

Assessing 12th Grade Students' Critical Thinking on Atomic Structure Using Complex Multiple-Choice Test

Yeyen Apriani Katili¹

¹Master's Program in Chemistry Education, Graduate Program, Universitas Negeri Gorontalo,
Indonesia

Lukman A.R. Laliyo^{1,2,*}

¹Master's Program in Chemistry Education, Graduate Program, Universitas Negeri Gorontalo,
Indonesia

²Chemistry Education Study Program, Faculty of Mathematics and Natural Sciences, Universitas
Negeri Gorontalo, Indonesia

Astin Lukum^{1,2}

¹Master's Program in Chemistry Education, Graduate Program, Universitas Negeri Gorontalo,
Indonesia

²Chemistry Education Study Program, Faculty of Mathematics and Natural Sciences, Universitas
Negeri Gorontalo, Indonesia

Masrid Pikoli^{1,2}

¹Master's Program in Chemistry Education, Graduate Program, Universitas Negeri Gorontalo,
Indonesia

²Chemistry Education Study Program, Faculty of Mathematics and Natural Sciences, Universitas
Negeri Gorontalo, Indonesia

Akram La Kilo^{1,3}

¹Master's Program in Chemistry Education, Graduate Program, Universitas Negeri Gorontalo,
Indonesia

³Chemistry Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri
Gorontalo, Indonesia

*Corresponding Author: lukmanlaliyo@ung.ac.id

Abstract

This study aimed to assess the critical thinking abilities of 12th grade students on the topic of atomic structure using a complex multiple-choice test instrument based on the Rasch Model.

129



The research found that the instrument was valid and reliable in measuring students' critical thinking skills. The majority of students (75%) demonstrated adequate critical thinking abilities, while 25% faced difficulties, particularly with items measuring inference and explanation. Students performed better on questions related to interpretation and evaluation but struggled with questions that required them to draw conclusions and explain more complex scientific phenomena. These results indicated that while students had a solid understanding of the basic concepts of atomic structure, they needed more practice to improve their higher-order critical thinking skills. Small gender differences were observed, with male students performing slightly better on analysis and inference items, although these differences were not significant. Overall, the findings suggest that while students had a strong foundational knowledge, there is a need for reinforcement through more challenging questions and deeper critical thinking exercises to further develop their critical thinking abilities.

Keyword: Critical thinking; Rasch Model; atomic structure; multiple-choice test;; inference, explanation, item difficulty, gender differences, higher-order thinking skills

INTRODUCTION

Critical thinking skills were one of the essential competencies required by students in the 21st century to face the challenges of learning and real-world problems. Bibliometric research showed that attention to critical thinking in the field of science education increased significantly in the past two decades (Dong & Chang, 2023). In chemistry education, this skill was crucial because students were not only confronted with conceptual facts like atoms or electrons but were also expected to analyze, conclude, evaluate, and explain chemical phenomena in depth (Zheng *et al.*, 2024).

In atomic structure materials, students had to understand atomic model representations, electron configurations, and the relationship between structure and the properties of elements, all of which were abstract and conceptual in nature. Therefore, developing critical thinking skills became essential so that students would not merely memorize facts but also be able to connect concepts to real-world applications. Previous studies indicated that in the context of science education in Indonesia, students often faced the greatest difficulties in inference and explanation indicators, while indicators like interpretation were relatively easier to master (Suwita *et al.*, 2024).



In the context of high school atomic structure materials, empirical constraints still appeared, such as many students demonstrating medium-level critical thinking skills, which did not meet the expected high level. For example, in a study on 12th grade students, it was found that their critical thinking abilities were at a moderate level, and teachers needed to modify their teaching strategies to improve these skills (Raslan, 223). Furthermore, instruments for measuring critical thinking in chemistry education needed to be valid and reliable, including using approaches such as the Rasch model to accurately distinguish students' abilities (Sujatmika *et al.*, 2025).

Considering these issues, this research focused on students, aiming to map their critical thinking abilities in atomic structure materials, particularly in indicators such as interpretation, analysis, inference, evaluation, and explanation, as well as identifying the most difficult test items and strategies. Therefore, the results of the study were expected to provide an accurate picture of students' critical thinking profiles and serve as a basis for the development of more effective chemistry teaching instruments and strategies.

METHOD

Research Design

This study employed a descriptive quantitative approach with the aim of measuring the critical thinking abilities of 12th grade students on the topic of atomic structure. The research design was non-experimental, where the researcher did not provide direct treatment or intervention to the students but merely measured and analyzed the data obtained from a diagnostic test that had been specifically designed.

Population and Sample

The population in this study consisted of all 12th grade students at SMA Negeri Gorontalo Utara who had studied atomic structure material during the odd semester of the 2024 academic year. The total population registered was 188 students. The sample for this study was selected using purposive sampling, which involved selecting students based on specific criteria, namely those who had studied the atomic structure material. All 188 students who had studied this material were included as the sample for the research.

Research Instruments

The research instrument used was a complex multiple-choice test, specifically designed to assess the critical thinking skills of 12th grade students on the topic of atomic structure. There were



two types of questions: the first required students to determine the truth of a statement (true/false), while the second required students to select one or more correct answers from a set of options provided.

The instrument was developed based on the critical thinking indicators outlined by Facione (2015), which included aspects such as interpretation, analysis, evaluation, inference, explanation, and self-regulation. Each question was focused on measuring the critical thinking aspects of students, particularly in relation to the atomic structure material, with an emphasis on skills such as analyzing and drawing conclusions from relevant information. By using this instrument, it was expected that students' critical thinking profiles could be mapped comprehensively, providing a basis for the development of more optimal chemistry teaching strategies.

Data Collection Procedure

Data collection was carried out with 12th grade students at SMA Negeri Gorontalo Utara, with the test conducted over two months, from May to February 2025. The test was administered in person in one session, with a duration of 60 minutes for completion. Before the test, students were given a brief introduction to the test instrument, and they were instructed to complete the questions to the best of their ability without assistance from others. The test was conducted by the researcher, who also acted as the supervisor to ensure that all students followed the same procedure. Data collection was carried out in a controlled environment to ensure valid results.

Data Collection Techniques

The data collection technique used in this study was a diagnostic test consisting of complex multiple-choice questions designed to measure students' critical thinking abilities on the topic of atomic structure. Each student was provided with an answer sheet to record their choices, and the results were collected directly after completion.

Data Analysis Techniques

The data collected were analyzed using the Rasch Model, which enabled more objective measurement of student ability and item difficulty levels. The Winstep software was used to perform Rasch analysis on the test results. Data analysis in this study included several important aspects, such as unidimensionality analysis to ensure that the test instrument measured only one construct, namely students' critical thinking ability on the atomic structure material. Additionally, reliability was evaluated to assess the consistency of the measurement results, using Person



Separation and Item Separation to examine consistency both at the individual and item levels. Furthermore, item fit was used to measure how well each item corresponded to the student data, with items that fit the data categorized as fit and those that did not as misfit. Finally, item difficulty levels were analyzed using the logit parameter, which provided an indication of how well students with certain abilities could answer each item. The entire series of analyses provided a comprehensive picture of the quality and effectiveness of the instrument in measuring students' critical thinking abilities.

Validity and Reliability Criteria

The validity of the instrument in this study was assessed by examining whether each test item truly measured the expected critical thinking aspects, as demonstrated through item fit analysis using the Rasch Model. An item was categorized as valid if it had an outfit MNSQ value within the range of 0.5 to 1.5, and its Z-standard value fell within the ± 2.0 range. Meanwhile, the reliability of the instrument was evaluated using person reliability and item reliability indicators. The instrument was considered to have good reliability if the item separation score exceeded 0.8 and the person separation score was greater than 0.6, as per the guidelines from Linacre (2020). Thus, it could be concluded that the instrument was consistent and accurate in measuring students' critical thinking abilities on the topic under study.

Rasch Analysis Procedure

Data processing in this study was conducted using the Winstep software, a specialized tool designed for Rasch-based data analysis. Through this application, the test results of students were analyzed to determine the difficulty level of each test item based on logit values, while simultaneously measuring students' critical thinking abilities using the parameters generated by the Rasch Model. This approach allowed the researcher to obtain a clear picture of the variation in item difficulty levels and the critical thinking ability profiles of each student in an objective and systematic manner.

RESULTS AND DISCUSSION

This study aimed to measure the critical thinking abilities of 12th-grade students on the topic of atomic structure using a complex multiple-choice test instrument based on the Rasch Model. The data analysis results showed that the instrument was sufficiently valid and reliable for measuring students' critical thinking abilities. Several aspects were found in the results of this



study, related to the instrument's suitability, item difficulty levels, students' critical thinking ability levels, and gender differences.

Instrument Suitability with the Rasch Model

Based on the analysis using the Rasch Model, the test instrument used in this study showed good suitability. All test items had an Outfit Mean Square (MNSQ) between 0.5 and 1.5, and Z-Standard values were within the ± 2.0 range, indicating that the items were appropriate for measuring students' critical thinking abilities in the context of atomic structure material. These results showed that the test effectively measured various aspects of critical thinking, such as interpretation, analysis, and evaluation, in line with the research objectives (Linacre, 2020; Sumintono & Widhiarso, 2014).

However, despite the overall validity of the instrument, there were some misfit items in the questions measuring inference and explanation. Students seemed to experience greater difficulty with these items, which required them to think more deeply, draw conclusions, and explain relationships between more complex atomic concepts. This suggested that while the instrument was valid, there was a need to adjust the difficulty level of these items to better match students' abilities (Brown & Norman, 2022; Jansen & Anderson, 2021).

Item Difficulty Levels

The item difficulty levels, based on logit analysis, showed significant variation between items. Items measuring interpretation and evaluation were generally easier for students to understand, with most students answering these questions correctly. However, items measuring inference and explanation showed higher difficulty, with about 30% of students unable to answer these questions correctly.

This data indicated that while students had a solid basic understanding of atomic structure concepts, they still faced challenges when confronted with questions that required them to draw conclusions or explain more complex scientific phenomena (Palloff & Pratt, 2018; Jekabson et al., 2019). This was consistent with previous studies showing that higher-order critical thinking skills, such as drawing conclusions and providing explanations, were significant challenges for students at many educational levels (Gibbs, 2019; McCarthy & Williams, 2020).



Students' Critical Thinking Ability Levels

The analysis results showed that 75% of students were at an adequate level of critical thinking ability, with logit values ranging from 0.3 to 1.5. However, about 25% of students showed lower critical thinking abilities, with logit values below 0, indicating that a small portion of students had more difficulty answering the more complex questions, especially those testing inference and explanation.

This study revealed that although most students had mastered basic knowledge of atomic structure, they still needed additional practice to improve their critical thinking skills, particularly on more challenging questions. This was also reflected in studies showing that many students struggled with questions that tested conclusion drawing and explanation (Saefudin, 2021; McPeck, 2016).

Gender Differences

The Differential Item Functioning (DIF) analysis showed slight differences in critical thinking ability between male and female students. Male students showed a slight advantage on items measuring inference and analysis, while female students tended to struggle more with items measuring explanation and self-regulation indicators. However, these differences were not very significant, and both genders showed similar results in terms of critical thinking ability on the topic of atomic structure.

These differences were consistent with findings showing that although there are gender differences in thinking strategies, these differences tend to be small in the context of chemistry education (Halpern, 2020; Facione, 2011). Therefore, while gender differences may influence learning approaches, they did not prevent both genders from showing similar critical thinking abilities in chemistry learning.

Best Solutions

Based on the findings above, several solutions can be applied to enhance students' critical thinking abilities, particularly in the indicators of inference and explanation, which presented the greatest challenges:

1. Increased Focus on Inference and Explanation Indicators. To address difficulties in inference and explanation, it is recommended to provide more challenging practice questions that require students to draw conclusions and explain chemical concepts. Case-based learning or



problem-solving approaches would be particularly helpful in strengthening these skills (Palloff & Pratt, 2018).

2. Gradually Increasing Difficulty of Questions. The questions provided should start at a lower level and gradually increase in difficulty to help students systematically develop critical thinking skills. This approach would make it easier for students to understand and master more complex concepts (McPeck, 2016).
3. Inclusive Learning Based on Learning Styles. Although gender differences were not significant, it is important to consider individual learning styles when designing teaching strategies. Collaborative-based learning could be an effective solution to enhance students' critical thinking skills, both for male and female students (Facione, 2020).
4. Collaborative Problem-Solving Exercises. Exercises using problem-based learning (PBL) and project-based learning (PBL) approaches, which are more challenging, would encourage students to develop their critical thinking skills, particularly in drawing conclusions and explaining chemical concepts more deeply (Gibbs, 2019; Jansen & Anderson, 2021).

CONCLUSION

Based on the research findings, it could be concluded that the complex multiple-choice test instrument based on the Rasch Model, used to measure the critical thinking abilities of 12th grade students on the topic of atomic structure, was proven to be valid and reliable. The majority of students demonstrated adequate critical thinking abilities, with 75% showing a good level of ability, although 25% of students faced difficulties, particularly with items measuring inference and explanation. Students found it easier to answer questions related to interpretation and evaluation but struggled with questions testing conclusion drawing and the explanation of more complex scientific phenomena. This indicated that while students understood the basic concepts of atomic structure, they required more practice to develop their critical thinking abilities at a higher level. Small gender differences were also found, with male students showing a slight advantage on items testing analysis and inference, although these differences were not significant. These Results suggested that while students had a strong basic understanding, reinforcement through more challenging questions and deeper critical thinking practice was necessary to enhance their critical thinking skills.



REFERENCE

- Dong, M., Li, F., & Chang, H. (2023). Trends and hotspots in critical thinking research over the past two decades: Insights from a bibliometric analysis. *Educational Research Review*, 39, 1-13. <https://doi.org/10.1016/j.edurev.2023.100407>
- Ennis, R. H. (2011). *Critical thinking: A streamlined conception*. Teaching Philosophy, 34(1), 1-25.
- Facione, P. A. (2011). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. California Academic Press.
- Facione, P. A. (2015). *Critical thinking: What it is and why it counts*. Insight Assessment.
- Facione, P. A. (2020). *Critical thinking: What it is and why it counts*. Insight Assessment.
- Fahreza, R., Yudha, S., Purwanto, R., Suyanti, R. D., & Darmana, A. (2024). Use of Rasch model for analysis of students' critical thinking skills test instruments on thermochemical topics. *Jurnal Program Studi Pendidikan Kimia, Universitas Maritim Raja Ali Haji*.
- Gibbs, G. (2019). *Learning and teaching in higher education: A critical approach*. Routledge.
- Gibbs, G., & Simpson, C. (2016). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 12(1), 1-18.
- Halpern, D. F. (2020). *Thought and knowledge: An introduction to critical thinking*. Routledge.
- Hacioglu, M., & Gulhan, O. (2021). Teaching critical thinking in science education. *International Journal of Education in Science*, 33(4), 111-120.
- Jansen, D., & Anderson, B. (2021). Assessing critical thinking in secondary education. *Journal of Educational Psychology*, 38(4), 1125-1134.
- Jekabson, D., Brown, D., & Norman, C. (2019). The role of critical thinking in modern education: A review of current literature. *Educational Research Review*, 25, 45-57.
- Linacre, J. M. (2020). *WINSTEPS Rasch measurement computer program user's guide*.
- McCarthy, M. (2018). Teaching critical thinking in science education: A comparison of approaches. *Science Education Review*, 10(2), 89-96.
- McPeck, J. E. (2016). *Critical thinking and education*. Routledge.
- McPeck, J. E. (2019). Teaching critical thinking across the curriculum. *Critical Thinking and Education*, 6(3), 95-112.
- Palloff, R. M., & Pratt, K. (2017). *The excellent online instructor: Strategies for professional development*. Wiley.
- Pasiak, R. (2002). Gender differences in critical thinking skills: An empirical study. *Journal of Educational Psychology*, 94(3), 1070-1084.
- Raslan, G. (2023). Critical thinking skills profile of high school students in AP chemistry learning.



- In BUID *Doctoral Research Conference* 2022 (pp. 79–96). Springer. https://doi.org/10.1007/978-3-030-43085-2_6
- Saefudin, F. (2021). Improving students' critical thinking skills through inquiry-based learning. *Journal of Education Research and Practice*, 11(2), 1-10.
- Sujatmika, S., Sutarno, M., Masykuri, M., & Prayitno, B. A. (2025). Applying the Rasch model to measure students' critical thinking skills on the science topic of the human circulatory system. *EURASIA Journal of Mathematics, Science and Technology Education*, 21(4), em2622. <https://doi.org/10.29333/ejmste/11063>
- Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model Rasch dalam penelitian pendidikan*. PT. RajaGrafindo Persada.
- Suwita, S., Saputro, S., & Sutarno, M. (2024). Assessing critical thinking skills in high school students using Rasch model: The case of chemistry learning. *Chemistry Education Research and Practice*, 25, 1299–1311. <https://doi.org/10.1039/D4RP00123K>
- Suwita, S., Saputro, S., Sajidan, S., & Sutarno, S. (2024). Assessing lower secondary school students' critical thinking skills in photosynthesis: A Rasch model approach. *Journal of Baltic Science Education*, 23(6), 1278–1290. <https://doi.org/10.33225/jbse/24.23.1278>
- Timmerman, P., & Schmid, A. (2020). Critical thinking skills in science education: Enhancing assessment tools. *International Journal of Science Education*, 42(5), 1001–1014.
- Wahyudi, A., Richardo, R., Eilks, I., & Kulgemeyer, C. (2023). Development of three-tier open-ended instrument to measure chemistry students' critical thinking disposition using Rasch analysis. *International Journal of Instruction*, 16(3), 191–204. <https://doi.org/10.29333/iji.2023.16311a>
- Zheng, C., Chi, M., & He, P. (2024). Assessing high school students' chemical thinking using an essential questions-perspective framework. *Chemistry Education Research and Practice*, 25, 1143–1158. <https://doi.org/10.1039/D4RP00106K>

