland

**DOI: 10.23917/ijolae.v5i3.22242**
Received: April 16th, 2023. Revised: August 16th, 2023. Accepted: August 31st, 2023
Available Online: September 22nd, 2023. Published Regularly: September, 2023

## Abstract

This study aimed to assess high school student's understanding of Banjarmasin by Rasch modeling, precisely the cognitive aspect. The research method was descriptive with a quantitative approach to assess the pattern of reactions and symptoms of Rasch data. According to the research findings, the person reliability (students) was +0.66 based on the Rasch modeling analysis, with the criterion satisfactory, indicating that the students knowledge was adequate. At the same time, Cronbachs' alpha score had a value of +0.71 and met the criteria of excellent. In addition, the students with the highest level of understanding were coded 127P12B with a logit value (person measure) of 2.52 and average students with logit value of -0.77 or <0. In contrast, the students with low abilities were coded 030P10B and 059L11B (same logit value, -3.27 or <0). Other data were INFIT MNSQ and OUTFIT MNSQ (person) with average values of +0.99 and +1.14 (closer to 1, the better), while the INFIT and OUTFIT ZSTD values were -0.1 and 0.0 respectively (closer to 0, better). The most difficult question was Q16, which had a logit score of 1.96; students' logit values carried this question. Students were regarded to have appropriate knowledge even though their ability exceeded the problem ability. In conclusion, the Rasch model-based pre-learning evaluation was found to be useful in measuring students' cognitive grasp of chemical bonding material. This study could serve as the primary reference for teachers in assessing students' level of knowledge before they begin learning. In addition to interpreting student knowledge through various Rasch data presentations, a study of the structure of questions with varying difficulty levels could be used to assess the full group of students' understanding of chemistry.

**Keywords:** cognitive aspect, learning progress, pre-learning evaluation, rasch modeling

***Corresponding Author:***
*Almubarak, Faculty of Teacher Training and Education, Lambung Mangkurat University, Indonesia*
*Email: almubarak_kimia@ulm.ac.id*

## 1. Introduction

Assessment in the field of education is a process that needs to be carried out by teachers (Sumintono, 2018a). Learning assessment is an activity that is always present in the school environment, especially in the context of understanding analysis (Barke et al., 2012; Mešić et al., 2019; Sumintono & Widhiarso, 2015). Without the treatment of learning assessment, teachers experience difficulties in assessing the progress and development of students, so learning assessment becomes essential (Barke et al., 2012; Setyaningsih et al., 2022; Sulistyanto et al., 2022). However, not all assessment forms can be representative and follow the goals desired by the teacher (Sumintono & Widhiarso, 2015). The Rasch model is an appropriate measurement technique that teachers can use to assess students' academic development (Sumintono & Widhiarso, 2015; Winarti et al., 2019; Winarti & Almubarak, 2019). The main objective of this research is to comprehend student response patterns through the presentation of Rasch statistical data. The response pattern is then analyzed to determine the student's understanding of chemical bonding material.

The rationale for research objectives is that many teachers are still unaware of the need for conducting assessments before beginning chemistry lessons in class (Sumintono & Widhiarso, 2015; Winarti et al., 2019). On the other hand, the teacher needs to conduct a chemistry learning assessment considering that many students have a negative stigma towards chemical material, so the Rasch model is essential. In terms of assessment application, the Rasch model is not intended for chemistry learning, but the cognitive analysis of students using Rasch is unique and innovative for teachers (Chan et al., 2020; Goldstein, 1979; Talib et al., 2018; Winarti et al., 2019). Pre-assessment with Rasch is very helpful in designing the chemistry learning process, especially in transforming preconceptions into scientific concepts (Barke et al., 2012; Mezirow, 1997).

Initial understanding is termed by Barke et al. (2012) as preconceptions. This term describes students' mental models in interpreting a phenomenon based on their observations without concretely seeing the suitability of scientific terminology (Barke et al., 2012). As a result, students' mental models form incoherent constructions, so the transformation of preconceptions into scientific concepts is complex. This situation makes students prone to misconceptions, so teachers need to detect the level of student understanding through understanding analysis (Barke et al., 2009, 2012; Bouw et al., 2021). Analysis of students' cognitive structures is a form of learning assessment that allows teachers to learn about their students' cognitive, behavioral, and other development (Grove & Bretz, 2007; Gumartifa et al., 2023; Quinlan et al., 2019; Rauch & Hartig, 2010; Supriyanto et al., 2022).

Assessment of learning using written tests is an essential technique for analyzing student's development and learning progress (Adimayuda et al., 2020; Campbell, 2015; Dewi et al., 2019; Sumintono, 2018b). In addition, an analysis of students' understanding before class activities is the main reference for teachers, especially in designing teaching concepts that suit students' learning needs (Barke et al., 2012; Chetty et al., 2019; Mezirow, 1997; Valtonen et al., 2021; Yessi, 2021). The analysis of students' understanding prior to learning is a form of teaching method used to investigate the structure of students' understanding of a chemical material.

Analysis using the Rasch model is considered an appropriate and practical measurement technique for representing students' abilities through data presentation formats (Adimayuda et al., 2020; Boone, 2016; Noben et al., 2021; Sumintono & Widhiarso, 2015; Talib et al., 2018; van der Lans et al., 2018). In the context of the measurement level, the Rasch model has a scientific perspective that everyone has the same opportunity to be right on every question (Sumintono, 2018b; Sumintono & Widhiarso, 2015). Furthermore, the Rasch model has a logarithmic approach to the measurement process, so the Rasch model's output analysis provides a high accuracy level in measurement (Adimayuda et al., 2020; Sumintono & Widhiarso, 2015). Furthermore, Rasch is more than just a strategy for measuring student abilities; it can also analyze the quality of questions in a statistical data format.

The research of (Chiang, 2015; van de Grift et al., 2019) found that the Rasch model can be an essential assessment for classroom learning. Rasch can examine each item answered by students to determine their abilities, while also identifying each student's learning progress. As a result, the Rasch analysis enables teachers to identify strengths and weaknesses based on interpreted data. The Rasch model not only measures and justifies students' cognitive capacities, but also the effectiveness of the assessment methods utilized, so that the strategy can be used to improve the quality of chemistry teaching and learning (Adimayuda et al., 2020; Chow et al., 2018; Mešić et al., 2019; Noben et al., 2021; Rabbitt, 2018; Sumintono, 2018b). The research concept analyzes students' abilities using the Rasch model as a primary assessment before the teacher begins the learning process.

## 2. Method

An adequate and appropriate method in the analysis process was required to determine the symptoms of the data that resulted from the Rasch model analysis. Since the results were statistical data with a logarithmic approach integrated into the Rasch model software, the quantitative method was the best choice for reaching this goal. The quantitative method was the best way to assess students' cognitive ability with regard to chemical material, specifically chemical bonds (Creswell, 2009; Sumintono & Widhiarso, 2015). The data symptoms assisted the teacher in determining the level of understanding of the subject matter. Teachers could utilize the findings of this analysis to design lessons according to their students' cognitive abilities.

### a. Rasch Modeling with the Dichotomous Model

In Rasch modeling, participant's abilities were shown through logit values and various data patterns resulting from Rasch modeling. Furthermore, this modeling examined not only the abilities of the participants, but also the abilities of the items, ensuring that the instruments utilized represent the context of measurement validity. The statement below is the Rasch modeling principle (Sumintono & Widhiarso, 2015):

*"a person having a greater ability than another person should have the greater probability of solving any item of the type in question, and similar, one item being more difficult than another means that for any person, the probability of solving the second item is the greater one"* (Rasch, 1960 in Bond and Fox, 2015).

According to the statement above, students with superior abilities had a greater chance of successfully answering one item. Using the same principle, one complex item reduced students' chances of answering correctly. In the Rasch Model, multiple choice questions were called Dichotomies, in which the Rasch modeling combines an algorithm that stated the results of probabilistic expectations of the item (i) and respondent (n), which was mathematically expressed (Runnels, 2012; Sumintono, 2018b; Sumintono & Widhiarso, 2015).

$$P_{ni}\,(X_{ni} = 1 |\, \beta_n, \delta_i) = \frac{e^{(\beta_n - \delta_i)}}{1 + e^{(\beta_n - \delta_i)}}$$

This formula, $P_{ni}$ ($X_{ni}$ = 1| $b_n$, $d_i$), was the probability of respondent (n) in item (i) to produce a correct answer (x = 1), with the respondent's ability, βn, and item difficulty level δi.

The above equation by Rasch could be simplified even more by plugging in the logarithm function and making it as follows:

$$Log\ P_{ni}\,(X_{ni} = 1 |\, \beta_n, \delta_i) = \beta_n - \delta_i$$

So, the probability of one's success could be written as:

Probability of success = respondent's ability – item difficulty level

The cognitive test was a dichotomous format (multiple choice) with 25 questions on chemical bonds. Chemical bonds were fundamental knowledge in learning chemistry because they were part of the material structure of atoms. In statistical language, students should have thorough understanding of the material because it was considered primary. Since the sample consisted of students who had studied chemical bonding material, it was assumed that the students could answer the questions presented. The participants were high school students from Banjarmasin, South Kalimantan.

Purposive sampling was used in this study, and the samples were taken after going through numerous considerations. The students had studied with certain materials experience knowledge transformation in the environment of learning, thus they tend to create their knowledge and compare it to new things they discovered (Barke et al., 2012). The scientific reason for selecting the sample strengthened the assumption that students could answer the questions presented. This expectation discrepancy could be seen from data symptoms as well as students' response patterns, while some teachers rarely made accurate measurements before the learning process occurred. This research was essential for teachers to obtain valid information regarding the analyzation of students learning needs before designing learning processes (Barke et al., 2012; Mezirow, 1997; Mezirow et al., 2019).

Rasch's cognitive test referred to several data size models: Infit, Outfit, Mean-Square Fit Statistics, and Standardized Fit Statistics (ZSTD). Infit is inlier sensitive or information-weighted fit, in other words, the sensitivity of the response pattern to the target item on the respondent (person) or vice versa. Outfit means outlier-sensitive fit, which measures the sensitivity of response patterns to items with a particular difficulty level from the respondent (person) or vice versa. For example, suppose the participant's response is incorrect (careless); in this case, the participant cannot answer simple questions despite their superior ability. Another circumstance exists in which participants can work on complicated questions but have limited ability (lucky guesses), which are relatively easy to detect.

Mean-Square Fit Statistics measures randomness, the amount of distortion in the measurement system. The expected value is between 0.5-1.5; if the value is less than this, it indicates that the data is too predictable (data overfit the model). Regarding statistical value, chi-square represents the degree of freedom; its value is always positive. Furthermore, ZSTD is a t-test for the hypothesis, which asks whether the data fits the model or not. The result is a z-value, i.e., unit deviation. It describes the impossibility of the data, i.e., its significance if it fits the model.

These criteria were absolute and constituted a reference in evaluating the generated data, implying that students' understanding of chemical material was judged based on these criteria. If the statistical result did not fall within the MNSQ, ZSTD, or Pt Mean Corr ranges, the student was considered to lack a solid understanding of the material being evaluated (chemical bonding). In addition to mapping students' abilities according to the interests of achievement grouping, the Rasch model might discover individuals whose response patterns were incorrect. The difference between the answers supplied by students depending on their skills and the ideal model was shown by the pattern of varied responses. That was, this approach could be used to determine the consistency of students' thinking and to uncover student dishonesty.

In the context of item analysis, the criteria above served as the foundation for determining whether or not the items performed normally when measured (Rizbudiani et al., 2021; Sihombing et al., 2018; Sumintono, 2018a; Sumintono & Widhiarso, 2015).

Boone et al. (2014) and Bond & Fox (2015) explained that outfit means-square, outfit z-standard, and point measure correlation were the criteria used to see the patterns of student responses to the items worked and the items themselves. If the three requirements (MNSQ, ZSTD, and Pt Mean Corr) were not met, they were inevitably inadequate and needed to be repaired or replaced. It would ensure that the level of student's understanding was tested using appropriate and quality questions (Masito et al., 2022; Sumintono & Widhiarso, 2015; Yudha, 2023). However, the value of the suitability of the items was strongly influenced by the size of the sample. Examination of participants (person) also adopted the same criteria. This criterion aimed to check the suitability of items that did not fit (misfit items) and the participants (person fit). The following criteria were referred to (Sumintono & Widhiarso, 2015):

1) Received Outfit Mean Square (MNSQ) value: $0.5 < MNSQ < 1.5$
2) Received Outfit Z-Standard (ZSTD) value: $-2.0 < ZSTD < +2.0$
3) Point Measure Correlation (Pt Mean Corr): $0.4 <$ Pt Measure Correlation $< 0.85$

## 3. Result and Discussion

Several conclusions from the research had been gathered and were likely to result in a reform in chemistry teaching. Finding 1 data from changeable map images or Wright maps revealed that student 127P12B had the highest ability with a logit score of 2.52, while the problem ability was 1.92. That was the ability of the students to outperform the problem. Finding 2, the scalogram data showed that only 14 students (or 10% of the total number of students) correctly answered question number 16 (Q16), indicating that Q16 was rated as the most difficult problem for chemical bonding material. Finding 3, person dependability (students) of +0.66 with the criterion "adequate", indicating that student understanding was adequate, even though Cronbach's alpha score had the criterion "good" with a value of +0.71. Finding 4 was based on scalogram data, which demonstrated that students with code 114 (114P12B) were not careful since the easiest questions could not be answered, while questions with a high difficulty level might. Furthermore, because the response patterns of 114 students were inconsistent, the assumption of "guessing" or lucky guessing was placed on 114 students. In the following paragraph, these findings were examined in depth, and the conversation became scientific evidence of the value of utilizing Rasch analysis.

Figure 1 was a variable or Wright map; the image compared the levels between student abilities (person) and questions (items). The left side was the pattern of student responses, while the right side was the distribution of item difficulty levels. The code # indicated 3 (three) students, and the sign (.) showed 1 to 2 students. It meant that if the code line was ###., the code showed > 5 people or equal ability. Students with the same ability were shown through the ### code, meaning that these students had the same logit value (person measure).

The maps variable showed several conditions; in condition 1, the average logit or person measure value was -0.77 (or <0). The average logit value was relevant to the figure below. On average, it could be seen that the number of students (#) below the 0 scale was more than the number of students with positions above the 0 scale. There were even some students who only had one correct answer through the scalogram data. In addition, the average students only answered questions with a difficulty level of <0, so students could only answer questions with a high level of ability with a logit score of >0 (and even then, a small proportion of students). In condition 2, items with a logit value > 1 were only answered by a few students, meaning that this item lacked function because only students with a logit value > 1 could answer the item correctly.

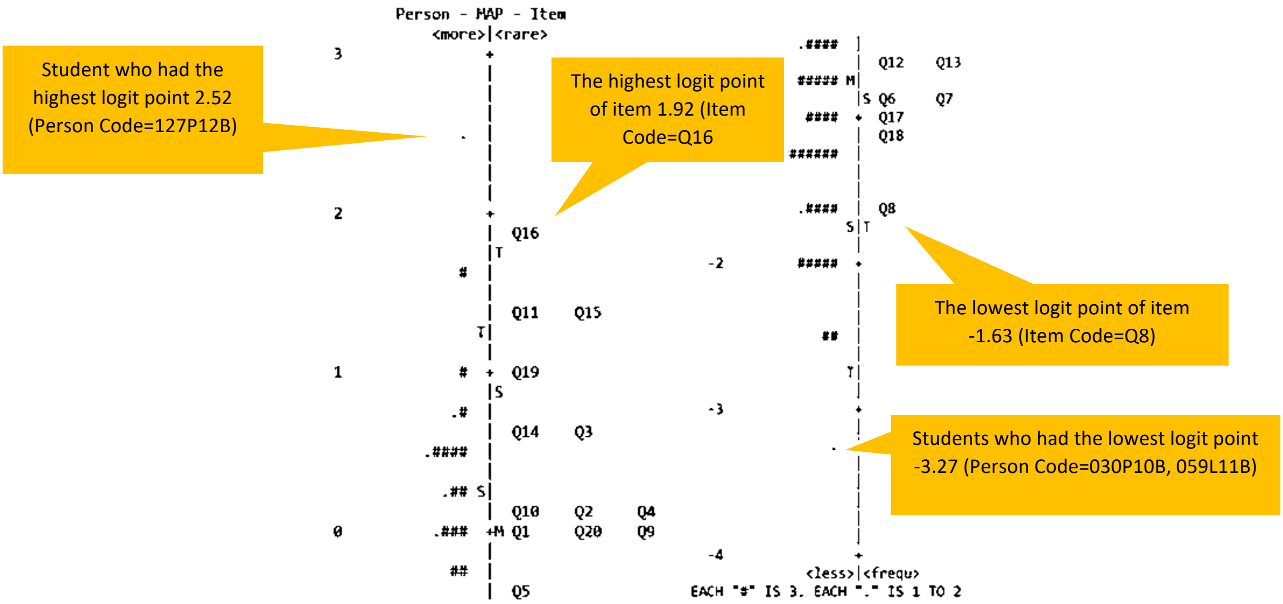


**Figure 1. Maps Variable (Wright Maps)**

Figure 1 above with the point code (**.**) (left side, person) was student 127P12B who had logit score of 2.52 with the highest ability, while students 030P10B and 059L11B was the students with the lowest ability who had logit score of -3.27. On the other hand, questions with a great difficulty level had a logit value of 1.92 (Q16 was question number 16), showed on the right side of Figure 1. Then, question number 8 or Q8 was a question with the lowest difficulty level (logit value = -1.63), meaning that Q8 was the easiest to answer. In the context of the matter, chemical bonds were relatively basic materials because the properties of matter were simple in terms of theoretical applications. Chemical bonding materials could also be the basis for determining the shape of molecules (Petrucci et al., 2011). Lewis' theory was the most dominating theory in understanding chemical bonding, making Lewis' thinking in describing bonding in compounds was used as a basis. Regarding the question with the most significant difficulty level (Q16), the following explained Q16.

*Problem Q16* - The number of lone pairs of electrons in the $N_2$ molecule is ...

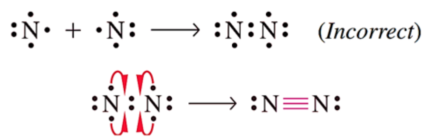
a. One pair
**b. Two pairs**
c. Three pairs
d. Four pairs
e. Five pairs

*Answer* - the $N_2$ molecule has **two lone pairs of electrons**.

The scalogram data showed that among the 154 students who answered the questions, only 14 persons answered question number 16 (Q16), or 10% of the total number of students correctly; meanwhile, Q16 was classified as a fundamental concept in understanding chemical bonding material. In terminology, the $N_2$ the molecule was two molecules that atoms shared more than one pair of electrons. In addition, Lewis' theory was the leading theory that could be used to broaden the understanding of Q16. The biggest asset for students to understand concretely about Q16 was the basic definition of chemical bonding, including the types of chemical bonds (ionic, covalent, polar covalent, non-polar covalent, hydrogen, etc.).

In addition to the basic definitions above, integrating chemical representation concepts was the most appropriate and scientific approach in developing students' understanding of chemical bonding (Barke et al., 2009; Cheng & Gilbert, 2009; Ryan & Herrington, 2014). Representation could assist students understood the material being studied in a concrete way, preventing ongoing misconceptions (Barke et al., 2012; Darmiyanti et al., 2017; Rusmansyah et al., 2021). Ten percent indicated that teachers had to realize the importance of reconstructing students' basic knowledge, especially in chemical bonding material. The expert emphasized that the involvement of the concept of representation could strengthen student's partial abilities so that students had a scientific understanding of learning chemistry (Barke et al., 2012; Justi et al., 2009; Ryan & Herrington, 2014).

The following identified the Q16 problem by describing the Lewis structure of the $N_2$ the molecule as an initial step.




Each N atom appeared to have only six outer electrons instead of the expected eight. This situation could be corrected by bringing four unpaired electrons into the region between the N atoms and using them as additional bond pairs. In this way, students could finally see the sharing of three pairs of electrons between the two N atoms. The bond between the two N atoms in $N_2$ was a triple covalent bond (≡). Regarding the triple covalent bond in $N_2$, the bonding was powerful and challenging to break in chemical reactions. This tremendous power made $N_{2(g)}$ inert. As a result, $N_{2(g)}$ coexisted with $O_{2(g)}$ in the atmosphere, forming only small amounts of nitrogen oxides at high temperatures. In addition, the lack of reactivity of $N_2$ and $O_2$ Reactivity was an essential condition for life on Earth. The humidity of $N_{2(g)}$ also made it difficult to synthesize nitrogen compounds. In addition, nitrogen was also one of the elements with the highest electronegativity of the elements (3.04 on the Pauling scale), surpassed only by chlorine (3.16), oxygen (3.44), and fluorine (3.98).

Figure 2 below was another strategy to help students understood chemical bonding, especially in the context of Q16, namely the dinitrogen molecule. The image was a visualization of the distribution of electrons $N_2$. There were seven pairs of electrons with atomic number 7 (7 electrons). In their ground state, they were regular in the electron configuration of $1s^2 2s^2 2p_x^1 2p_y^1 2p_z^1$.

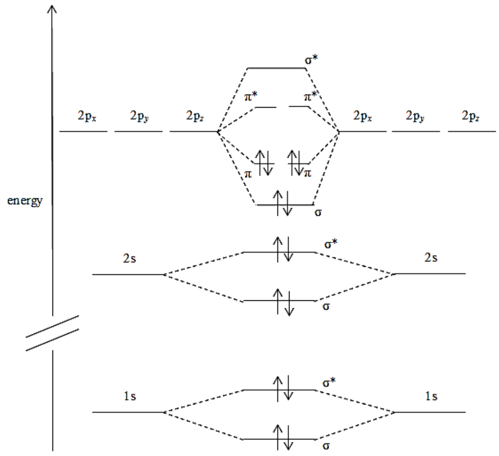



**Figure 2. Dinitrogen Molecular Orbital Diagram, $N_2$ (Wikipedia, 2023)**

Therefore, there were five valence electrons in the 2s and 2p orbitals, three of which (the p electrons) were unpaired (note the electrons circled in red in Figure 2). That was, the $N_2$ the molecule had two pairs of free electrons, while only 10% of the students answered correctly with the number 16. Explanation of research results based on Rasch data and scientific terminology regarding problem number 16 provided an overview of students' level of understanding. In addition, this assessment could be a reference for teachers in developing and designing chemistry lessons so students could achieve scientific understanding.

Figure 2 was a model that needed to be understood by students. Expert studies explained that a model was always initially produced in students' minds, which was then called a mental model (Justi et al., 2009). For communication purposes, the model had to be expressed in different modes of representation (concrete, verbal, mathematical, visual, gestural, or mixed). Aspects at the macro and sub-micro levels could be expressed in all representation models, while aspects at the symbolic level were generally expressed in verbal or mathematical modes. It showed the importance of using representational models to comprehensively understand the three levels at which a particular entity could be modeled (Barke et al., 2012; Cheng & Gilbert, 2009; Justi et al., 2009). Figure 2 needed to be understood by students to obtain scientific knowledge about chemical bonding material. Impact: if the teacher carried out advanced cognitive analysis, all students were expected to answer questions for this context correctly.

**Table 1. Statistics Summary of 154 Measured Person**

| | Total Score | Count | Measure | Model Error | INFIT | | OUTFIT | |
|---|---|---|---|---|---|---|---|---|
| | | | | | MNSQ | ZSTD | MNSQ | ZSTD |
| Mean | 7.2 | 20.0 | -.77 | .55 | .99 | -.1 | 1.14 | .0 |
| S.D. | 3.6 | .0 | 1.00 | .09 | .22 | 1.0 | .65 | 1.2 |
| Max. | 18.0 | 20.0 | 2.52 | 1.04 | 1.57 | 2.6 | 4.49 | 3.1 |
| Min. | 1.0 | 20.0 | -3.27 | .49 | .60 | -2.4 | .33 | -2.2 |
| Real RMSE | | .59 | TRUE SD | .81 | Separation | 1.38 | Person Reliability | .66 |
| Model RMSE | | .56 | TRUE SD | .83 | Separation | 1.48 | Person Reliability | .69 |

S.E. Of Person MEAN = .08
Person RAW Score-to-Measure Correlation = .99
Cronbach Alpha (KR-20) Person RAW Score "Test" Reliability = .71

The statistical summary table provided comprehensive information about the quality of student response patterns, so this Rasch data was also essential in analyzing student understanding levels. The logit value of -0.77 was the average of all students' working scores on the given items. The average student whose score less than 0.0 suggested that students did not fully understand the chemical bonding material, although the Cronbach alpha score (AC) using the "good" criterion was +0.71. The AC value indicated the measurement reliability or interaction between the person (students) and the overall item (question). It also implied that the instrument used could provide reliable information as a data collection tool.

Person reliability (students) based on the Rasch modeling analysis was +0.66 with the criterion "adequate", meaning that students' understanding was at an adequate level; meanwhile, Cronbach's alpha score had the criterion "good" with a value of +0.71. The "adequate" criterion and previous results analysis revealed that students needed to undergo transformation in order for their understanding of chemical bonding and other materials to improve. These numbers were quite helpful in determining the overall degree of understanding among students.

Furthermore, other data were INFIT and OUTFIT MNSQ (person) with an average value of +0.99 and +1.14 (closer to 1, the better), meanwhile the value of INFIT and OUTFIT ZSTD, the respective values were -0.1 and 0.0 (closer to 0, the better). The MNSQ INFIT and OUTFIT values indicated the randomness or amount of distortion in the measurement system with an expected value of 0.5 – 1.5. The INFIT and OUTFIT MNSQ data met these requirements, indicating that the test presentation utilized in the assessment was in excellent enough condition that student responses were indicative of research objectives (Sumintono & Widhiarso, 2015). In addition to the MNSQ value, the INFIT and OUTFIT ZSTD values were in the criteria of -1.9 – 1.9 with the description of the data having a logical estimate, meaning that the measurements taken had no deviation from what was being measured. In this context, measurements regarding students' understanding of chemical bonding material could be used as a learning assessment (Sumintono & Widhiarso, 2015).

Interpretation of students' understanding could also be seen from the Scalogram data (Figure 3). Student 127P12B was the student who had the highest score with a logit score of 2.52 and only had two wrong answers. The student was considered to have an excellent understanding of chemical bonding based on the response pattern of the Rasch data obtained. Furthermore, the response pattern of 127P12B students was deemed satisfactory because, in addition to receiving only two incorrect answers, these two numbers were questions of the most significant difficulty. As a result of the regularity of their responses, 127P12B students had a good level of understanding of the subject matter. Then, Students 53, 116, and 126 had the same logit score (1.92), while the response patterns of these students were considered careless. Their response patterns were considered careless because they were able to answer the questions with a high degree of difficulty, but not the easiest ones. The data showed that students with this code understood the chemical bonding material. However, the teacher might need to review several parts of the material again as a learning assessment.

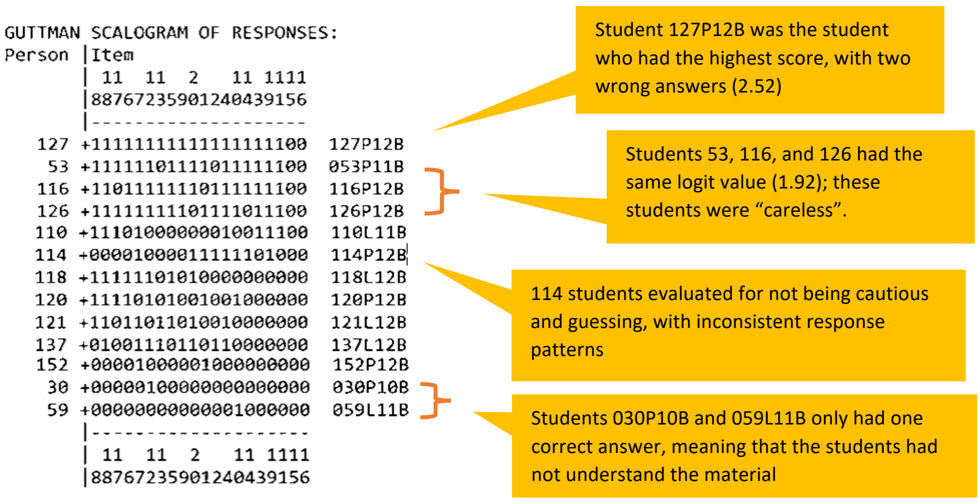



**Figure 3. Student's responses from Scalogram**

Student 114 (114P12B) was a student who was deemed careless since they were able to answer the questions with a high degree of difficulty, but not the easiest ones. Furthermore, because student 114's response patterns were inconsistent, the guessing or lucky guess assumption was placed on them. Students with these impairments required specialized learning procedures in order to gain a scientific understanding of chemical material. Unlike student 114P12B, students 030P10B and 059L11B only had one right answer, and their logit scores were -3.27, which was considerably below 0.0. Furthermore, students 030P10B and 059L11B only guessed one number correctly, with similar response patterns. This suggested that these students did not have a solid understanding of the chemical bonding material. As a result, the visualization of this scalogram data was also vital to the teacher in analyzing student understanding.

The difference in student logit values in the scalogram figure showed a finding that teachers needed to facilitate students from these findings. There were several recommendations in the scalogram data situation, namely the teacher had to group students based on the symptoms shown. For example, the coding for students in the criteria for guessing, lowest score, and cheating was put together; the grouping was then evaluated to determine the location of the student's problems from the questions they were working on. Then, the teacher offered to conduct peer tutoring or enrichment so that students' knowledge improved.

The teacher then ranked the students based on their logit value. The most scientific assumption was that students with similar logits had similar levels of understanding. The same logit value was also likely to have the same response pattern, so this tip was useful for teachers in implementing learning strategies that met students' learning needs. That was, teachers could utilize the data from each Rasch model to assess students' understanding of chemistry.

Based on the findings described above, it could be concluded that the Rasch model was the primary reference for teachers in analyzing students' level of understanding, especially the analysis of learning needs before the learning process began. Pre-learning analysis was critical since the teacher needed to know how much students understood before they began learning. Analysis with the Rasch model also helped teachers to design learning by determining learning models, scenarios, media, and assignments so that students were not burdened with learning. A research showed that the Rasch model could be used as a reference in increasing student understanding, learning quality, and guidelines in designing chemistry learning strategies based on an analysis of student understanding (Winarti et al., 2019). In the context of measurement, the Rasch model produced two categories, namely participant ability and item ability, in the same calculation, resulting in a scientific approach to learning assessment implementation (Andrich, 2010; Andrich & Pedler, 2019; Sumintono & Widhiarso, 2015). A similar research came from (Mamat et al., 2014) who stated that the Rasch model offered a reliable solution in producing accurate analytical skills, data presentation scores for students, and scientific reliability scores for student answers.

The Rasch modeling provided many perspectives in assessing student's complete understanding, while some teachers rarely conducted needs analysis before designing learning process. The Rasch model-based analysis might be the best option for teachers in this situation. In addition to interpreting student knowledge through various Rasch data presentations, a study of the structure of questions with varying difficulty levels could be used to assess the full group of students' understanding of chemistry. Rasch modeling was essential to learning chemistry since it might be applied in many sorts of science resources and student grade levels.

## 4. Conclusion

The conclusions were based on the research, which showed that 127P12B students had the highest ability with a logit score of 2.52; on variable maps or Wright maps, the question ability was 1.92. Furthermore, the scalogram data revealed that only 14 students (or 10% of the total number of students) correctly answered question number 16 (Q16), indicating that Q16 was ranked as the most difficult question. According to the Rasch modeling analysis, person reliability (students) was +0.66 with the criterion "adequate", indicating that student understanding was adequate, even though Cronbach's alpha score had the criterion "good" with a value of +0.71. Another Scalogram finding was that students with code 114 (114P12B) were regarded as careless because they were able to answer the questions with a high degree of difficulty, but not the easiest ones. These findings suggested that utilizing the Rasch model for analysis was a creative and successful technique for teachers to determine students. Rasch data presentation made it easier for teachers to justify student's understanding patterns, therefore Rasch results provided a good chance for teachers to design learning based on their students' needs.

Person reliability (students) based on Rasch modeling analysis is +0.66 with the criterion "adequate", indicating that student understanding was still at an adequate level, even though Cronbach's alpha score had the criterion "good" with a value of +0.71. Even if Cronbach's alpha score was a good criterion, data person reliability was high enough that the average student did not have a real understanding of the chemical bonding material. It meant that teachers needed to monitor students' understanding of learning concepts that encouraged expanded knowledge, such as using life context-based approaches, appropriate learning media, doing laboratory experiments, and possibly project-based assignments.

Contributively, teachers could use the findings of this research as a primary reference when assessing students' understanding of other chemistry materials. Adoption of these findings by teachers aided in the acquisition of new findings in following study utilizing various scientific methodologies, resulting in novelties and new findings in the field of education.

## 5. References

Adimayuda, R., Aminudin, A. H., Kaniawati, I., Suhendi, E., & Samsudin, A. (2020). A Multitier Open-Ended Momentum and Impulse (MOMI) Instrument: Developing and Assessing Quality of Conception of 11th Grade Sundanese Students with Rasch Analysis. *International Journal of Scientific and Technology Research*, *9*(2), 4799–4804.

Andrich, D. (2010). Educational Measurement: Rasch Models. *International Encyclopedia of Education*, 111–122.

Andrich, D., & Pedler, P. (2019). A Law of Ordinal Random Error: The Rasch Measurement Model and Random Error Distributions of Ordinal Assessments. *Measurement: Journal of the International Measurement Confederation*, *131*, 771–781. https://doi.org/10.1016/j.measurement.2018.08.062

Barke, H.-D., Harsch, G., & Schmid, S. (2012). Essentials of Chemical Education. In *Angewandte Chemie International Edition, 6(11), 951–952.* Springer.

Barke, H.-D., Hazari, A., & Yitbarek, S. (2009). *Misconceptions in Chemistry (Addresing Perceptions in Chemical Education)*. Sense Publisher. https://doi.org/10.1007/978-3-540-70989-3_2

Boone, W. J. (2016). Rasch Analysis For Instrument Development: Why,When,And How? *CBE Life Sciences Education*, *15*(4). https://doi.org/10.1187/cbe.16-04-0148

Bouw, E., Zitter, I., & de Bruijn, E. (2021). Designable Elements Of Integrative Learning Environments At The Boundary Of School And Work: A Multiple Case Study. In *Learning Environments Research* (Vol. 24, Issue 3). Springer Netherlands. https://doi.org/10.1007/s10984-020-09338-7

Campbell, M. L. (2015). Multiple-Choice Exams and Guessing: Results from a One-Year Study of General Chemistry Tests Designed to Discourage Guessing. *Journal of Chemical Education*, *92*(7), 1194–1200. https://doi.org/10.1021/ed500465q

Chan, S., Looi, C., & Sumintono, B. (2020). A Rasch Model Measurement Analysis.

*Journal of Computers in Education*, *8*(2), 213–236.

Cheng, M., & Gilbert, J. K. (2009). Introduction: Macro, Submicro and Symbolic Representations and the Relationship Between Them: Key Models in Chemical Education. In J. K. Gilbert & D. Treagust (Eds.), *Multiple Representations in Chemical Education: Models and Modeling in Science Education* (p. 369). Springer.

Chetty, N. D. S., Handayani, L., Sahabudin, N. A., Ali, Z., Hamzah, N., Rahman, N. S. A., & Kasim, S. (2019). Learning Styles And Teaching Styles Determine Students' Academic Performances. *International Journal of Evaluation and Research in Education*, *8*(4), 610–615. https://doi.org/10.11591/ijere.v8i3. 20345

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

Creswell, J. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). SAGE Publications, Inc.

Darmiyanti, W., Rahmawati, Y., Kurniadewi, F., & Ridwan, A. (2017). Analisis Model Mental Siswa Dalam Penerapan Model Pembelajaran Learning Cycle 8E Pada Materi Hidrolisis Garam. *JRPK: Jurnal Riset Pendidikan Kimia*, *7*(1), 38–51. https://doi.org/10.21009/jrpk.071.06

Dewi, R. K., Wardani, S., Wijayati, N., & Sumarni, W. (2019). Demand of ICT-Based Chemistry Learning Media In The Disruptive Era. *International Journal of Evaluation and Research in Education*, *8*(2), 265–270. https://doi.org/10.11591/ijere.v8i2.1710 7

Goldstein, H. (1979). Consequences of Using the Rasch Model for Educational Assessment. *British Educational Research Journal*, *5*(2), 211–220. https://doi.org/10.1080/0141192790050 207

Grove, N., & Bretz, S. L. (2007). CHEMX: An Instrument To Assess Students' Cognitive Expectations For Learning Chemistry. *Journal of Chemical Education*, *84*(9), 1524–1529. https://doi.org/10.1021/ed084p1524

Gumartifa, A., Syahri, I., Siroj, R. A., Nurrahmi, M., & Yusof, N. (2023). Perception of Teachers Regarding Problem-Based Learning and Traditional Method in the Classroom Learning Innovation Process. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *5*(2), 151–166.

Justi, R., Gilbert, J. K., & Ferreira, P. F. . (2009). The Application of a 'Model of Modelling' to Illustrate the Importance of Metavisualisation in Respect of the Three Types of Representation. In D. Treagust & J. K. Gilbert (Eds.), *Multiple Representations in Chemical Education, Models and Modeling in Science Education*. Springer.

Mamat, M. N., Maidin, P., & Mokhtar, F. (2014). Simplified Reliable Procedure for Producing Accurate Student's Ability Grade Using Rasch Model. *Procedia - Social and Behavioral Sciences*, *112*(Iceepsy 2013), 1077–1082. https://doi.org/10.1016/j.sbspro.2014.01 .1272

Masito, F., Oka, I. G. A. A. M., Cahyadi, C. I., Komalasari, Y., Anes, A. M., & Risdianto, E. (2022). Scientific Argumentation Skills Through The Rasch Model on Analysis of Survey Data on The Importance of Aviation Vocational Education in Indonesia. *Journal of Innovation in Educational* *and Cultural Research*, *3*(3), 487–498. https://doi.org/10.46843/jiecr.v3i3.178

Mešić, V., Neumann, K., Aviani, I., Hasović, E., Boone, W. J., Erceg, N., Grubelnik, V., Sušac, A., Glamočić, D. S., Karuza, M., Vidak, A., AlihodŽić, A., & Repnik, R. (2019). Measuring Students' Conceptual Understanding of Wave Optics: A Rasch Modeling Approach. *Physical Review Physics Education Research*, *15*(1), 1–20. https://doi.org/10.1103/PhysRevPhysEducRes.15.010115

Mezirow, J. (1997). Transformative Learning: Theory to Practice. *New Directions for Adult and Continuing Education*, *1997*(74), 5–12. https://doi.org/10.1002/ace.7401

Mezirow, J., Cottafava, D., Cavaglià, G., Corazza, L., Raikou, N., Chu, S. Y., Garcia, S., Schnitzler, T., Pappamihiel, N. E., Moreno, M., Mensah, F. M., Vatalaro, A., Szente, J., Levin, J., Buechner, B., Dirkx, J., Konvisser, Z. D., Myers, D., Peleg-Baker, T., … Wikan, G. (2019). Culturally Responsive Teaching Efficacy Beliefs Of In-Service Special Education Teachers. *Journal of Hispanic Higher Education*, *10*(2), 993–1013. https://doi.org/10.1108/IJSHE-05-2019-0168

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

Quinlan, D., Vella-Brodrick, D. A., Gray, A., & Swain, N. (2019). Teachers Matter: Student Outcomes Following a Strengths Intervention are Mediated by Teacher Strengths Spotting. *Journal of Happiness Studies*, *20*(8), 2507–2523. https://doi.org/10.1007/s10902-018-0051-7

Rabbitt, M. P. (2018). Causal Inference With Latent Variables From The Rasch Model As Outcomes. *Measurement: Journal of the International Measurement Confederation*, *120*(January), 193–205. https://doi.org/10.1016/j.measurement.2018.01.044

Rauch, D. P., & Hartig, J. (2010). Multiple-Choice Versus Open-Ended Response Formats of Reading Test Items: A Two-Dimensional IRT Analysis. *Psychological Test and Assessment Modeling*, *52*(4), 354–379.

Rizbudiani, A. D., Jaedun, A., Rahim, A., & Nurrahman, A. (2021). Rasch Model Item Response Theory (IRT) To Analyze The Quality of Mathematics Final Semester Exam Test on System of Linear Equations In Two Variables (SLETV). *Al-Jabar : Jurnal Pendidikan Matematika*, *12*(2), 399–412. https://doi.org/10.24042/ajpm.v12i2.9939

Runnels, J. (2012). Using The Rash Model to Validate A Multiple Choice English Achievement Test. *International Journal of Language Studies*, *6*(4), 141–155.

Rusmansyah, Almubarak, Hamid, A., & Analita, R. N. (2021). Analyze Mental Model of Prospective Chemistry Teachers With Chemical Representation Teaching Material Based on 8E Cycle Learning Model. *AIP Conference Proceedings*, *2331*, 0–7. https://doi.org/10.1063/5.0041732

Ryan, S., & Herrington, D. G. (2014). Sticky Ions : A Student-Centered Activity Using Magnetic Models to Explore the Dissolving of Ionic Compounds. *Journal of Chemical Education*, *91*, 860–863.

Setyaningsih, E., Agustina, P., Anif, S., Ahmad, C. N. C., Sofyan, I., Saputra, A., Salleh, W. N. W. M., Shodiq, D. E., Rahayu, S., & Hidayat, M. L. (2022). PBL-STEM Modul Feasibility Test for Preservice Biology Teacher. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *4*(2), 118–127. https://doi.org/10.23917/ijolae.v4i2.15980

Sihombing, R. U., Naga, D. S., & Rahayu, W. (2018). a Rasch Model Measurement Analysis on Science Literacy Test of Indonesian Students: Smart Way to Improve the Learning Assessment. *Indonesian Journal of Educational Review*, *6*(1), 44–55.

Sulistyanto, H., Anif, S., Sutama, S., Narimo, S., Sutopo, A., Haq, M. I., & Nasir, G. A. (2022). Education Application Testing Perspective to Empower Students' Higher Order Thinking Skills Related to The Concept of Adaptive Learning Media. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *4*(3), 257–271.

Sumintono, B. (2018a). *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. (2018b). *Rasch Model Measurements as Tools in Assesment for Learning*. *October 2017*. https://doi.org/10.2991/icei-17.2018.11

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*. Penerbit Trim Komunikata.

Supriyanto, S., Munadi, S., Daryono, R. W., Tuah, Y. A. E., Nurtanto, M., & Arifah, S. (2022). The Influence of Internship Experience and Work Motivation on Work Readiness in Vocational Students: PLS-SEM Analysis. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, *5*(1), 32–44. https://doi.org/10.23917/ijolae.v5i1.20033

Talib, A. M., Alomary, F. O., & Alwadi, H. F. (2018). Assessment of Student Performance for Course Examination Using Rasch Measurement Model: A Case Study of Information Technology Fundamentals Course. *Education Research International*, *2018*. https://doi.org/10.1155/2018/8719012

Valtonen, T., Leppänen, U., Hyypiä, M., Kokko, A., Manninen, J., Vartiainen, H., Sointu, E., & Hirsto, L. (2021). Learning Environments Preferred By University Students: A Shift Toward Informal and Flexible Learning Environments. *Learning Environments Research*, *24*(3), 371–388. https://doi.org/10.1007/s10984-020-09339-6

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 der Lans, R. M., van de Grift, W. J. C. M., & van Veen, K. (2018). Developing an Instrument for Teacher Feedback: Using the Rasch Model to Explore Teachers' Development of Effective Teaching Strategies and Behaviors. *Journal of Experimental Education*, *86*(2), 247–264. https://doi.org/10.1080/00220973.2016.1268086

Wikipedia. (2023). *Nitrogen*. Wikimedia.

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

Winarti, A., Almubarak, & Annurc, S. (2019). How Does The Rasch Model Justify Multiple Choice Question Items As A Measure of Student Understanding Of Acid-Base Material At The Sub-Microscopic Level? *International Journal of Innovation, Creativity and Change*, *7*(11), 344–360.

Yessi, M. (2021). Analisis Literasi Digital Peserta Didik Melalui Pemanfaatan Media Pembelajaran Berbasis Android Smart Apps Creator (SAC) Dan Instagram Dalam Pembelajaran Koloid. *Jurnal Riset Pendidikan Kimia*, *7*(1), 38–51.

Yudha, R. P. (2023). Higher Order Thinking

Skills (HOTS) Test Instrument: Validity and Reliability Analysis With The Rasch Model. *Eduma : Mathematics Education Learning and Teaching*, *12*(1), 21. https://doi.org/10.24235/eduma.v12i1.9468