



# European Journal of Educational Research

Volume 10, Issue 4, 1769 - 1779.

ISSN: 2165-8714

<http://www.eu-jer.com/>

## A Rasch Analysis of Item Quality of the Chemical Literacy Assessment for Investigating Student's Chemical Literacy on Chemical Rate Concepts

**Apriliya Dwi Setyorini\***   
Universitas Sebelas Maret,  
INDONESIA

**Sri Yamtinah**   
Universitas Sebelas Maret,  
INDONESIA

**Lina Mahardiani**   
Universitas Sebelas Maret,  
INDONESIA

**Sulistyo Saputro**   
Universitas Sebelas Maret,  
INDONESIA

*Received: January 9, 2021 • Revised: May 29, 2021 • Accepted: August 10, 2021*

**Abstract:** Assessment is a topic that continues to be developed in science education research. Assessment evaluates not only students' cognitive abilities but also their thinking skills. Therefore, in this study, an assessment that could measure students' chemical literacy was developed. Chemical literacy is a thinking skill that students must develop as part of their chemistry learning. The goal of this study was to assess item quality, as well as student chemical literacy on the concept of chemical rate. The Rasch model was employed to analyze the data in this study. The results of this study depict that the developed assessment had sufficient reliability and validity to be used to assess students' chemical literacy. Furthermore, the analysis of the students' responses to the items revealed that many students did not understand or were unaware of the context presented. These findings suggest that students' chemical literacy in the material for the reaction rate is still lacking and needs to be improved. As a result, the teacher's role in assisting students in improving their chemical literacy through chemistry learning is critical.

**Keywords:** *Assessment, chemical rate, Rasch analysis, students' chemical literacy.*

**To cite this article:** Setyorini, A. D., Yamtinah, S., Mahardiani, L., & Saputro, S. (2021). A Rasch analysis of item quality of the chemical literacy assessment for investigating student's chemical literacy on chemical rate concepts. *European Journal of Educational Research*, 10(4), 1769-1779. <https://doi.org/10.12973/eu-jer.10.4.1769>

### Introduction

Assessment is one of the topics that are frequently discussed in science education research. The assessment is a process in which teachers and students work together to produce a result that will be used for feedback and the improvement of future learning (Black & Wiliam, 1998). Because students' understanding of concepts changes over time, assessment is always discussed. Therefore, in order to diagnose students, an assessment that can provide information about students' level of understanding, whether they already understand, misunderstand, or even lack conceptions, must be developed (Johnstone et al., 1977). Diagnostic tests are typically undertaken in various formats, the most common of which is multiple-choice. This is because multiple-choice questions can provide a quick analysis and a reliable score (Caleon & Subramaniam, 2010; McClary & Bretz, 2012). Assessment to diagnose students' cognitive understanding has made a lot of progress. However, the assessment can also be designed to assess students' thinking skills (Sadler & Zeidler, 2009).

Currently, only a few assessments can be used to analyze the thinking skills required by students in the 21<sup>st</sup> century. One of the thinking skills that students need to have is chemical literacy, which is a part of scientific literacy. Similar to scientific literacy, chemical literacy must also be improved by learning chemistry (Cigdemoglu et al., 2017). There have been many studies on chemical literacy as knowledge and skills that must be present in learning chemistry. However, research on how literacy is taught and measured is considered new (Celik, 2014). Thus, it is necessary to develop an assessment to evaluate students' chemical literacy (Shwartz et al., 2006).

In this study, this assessment was developed to analyze students' chemical literacy, by following the dimensions of chemical literacy by Shwartz et al. (2006), including chemical content knowledge, chemistry in context, high-order learning skills, and affective aspects. The assessment was developed with multiple-choice questions, which could give simple and quick analysis, be objective, and be used to analyze students' conceptions. Only three dimensions were used in this assessment: chemical content knowledge, chemistry in context, and high-order learning skills. The affective aspect was not included because it differed from the others, necessitating the use of other assessments to evaluate it.

#### \* Corresponding author:

Apriliya Dwi Setyorini, Universitas Sebelas Maret, Chemistry Education Department, Postgraduate Program, Indonesia. ✉ [apriyadwi47@gmail.com](mailto:apriyadwi47@gmail.com)

Moreover, tools for analyzing student test results are required to obtain good test analysis results. The Rasch model is one of the tools that can be used to analyze the student's response. This model is useful since it can analyze students' responses after taking the test based on the difficulty level of the item as well as the level of students' ability (Englehard, 2013). The Rasch model is also distinct in that the data used are calibrated to take the form of interval data, the relationship between the difficulty of the item, and the students' ability to calculate the Rasch model independently (Bond & Fox, 2012; Boone et al., 2014). Furthermore, many people believe that Rasch analysis is a suitable approach to measurement in the field of social science (Bond & Fox, 2012; Boone et al., 2014; Liu, 2010).

The assessment developed in this study was limited to the chemical rate concepts. This is due to the fact that the chemical rate is one of the concepts that is both difficult to teach and challenging for students to grasp (Cakmakci et al., 2006). Besides, the chemical rate is frequently misunderstood (Bain & Towns, 2016; Chairam et al., 2015). Therefore, it is essential to conduct a proper assessment of the students' understanding of this chemical rate concept. The study aims to assess the item' quality as well as to determine the students' chemical literacy on the chemical rate concept.

## Literature Review

### *Chemical Literacy*

Science literacy is one of the primary goals of science education and is an essential component of science education worldwide (American Association for the Advancement of Science [AAAS], 1993; Coll & Taylor, 2009; DeBoer, 2000; National Research Council [NRC], 1996). Scientific literacy is essential for students to become well-informed about how to perceive science issues in order to make decisions in a variety of subjects that affect their daily lives (Dijk, 2011; Tal & Dierking, 2014). The PISA 2015 framework refers to the importance of understanding not only scientific ideas and theories but also general methods associated with scientific research and how this allows science to improve (De Jong, 2012). Someone with scientific literacy should understand scientific knowledge, science application, cognitive skills, the ability to solve problems using scientific knowledge, the nature of science, the ethics of scientists, as well as their relevance to culture, society, and technology (Barlia, 2016; Holbrook & Rannikmae, 2007; Liu, 2009; National Research Council [NRC], 2007; Norris & Phillips, 2003; Preczewski et al., 2009). Science literacy has been recognized as important in our current civilization since it has been twisted with science and technology. Furthermore, science literacy includes the following components: (1) understanding natural science concepts, (2) understanding and critically judging scientific content, and (3) preparing members of society to face real-life scientific and technological contexts (Miller, 1983; Norris & Phillips, 2003; Shwartz et al., 2006). Because scientific literacy is a broad concept, teaching any specific subject in science learning should provide students with the opportunity to practice to be scientifically literate. Teaching chemistry develops students' chemistry literacy, as well as their scientific literacy in general (Shwartz et al., 2006).

Chemistry is an important branch of science. Chemistry topics typically include learning materials that are relevant to many disciplines, such as health, geography, physics, environment, and economics (Brown et al., 2000). Understanding chemistry is also very significant for most people because chemicals and chemical products are related to daily lives (Gilbert & Treagust, 2009). In addition, chemical literacy entails knowledge of the nature of matter, chemical reactions, laws, and chemical theory, as well as general chemical applications in everyday life (Barnea et al., 2010). Someone who is chemically literate is someone who can apply knowledge of chemistry in everyday situations (Tsaparlis, 2000). Thus, understanding chemistry will enable people to participate in public debates and solve problems in their daily lives. Chemical literacy involves knowledge of chemistry as well as the necessary skills for a chemistry-based understanding of social and scientific issues. Chemical literacy consists of three components: (1) key concepts in general chemistry, such as elements, symbols, processes, and models; (2) understanding of what expert chemists in academia and industry do; and (3) social context, or the ability to place chemistry in a real-world context (Dori et al., 2018; Holman & Hunt, 2002; Shwartz et al., 2013). Moreover, chemical literacy encompasses an affective component, which refers to an interest in chemistry-related topics (Shwartz et al., 2006). Because chemical literacy is an important aspect or skill for students when learning chemistry, it is necessary to assess their chemical literacy. The purpose of this assessment is to determine the level of chemical literacy among students. As a result, teachers will receive feedback on how to improve chemistry learning, which will help students improve their chemical literacy. Chemistry exams used to assess students' chemical literacy must include dealing with given information in a chemistry problem, as well as students' knowledge of how to use chemistry and skills to include information from an everyday context (Witte & Beers, 2003). In these assessments, the context can be an industrial process, an environmental issue, or any other issues from school, home, or science in laboratories.

### *Rasch Model*

Rasch model is a measurement method that is part of a group called item response theory (IRT) and is frequently used to analyze data in educational research (Khine, 2020). Rasch model analysis provides an alternative method for investigating the psychometric properties of measures and addressing response bias (Bradley et al., 2015). The analysis approach presented by this model can be applied to promote the test item as an essential tool in assessment for learning (Suminto & Widhiarsho, 2015). The data are analyzed using WINSTEPS (Linacre, 2011) to estimate students

and item measures. For the dichotomous data in the Rasch model, the probability that a student will answer an item correctly is determined by the variation in the level of student's performance and item's difficulty (Bond & Fox, 2012). The Rasch model produces an interval scale with units of "log-odds," which is also known as logits.

Rasch analysis measures item difficulty as an independent unit. Item difficulty is measured on the same interval scale as students' ability, allowing item targeting to be estimated. The Rasch model determines the opportunity that a student will correctly respond to an item based on the difference between the student's overall performance level and the difficulty of the item. This is measured using the following equation:

$$\ln\left(\frac{P_{ni}}{1 - P_{ni}}\right) = B_n - D_i$$

where  $P_{ni}$  is the probability that student  $n$  of overall performance level  $B_n$  answers correctly to item  $i$  with a difficulty of  $D_i$  (Bond & Fox, 2012; Liu & Boone, 2006). How well the items predict student responses based on ability and difficulty measures are delivered by fit statistics (Linacre, 2011). Fit statistics are used for both item difficulty and students' ability measures to the Rasch model and a powerful tool for deciding item performing for a given target population. Items that give unexpected student responses and students who give unexpected answers can be identified by fit statistics.

In addition to the fit statistic, the option probability curve depicts the likelihood that the student will select each answer option presented in the Rasch analysis. It is critical to examine the curve for each option in multiple-choice items because the shape of the curve provides information about the students' overall understanding. The response options reveal information about students' comprehension of the context. Analyzing the option probability curve for each answer choice thus provides more qualitative information about the student's understanding of each context on the item (Herrmann-abell & Deboer, 2016).

### Methodology

This study employed a mixed-methods research design, combining quantitative and qualitative research (Creswell, 2006). This study began with the development of item questions based on the chemical literacy dimensions, which were then randomly tested on students. The results were analyzed using the Rasch model to determine the feasibility of the developed assessment as well as to analyze students' performance.

#### *Item Development*

Shwartz et al. (2006) developed this assessment based on the dimensions of chemical literacy. The first dimension is understanding chemical content. Students who can explain chemical processes and energy changes at the macroscopic and microscopic levels using chemical language can also perform scientific investigations based on their knowledge in this dimension. The items in the content knowledge dimension consist of chemical reactions related to the graph of the reaction rate and the effect of concentration on the reaction rate. Furthermore, the second dimension is chemistry in context. Students can use their chemistry knowledge to explain everyday phenomena and make effective decisions in social situations involving chemistry. Moreover, the items in the dimension of chemistry in context include the examples of chemical phenomena in everyday life. The third dimension is higher-order learning skills. Students can ask questions, investigate information, and weigh the benefits and drawbacks of a phenomenon related to their chemical knowledge. The items created were derived from three dimensions and the questions were designed using the chemical rate concept. The items on the high-order learning skills dimension cover the examples of phenomena in nature and industry, as well as their roles in the human body related to chemical rates.

Also, the items developed were validated by high school chemistry teachers, chemistry education lecturers, and educational evaluation experts. This assessment covered 14 multiple-choice items that had been validated by nine experts. The validation was performed to determine the validity content of the items. The most common method for measuring content validity is the Item-level CVI (I-CVI), while the other method to measure content validity is Scale-level CVI (S-CVI), which can be determined using Universal Agreement among expert (S-CVI/UA) or the Average CVI (S-CVI/Ave) (Lynn, 1986). The I-CVI is considered excellent if its value is 0.78. The S-CVI with a value between 0.80 and 0.90 is acceptable. The computation analysis yielded an average I-CVI of 0.98; S-CVI/UA of 0.85; and S-CVI/Ave of 0.98. These results indicate that the constructed items have excellent and acceptable content validity.

These questions were used to assess students' chemical literacy as well as their understanding of the chemical rate concept. These items were classified into four items of chemical content knowledge, four items of chemistry in context dimension, and six items of high-order learning skills. The assessment was then administered to students and analyzed using the Rasch model. Table 1 depicts the distribution of items.

Table 1. Distribution of Items on Chemical Literacy Dimensions

Dimension	Number of Item
Chemical content knowledge	1,2,5,6
Chemistry in context	7,8,9,10
High-order learning skills	3,4,11,12,13,14

### Participants

The participants in this study were 277 students of 11<sup>th</sup> grade from three different schools taken with random sampling. The schools are located in Surakarta, Central Java, Indonesia. The 11<sup>th</sup>-grade students were chosen because the chemical rate concept was given in this grade.

### Data Collection

Participants in this study were chosen randomly. The students participating in this study received previous learning chemical rate concept as well. This test was given immediately after participants had learned the chemical rate concept taught by each teacher at the representative school.

### Data Analysis

The data were analyzed using the Winstep software (Linacre, 2006). The analysis results were used to compare students' responses and the items' difficulty level. The easier item had the lower logit value, while the more difficult item had higher logit. This is also portrayed on the items' Wright map. The easier items are at the bottom, while the more difficult items are at the top. The Rasch model would be used to quantify the item quality, the item and person reliability, item and person separation, and Alpha Cronbach value would be presented, as well as the item validity in the form of the infit value and the outfit MNSQ, Pt-measure correlation value, and unidimensionality. Besides, students' chemical literacy was qualitatively presented on probability items. The students' choice indicated their understanding of the context item.

## Results

### The Quality of the Items of Chemical Literacy Assessment

#### Reliability

This assessment's reliability was divided into the person and item reliability, as well as Cronbach's Alpha value. This assessment was also used to test the separation. Table 2 shows the reliability of this assessment.

Table 2. Reliability of Chemical Literacy Assessment

	Person	Item
Reliability	0.79	0.99
Separation	1.93	8.50
Alpha Cronbach	0.85	

Table 2 presents that the Cronbach Alpha value was 0.85, indicating that the developed items were reliable. An assessment resulting a Cronbach alpha value of 0.7 indicates acceptable level of reliability, while an alpha value of 0.8 is good and 0.9 is very good (Bond & Fox, 2012). The person's reliability was 0.79, signifying that the participants in this study were diverse and reliable. Meanwhile, the item reliability was 0.99, suggesting that this item was highly trustworthy. Moreover, separation of person was considered quite good because the value was close to two, and separation of item was very good because the value was greater than three. A good separation value is >2 for the person and >3 for the items (Bond & Fox, 2012; Duncan et al., 2003).

#### Validity

Validation is the process of seeking empirical evidence from the assessment developed (Stevens et al., 2010). The item fit statistic demonstrates the Rasch model validation. The fit statistical items for this assessment are presented in Table 3.

Table 3. Item Fit Statistic of Chemical Literacy Assessment

Item	Item logit	Standard error	Infit MNSQ	Outfit MNSQ	Pt-measure correlation
8	3.40	0.25	0.84	0.46	0.56
14	1.64	0.17	1.29	1.37	0.45
7	1.27	0.16	0.93	0.80	0.60
4	0.73	0.15	1.13	1.23	0.53
3	0.52	0.15	0.98	0.91	0.60
12	0.47	0.15	1.31	1.54	0.47
13	0.47	0.15	0.79	0.63	0.67
11	0.12	0.15	0.84	0.77	0.66
1	-0.59	0.16	0.82	0.71	0.66
2	-0.96	0.16	1.48	1.69	0.40
6	-1.17	0.17	1.10	0.98	0.55
10	-1.23	0.17	0.86	0.68	0.63
5	-1.46	0.17	0.75	0.52	0.66
9	-3.21	0.24	0.72	0.29	0.56

Table 3 demonstrates that the MNSQ values, both infit and outfit, mostly ranged between 0.5 and 1.5, meaning that the developed items can measure the expected indicators. MNSQ infit and outfit values of  $0.5 < \text{MNSQ} < 1.5$  are eligible (Bond & Fox, 2012). The Pt-measure correlation was greater than 0.3, noting that these items could discriminate against students' abilities (Li et al., 2016; Linacre, 2011). Furthermore, all items had a standard error less than 0.5, indicating that the items developed were accurate. Analysis of the items revealed unidimensional properties based on the value of the raw variance explained as well as the value of the unexplained variance. The raw variance value that shows good construct validity is  $>40\%$  and the eigenvalue for each contrast is  $<2$  (Linacre, 2011). The raw variance value explained by this assessment was 44% and the unexplained variance value for each contrast on each contrast had an eigenvalue less than 2 (1.2 to 1.9), indicating that the items being developed had unidimensional properties, or they have a good performance to measure. The developed measure items were classified as easy, with negative logit, medium that close to 0, and difficult, with logit values greater than 0. This demonstrates that the items were evenly distributed, as visible on the Wright map in Figure 1.

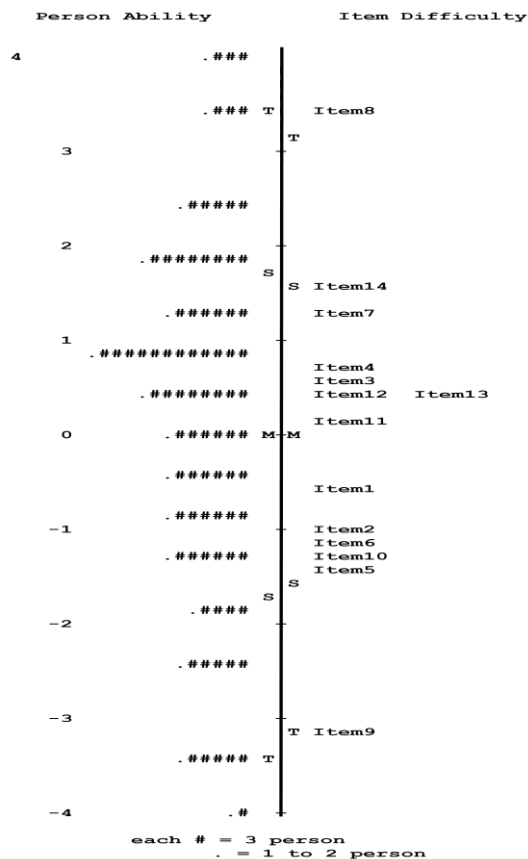


Figure 1. Wright Map of Item and Person

The questions ranged from easy to difficult, as illustrated in Figure 1. The easy items are listed at the bottom, and the difficult items are listed above. Similarly, the students' abilities are itemized on the left, where students with low abilities are listed at the bottom and those with high abilities are listed at the top. Based on these findings, the items produced could represent the students' abilities ranging from low to high.

#### Students' Chemical Literacy on Chemical Rate

The option probability of each item answered by the students revealed their responses. In this assessment, the options selected by students for each item could be analyzed as students' understandings of the chemical rate concept. The responses provided an overview of whether students perceived the concepts and contexts or whether a concept was incorrect.

The researcher selected three sample items, including item 8 (the most difficult), item 11 (the medium), and item 9 (the easiest). Items 8 and 9 were from the chemistry in context dimension, while item 11 was from the high-order learning skills dimension. Figure 2 shows the option probability for item 8.

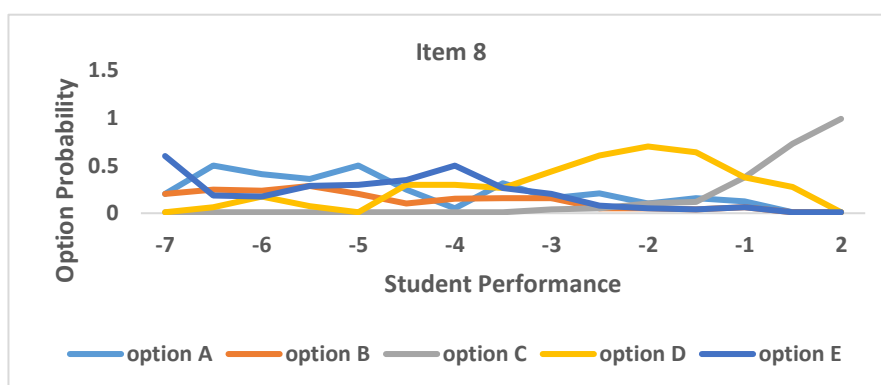


Figure 2. Option Probability of Item 8

Item 8 was the most difficult for students to complete. This item presents the avocado experiment, in which the avocado is first left at room temperature. Following that, the avocado is given an additional treatment by spraying ethylene on its surface. Students were asked to figure out why the selected conditions cause the avocado to ripen faster. Option C is the correct answer; the addition of ethylene from the outside increases the ethylene concentration on the surface of the fruit produced by the avocado, causing the fruit to ripen faster.

Students with a positive logit value are those who can work on the items being tested. Students with negative logit values include those having lower ability to work on the items being tested. Figure 2 shows that students with low understanding, those with negative logit scores, guessed a lot. This is because students who did not understand the concept were more likely to select options A, B, D, and E. Students having a high understanding with logits between -1 and -2 were more likely to choose option D. Option D was, in fact, the most popular among students. This option contains information that the addition of ethylene is for catalysts that can accelerate fruit ripening. However, this assumption is incorrect because the fruit produces its own ethylene during the ripening process, and the addition of ethylene concentration accelerates fruit ripening. As a result, many students answered item 8 incorrectly.

Further, the following is the analysis of the medium item of item 11, and the option probability of the item is presented in Figure 3.

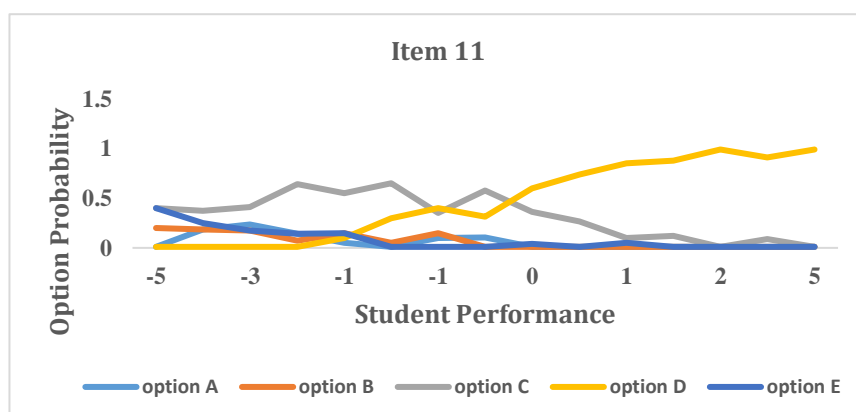


Figure 3. Option Probability of Item 11

Item 11 contains a question about the causes of acid rain, and the correct answer is option D. Students with a low level of understanding and a negative logit score guessed a lot on this item. It was indicated by the numerous probability options A, B, C, and E. Option D was chosen by more students with a higher level of understanding and a positive logit value. Option D states that acid rain is caused by human activities that increase the levels of gases such as  $\text{SO}_2$  and  $\text{SO}_3$ . Students, on the other hand, preferred option C, mentioning an increase in  $\text{CO}_2$  gas and chlorine gas. However,  $\text{CO}_2$  gas and chlorine gas are more precisely the cause of ozone layer depletion. Meanwhile, the rising levels of  $\text{SO}_2$  and  $\text{SO}_3$  gases contribute significantly to acid rain. These findings indicate that many students were unaware of acid rain. The majority of them were mindful of one of the factors influencing the chemical rate, namely the addition of concentration, but they were clueless of such context.

The next is the analysis result of item 9, the easiest item. The option probability of this item is presented in Figure 4.

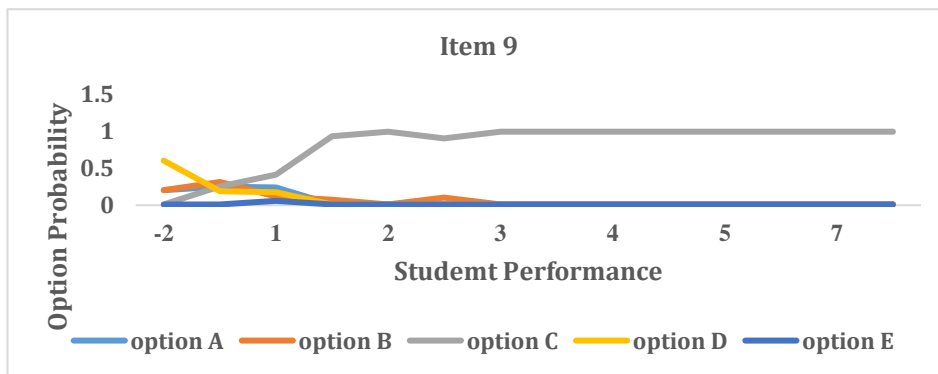


Figure 4. Option Probability of Item 9

Figure 4 shows that nearly all students selected the correct answer, option C. This item posed a question about a comparison of two cooking circumstances for vegetables. The first condition is cutting the vegetables into large pieces, and the second condition is cutting the vegetables into small pieces. The students were asked to select the condition that causes vegetables to cook faster. Students with a lack of comprehension answered this item incorrectly. A logit value of -2 indicated that the D option was frequently chosen. Option D states that if more water is added, both conditions will trigger the vegetables to cook faster, whereas Option C states that the second condition, slicing vegetables into small pieces, is the reason. Small pieces of vegetables have larger surface areas and this makes the vegetables cook quicker. Students having the understanding that adding water or increasing volume can accelerate the reaction rate provided incorrect answers because, in this context, this condition can not make vegetables cook quicker.

### Discussion

Based on the analysis, the developed assessment met the criteria in terms of reliability and validity aspects. The items had a good Cronbach alpha value, large separation, and good construct validity, as well as were quite thorough and able to discriminate against abilities. The validity was depicted from the MNSQ, Pt-measure correlation, standard error, and unidimensionality value. The MNSQ value, Pt-measure correlation, and unidimensionality value indicate that the constructed items have a good quality and can be used for measurement according to the specified indicators (Bond & Fox, 2012; Linacre, 2011). The Wright map results show a scattering of items ranging from easy to difficult. Furthermore, the population was slightly higher than the item difficulty average. The average difficulty item was 0.00, while the average student ability was 0.16, indicating that the students' ability was slightly higher than the average item difficulty level. This means that the students' level of understanding was nearly equal to the average item difficulty level. In addition, the selected respondents had varying abilities, as evidenced by the person's reliability of 0.79. A person's reliability close to or greater than 0.8 indicates that the respondent had good reliability and used a variety of abilities (Xie et al., 2014). Respondents' various abilities could also be used to evaluate the instrument's performance.

Some items, however, had lower MNSQ outfit values. This could be due to inconsistencies in student responses as well as a large number of students guessing (Boone et al., 2014). The data showed that unexplained raw variance value was 14%, meaning that 14% of the students were unpredictable in answering questions. Students' unpredictable responses reduced the instruments' fit. Nonetheless, the developed assessment was effective in measuring students' chemical literacy on the chemical rate concept.

Moreover, after being tested for this assessment, students' perceptions were found to be diverse. In item number 8, students did not understand the context and they responded that ethylene was a catalyst. This is because students understood catalysts to be materials that can speed up reactions (Kingir & Geban, 2012; Yalçinkaya et al., 2012). In this case, ethylene is added to increase the concentration of ethylene produced by the fruit itself, allowing the fruit to ripen more quickly. This means that students did not understand the context and neither did they maximize chemistry in context.

However, many students correctly answered item 9, which is a simple question. This item assessed the chemistry dimension in context as well. This suggests that people could grasp the concept that expanding the surface area of a reactant increases particle collision (Yalçınkaya et al., 2012). The process of cooking vegetables is sped up by cutting them into small pieces. Moreover, the exemplified cases are very close to the students' everyday lives. This is beneficial to students because they can observe events in their daily lives that are closely related to chemistry. A person with good chemical literacy skills will always apply his knowledge in everyday life (Holman & Hunt, 2002).

In addition, many students have misconceptions about or do not comprehend the context presented in item 11. Some students were unaware of the cause of acid rain. Despite their incorrect answers, students understood the concept of a linear relationship between concentration and reaction rate (Cakmakci et al., 2006; Turanyi & Toth, 2013). Based on the analysis, it was discovered that the students were not well literate in chemistry, as evidenced by their incorrect answers to the context presented. Previous research disclosed that many students had difficulty in understanding chemical concepts to everyday life and making decisions on social problems related to chemical concepts (Broman & Parchmann, 2014; Cigdemoglu & Geban, 2015; Parchmann, et al., 2006; Yapicioglu & Aycan, 2018). In the context of chemistry, researchers agree on the importance of associating chemical concepts with the phenomena in everyday life to make abstract chemical concepts more concrete (Pabuccu & Erduran, 2016; Sevian et al., 2018). Putting students in contextual situations provides a clear picture of the application of the previously learned knowledge and improves skills relating knowledge to the relevance of everyday life (Broman, et al., 2018). Therefore, chemistry education must address everyday issues, technology, and social and environmental issues. This enables students to understand how chemistry relates to daily life and to apply it in decision making; in other words, students who are literate in chemistry (Shwartz et al., 2013; Tsaparlis, 2000). The student's inability to relate classroom learning to real-life situation stems from the lack of opportunities to learn to connect concepts and phenomena in everyday life (Wu, 2003). Teachers must facilitate students to develop their chemical literacy through chemistry learning. As a result, learning chemistry helps students understand not only the concept of chemistry, but also the role of chemistry in everyday life, technology, and social life.

### Conclusion

Chemical literacy is a critical thinking skill that students must develop as part of their chemistry education. Based on the findings, the developed chemical literacy assessment has high reliability and validity for measuring students' chemical literacy. After being tested on students, it was discovered that many students' perceptions were incorrect. Besides, many students were unaware of the contexts in everyday life. These results indicate that the students' chemical literacy needs to be improved. Experts in chemistry education agree that learning chemistry must foster scientific literacy in general, and chemical literacy in particular (Dori et al., 2008; Talanquer & Sevian, 2013). The chemistry learning provided by the teacher must be capable of assisting students in developing their chemical literacy. Chemical literacy is also critical for teachers to enhance because of its impact on social life, the economy, and decision making (Avargil et al., 2013; Dori et al., 2018). Teachers can continue to provide cases relating to phenomena and items found in everyday life in the future (Kubiato, 2015). This is necessary for students' chemical literacy to improve.

### Recommendations

The researcher made several recommendations based on the findings. The assessment being developed was still limited to the reaction rate material, so future research with other materials in the chemistry subject can be developed. Those who are interested in this topic can combine chemical literacy dimensions with other chemistry subjects. However, when selecting material, future researchers must consider whether the material's characteristics are compatible or not with the chemical literacy dimensions to effectively measure students' chemical literacy.

### Limitations

Although this research was carried out following scientific procedures, several limitations remain. The material chosen was limited to the reaction rate, and 14 items were created for each chemical literacy dimension. There were four items in the chemical content knowledge section, four items in the chemistry in context section, and six items in the high-order learning skills section.

### Acknowledgments

The authors would like to express their gratitude to the Institute of Research and Community Service (LPPM), Universitas Sebelas Maret, Surakarta, Indonesia, for funding master's thesis research.

### Authorship Contribution Statement

Setyorini: Conceptualization, design, analysis, writing. Yamtinah: editing/reviewing, supervision. Mahardiani: editing/reviewing, supervision. Saputro: editing/reviewing, supervision.



## References

- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. Oxford University Press.
- Avargil, S., Herscovitz, O., & Dori, Y. J. (2013). Challenges in the transition to large-scale reform in chemical education. *Thinking Skills and Creativity*, 10, 189–207. <https://doi.org/10.1016/j.tsc.2013.07.008>
- Bain, K., & Towns, M. H. (2016). A review of research on the teaching and learning of chemical kinetics. *Chemistry Education Research and Practice*, 17(2), 246–262. <https://doi.org/10.1039/C5RP00176E>
- Barlia, L. (2016). Patterns of conceptual change process in elementary school students' learning of science. *Journal of Turkish Science Education*, 13(2), 49–60.
- Barnea, N., Judy, Y., & Hofstein, A. (2010). Development and implementation of inquiry-based and computerized-based laboratories : Reforming high school chemistry in Israel. *Chemistry Education Research and Practice*, 11(3), 218–228. <https://doi.org/10.1039/C005471M>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Bond, T. G., & Fox, C. M. (2012). *Applying the Rasch model: Fundamental measurement in the human sciences*. Routledge.
- Boone, B., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer. <https://doi.org/10.1007/978-94-007-6857-4>
- Bradley, K., Peabody, M., Akers, K., & Knutson, N. (2015). Rating scales in survey research: Using the Rasch model to illustrate the middle category measurement flaw. *Survey Practice*, 8(2), 1–14. <https://doi.org/10.29115/SP-2015-0001>
- Broman, K., Bernholt, S., & Parchmann, I. (2018). Using model-based scaffolds to support students solving context-based chemistry problems. *International Journal of Science Education*, 40(10), 1176–1197. <https://doi.org/10.1080/09500693.2018.1470350>
- Broman, K., & Parchmann, I. (2014). Students' application of chemical concepts when solving chemistry problems in different contexts. *Chemistry Education Research and Practice*, 15(4), 516–529. <https://doi.org/10.1039/c4rp00051j>
- Brown, T. L., LeMay, H. E., & Bursten, B. E. (2000). *Chemistry: The central science* (9th ed.). Prentice-Hall.
- Cakmakci, G., Leach, J., & Donnelly, J. (2006). Students' ideas about reaction rate and its relationship with concentration or pressure. *International Journal of Science Education*, 28(15), 1795–1815. <https://doi.org/10.1080/09500690600823490>
- Caleon, I., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education*, 32(7), 37–41. <https://doi.org/10.1080/09500690902890130>
- Celik, S. (2014). Chemical literacy levels of science and mathematics teacher candidates. *Australian Journal of Teacher Education*, 39(1). <https://doi.org/10.14221/ajte.2014v39n1.5>
- Chairam, S., Klahan, N., & Coll, R. K. (2015). Exploring secondary students' understanding of chemical kinetics through inquiry-based learning activities. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(5), 937–956. <https://doi.org/10.12973/eurasia.2015.1365a>
- Cigdemoglu, C., Arslan, H. O., & Cam, A. (2017). Argumentation to foster pre-service science teachers' knowledge, competency, and attitude on the domains of chemical literacy of acids and bases. *Chemistry Education Research and Practice*, 18(2), 288–303. <https://doi.org/10.1039/C6RP00167J>
- Cigdemoglu, C., & Geban, O. (2015). Improving students' chemical literacy level on thermochemical and thermodynamics concepts through context-based approach. *Chemistry Education Research and Practice*, 16(2), 302–317. <https://doi.org/10.1039/C5RP00007F>
- Coll, R. K., & Taylor, N. (2009). Exploring international perspectives of scientific literacy: An overview of the special issue. *International Journal of Environmental Science Education*, 4(3), 197–200.
- Creswell, J. W. (2006). Choosing a mixed method design. In J. W. Creswell & V. L. P. Clark (Eds.), *Designing and conducting mixed methods research* (pp. 58–89). Sage Publications.
- DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, 37(6), 582–601. [https://doi.org/10.1002/1098-2736\(200008\)37:6%3C582::AID-TEA5%3E3.0.CO;2-L](https://doi.org/10.1002/1098-2736(200008)37:6%3C582::AID-TEA5%3E3.0.CO;2-L)

- De Jong, J. H. (2012, March). *Framework for PISA 2015: What 15-years-old should be able to do* [Paper presentation]. 4th Annual Conference of Educational Research Center. Broumana, Lebanon.
- Dijk, E. M. V. (2011). Portraying real science in science communication. *Science Education*, 95(6), 1086–1100. <https://doi.org/10.1002/sce.20458>
- Dori, Y. J., Avargil, S., Kohen, Z., & Saar, L. (2018). Context-based learning and metacognitive prompts for enhancing scientific text comprehension. *International Journal of Science Education*, 40(10), 1198–1220. <https://doi.org/10.1080/09500693.2018.1470351>
- Dori, Y. J., Sasson, I., & Technion, S. (2008). Chemical understanding and graphing skills in an honors case-based computerized chemistry laboratory environment: The value of bidirectional visual and textual representations. *Journal of Research in Science Teaching*, 45(2), 219–250. <https://doi.org/10.1002/tea.20197>
- Duncan, P. W., Bode, R. K., Lai, S. M., & Perera, S. (2003). Rasch analysis of a new stroke-specific outcome scale: The stroke impact scale. *Archives of Physical Medicine and Rehabilitation*, 84, 950–963. [https://doi.org/10.1016/S0003-9993\(03\)00035-2](https://doi.org/10.1016/S0003-9993(03)00035-2)
- Englehard, G. (2013). *Invariant measurement, using Rasch models in the social, behavioral and health sciences*. Routledge.
- Gilbert, J. K., & Treagust, D. F. (2009). *Introduction: Macro, submicro and symbolic representations and the relationship between them: Key models in chemical education multiple representations in chemical education*. Springer.
- Herrmann-abell, C. F., & Deboer, G. E. (2016, April 8-12). *Using Rasch modeling and option probability curves to diagnose students' misconceptions* [Paper presentation]. Annual Meeting of the American Educational Research Association. Washington, DC.
- Holbrook, J., & Rannikmae, M. (2007). The nature of science education for enhancing scientific literacy. *International Journal of Science Education*, 29(11), 1347–1362. <https://doi.org/10.1080/09500690601007549>
- Holman, J., & Hunt, A. (2002). What does it mean to be chemically literate? *Education in Chemistry*, 39(1), 12–14.
- Johnstone, A. H., Macdonald, J. J., & Webb, G. (1977). Misconceptions in school thermodynamics. *Physics Education*, 12(4), 248–251. <https://doi.org/10.1088/0031-9120/12/4/011>
- Khine, M. S. (2020). Objective measurement in psychometric analysis. In M. S. Khine (Ed.), *Rasch measurement applications in quantitative educational research* (pp. 3–7). Springer Nature Singapore Pte Ltd. [https://doi.org/10.1007/978-981-15-1800-3\\_1](https://doi.org/10.1007/978-981-15-1800-3_1)
- Kingir, S., & Geban, O. (2012). The effect of conceptual change approach on students' understanding of reaction rate concepts. *Hacettepe University Journal of Education*, 43, 306–317.
- Kubiatko, M. (2015). Is chemistry attractive for pupils? Czech pupils' perception of chemistry. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(4), 855–863. <https://doi.org/10.12973/eurasia.2015.1444a>
- Li, C., Romero, S., Bonilha, H. S., Simpson, K. N., Simpson, A. N., Hong, I., & Velozo, C. A. (2016). Linking existing instruments to develop an activity of daily living item bank. *Evaluation & the Health Professions*, 41(1), 25–43. <https://doi.org/10.1177/0163278716676873>
- Linacre, J. M. (2006). Data variance explained by Rasch measures. *Rasch Measurement Transactions*, 20(1), 1048–1051.
- Linacre, J. M. (2011). *A user's guide to Winsteps & Ministep Rasch-Model computer programs*. Winsteps. <http://www.winsteps.com/a/winsteps.pdf>
- Liu, X. (2009). Beyond science literacy: Science and the public. *International Journal of Environmental & Science Education*, 4(3), 301–311.
- Liu, X. (2010). *Using and developing measurement instruments in science education: A Rasch modeling approach*. Information Age.
- Liu, X., & Boone, W. J. (2006). Introduction to Rasch measurement in science education. In X. Liu & W. J. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 1–22). JAM Press.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385. <https://doi.org/10.1097/00006199-198611000-00017>
- Mcclary, L. M., & Bretz, S. L. (2012). Development and assessment of a diagnostic tool to identify organic chemistry students' alternative conceptions related to acid strength. *International Journal of Science Education*, 34(15), 2317–2341. <https://doi.org/10.1080/09500693.2012.684433>
- Miller, J. D. (1983). *The American people and science policy*. Pergamon.
- National Research Council. (1996). *The national science education standards*. The National Academies Press.

- National Research Council. (2007). *Talking science to school: Learning and teaching science in grades K-8*. The National Academies press.
- Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, 87(2), 224–240. <https://doi.org/10.1002/sce.10066>
- Pabuccu, A., & Erduran, S. (2016). Investigating students' engagement in epistemic and narrative practices of chemistry in the context of a story on gas behavior. *Chemistry Education Research and Practice*, 17(3), 523–531. <https://doi.org/10.1039/C6RP00011H>
- Parchmann, I., Grasel, C., Baer, A., Nentwig, P., Demuth, R., & Ralled, B. (2006). "Chemie im Kontext": A symbiotic implementation of a context-based teaching and learning approach. *International Journal of Science Education*, 28(9), 1041–1062. <https://doi.org/10.1080/09500690600702512>
- Preczewski, J. P., Mittler, A., & Tillotson, W. (2009). Perspectives of German and US students as they make meaning of science in their everyday lives. *International Journal of Environment & Science Education*, 4(3), 247–258.
- Sadler, T. D., & Zeidler, D. L. (2009). Scientific literacy, PISA, and socioscientific discourse: Assessment for progressive aims of science education. *Journal of Research in Science Teaching*, 46(8), 909–921. <https://doi.org/10.1002/tea.20327>
- Sevian, H., Dori, Y. J., & Parchmann, I. (2018). How does STEM context-based learning work: What we know and what we still do not know. *International Journal of Science Education*, 40(10), 1095–1107. <https://doi.org/10.1080/09500693.2018.1470346>
- Shwartz, Y., Ben-zvi, R., & Hofstein, A. (2006). The use of scientific literacy taxonomy for assessing the development of chemical literacy among high-school students. *Chemistry Education Research and Practice*, 7(4), 203–225. <https://doi.org/10.1039/B6RP90011A>
- Shwartz, Y., Dori, Y. J., & Treagust, D. (2013). *How to justify formal chemistry education, to outline its objectives and to assess them, in teaching chemistry – a studybook: A practical guide and textbook for student teachers, teacher trainees and teachers*. Sense Publishers.
- Stevens, S. Y., Delgado, C., & Krajcik, J. S. (2010). Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal of Research in Science Teaching*, 47(6), 687–715. <https://doi.org/10.1002/tea.20324>
- Suminto, B., & Widhiarsho, W. (2015). *Aplikasi pemodelan Rasch pada assessment pendidikan* [Application of Rasch modelling in educational measurement]. Trim Komunikata.
- Tal, T., & Dierking, L. D. (2014). Learning science in everyday life. *Journal of Research in Science Teaching*, 51(3), 251–259. <https://doi.org/10.1002/tea.21142>
- Talanquer, V., & Sevian, H. (2013). Chemistry in past and new science frameworks and standards: Gains, losses, and missed opportunities. *Journal of Chemical Education*, 91(1), 24–29. <https://doi.org/10.1021/ed400134c>
- Tsaparlis, G. (2000). The states-of-matter approach (SOMA) to introductory chemistry. *Chemistry Education Research and Practice*, 1(1). <https://doi.org/10.1039/A9RP90017A>
- Turanyi, T., & Toth, Z. (2013). Hungarian university students' misunderstandings in thermodynamics and chemical kinetics. *Chemistry Education Research and Practice*, 14, 105–116. <https://doi.org/10.1039/c2rp20015e>
- Witte, D., & Beers, K. (2003). Testing of chemical literacy (Chemistry in context in the Dutch national examination). *Chemical Education International*, 4(1), 1–15.
- Wu, H. (2003). Linking the microscopic view of chemistry to real-life experiences: Intertextuality in a high-school science classroom. *Science Education*, 87(6), 868–891. <https://doi.org/10.1002/sce.10090>
- Xie, Q., Zhong, X., Wang, W., & Lim, C. P. (2014). Development of an item bank for assessing generic competences in a higher-education institute: A Rasch modelling approach. *Higher Education Research & Development*, 33(4), 821–835. <https://doi.org/10.1080/07294360.2013.863847>
- Yalçinkaya, E., Taştan-Kırık, O., Boz, Y., & Yıldırım, D. (2012). Is case-based learning an effective teaching strategy to challenge students' alternative conceptions regarding chemical kinetics? *Research in Science & Technological Education*, 30(2), 151–172. <https://doi.org/10.1080/02635143.2012.698605>
- Yapicioglu, A. E., & Aycan, S. (2018). Pre-service science teachers' decisions and types of informal reasoning about the socioscientific issue of nuclear power plants. *Educational Policy Analysis and Strategic Research*, 13(1), 31–53. <https://doi.org/10.29329/epasr.2018.137.2>