DEVELOPMENT OF A MENTAL MODEL DIAGNOSTIC TEST USING PREDICT OBSERVE EXPLAIN APPROACH ON CORROSION CONCEPTS

Samuel Indra Pratama*[1], Tuszie Widhiyanti[2], Wiji Wiji[3]

[1,2,3] Chemistry Education, Faculty of Mathematics and Natural Sciences, Indonesia University of Education

Corresponding author: samuelsitorus187@gmail.com

Abstract. *This study aimed to develop a Predict–Observe–Explain (POE)-based diagnostic test on corrosion concepts to identify students' mental model profiles. The research employed a Design and Development Research (DDR) approach, involving design, development, and limited evaluation through expert validation, readability testing, and limited trials. The instrument consisted of POE-based items supported by images, videos, and observation tables, integrating macroscopic, submicroscopic, and symbolic representations. The results indicated good content validity based on evaluations from five experts, with revisions made to the third indicator and several items. The instrument was subsequently refined based on expert suggestions before undergoing readability and reliability testing. Readability results showed that all items were highly comprehensible, with an average score above 3.7. Reliability testing using Cronbach's Alpha yielded a coefficient of 0.917, indicating very high reliability. Therefore, the developed instrument is valid, reliable, and suitable for identifying students' mental model profiles*.

Keywords: *diagnostic test, mental model, POE, corrosion, chemistry education*

INTRODUCTION

Current learning strategies are not yet optimal in facilitating students' ability to construct relationships among multiple levels of representation in the concept of corrosion. Several studies have attempted to develop instructional strategies to enhance students' understanding of corrosion concepts. For instance, analogy-based instructional strategies (ICBA) have been shown to improve students' conceptual understanding compared to conventional instruction without analogies [1]. Similarly, the use of multimedia-based intertextual e-modules has been reported to support students in developing more integrated representations by linking particle-level visualizations, symbolic representations, and everyday phenomena [2]. However, despite these improvements, some students still exhibit misconceptions.

A common misconception among students is the assumption that rusting of iron occurs solely as a result of a reaction between iron and oxygen. This understanding represents an oversimplification based on macroscopic observations, in which students observe iron exposed to air (containing oxygen) undergoing rusting. However, from a fundamental perspective, corrosion also involves the presence of water and electrolytes as a medium required to complete the electrochemical cell circuit. This indicates that students have not yet developed an understanding of the corrosion process at the submicroscopic level. When students assume that corrosion only involves Fe and O_2, they tend to overlook other significant environmental factors, such as pH and electrolyte solutions, which play a crucial role in accelerating the rate of corrosion. Difficulties in understanding corrosion concepts are often further exacerbated by misconceptions related to prerequisite concepts, particularly reduction–oxidation (redox) reactions at the symbolic level. Studies have shown that students frequently experience difficulties in determining the oxidation states of elements. A common example is the misinterpretation of the oxidation state of oxygen in the O_2 molecule, where some students assume that the total oxidation state of the two oxygen atoms is –2, whereas the oxidation state of each oxygen atom is 0. This conceptual error at the symbolic level leads to students' inability to balance anode and cathode half-reactions, which are fundamental to the electrochemical mechanisms underlying corrosion. Therefore, conceptual difficulties in corrosion reflect a lack of mastery of fundamental redox concepts, which are inherently abstract [2], [3].

A mental model is an internal representation that individuals construct to understand, explain, and predict phenomena or concepts, including those encountered in chemistry learning [4,5,6]. Students often experience difficulties in understanding chemical concepts due to limitations in constructing accurate mental models. These difficulties arise from students' limited ability to connect abstract concepts with observable or real-world representations [6], [7], [8]. Students' mental models can be identified through the evaluation of their conceptual understanding. This identification is typically conducted using diagnostic test instruments designed to accurately map students' strengths and weaknesses [9], [10]. Mental model diagnostic tests are generally developed by integrating the three levels of representation (macroscopic, submicroscopic, and symbolic) thereby enabling the identification of students' cognitive structures and encouraging them to explain phenomena conceptually [11]. Therefore, the results of diagnostic assessments can serve as a basis for designing more effective instructional strategies in the future.

The Predict–Observe–Explain (POE) learning strategy was first introduced by White and Gunstone as an approach that engages students in predicting experimental outcomes, observing phenomena, and explaining the consistency or discrepancy between their predictions and observations. This strategy has been shown to be effective in promoting students' conceptual understanding and scientific skills [12]. Subsequently, the POE stages were adapted into a diagnostic test instrument to identify students' conceptual understanding [13]. The POE-based diagnostic test is considered appropriate for identifying students' mental models, particularly in chemistry learning, because it accommodates the three levels of chemical representation: macroscopic, submicroscopic, and symbolic. This instrument is also aligned with the stages of the scientific approach, namely observing, questioning, collecting data, associating, and communicating, thereby enabling a more comprehensive assessment of students' chemical understanding. Through the POE framework, students are provided with opportunities to express their mental models more freely without being restricted by predetermined answer choices. Previous studies have shown that POE-based diagnostic tests are effective in revealing students' mental models and misconceptions across various chemistry concepts and in mapping their understanding at the macroscopic, submicroscopic, and symbolic levels [14], [15], [16]. These findings indicate that POE functions not only as an assessment tool but also as a diagnostic approach for examining the depth of students' conceptual understanding in chemistry learning.

In the developed instrument, each item was structured according to the three main stages of POE. In the prediction stage, students were asked to predict the outcome of a corrosion phenomenon based on their prior knowledge. In the observation stage, students observed the phenomenon through experimental images or videos and recorded their observations on the provided response sheet. In the explanation stage, students compared their predictions with the observed results and explained the corrosion phenomenon using macroscopic, submicroscopic, and symbolic representations. Students' responses at each stage provided a more comprehensive description of their mental model profiles [14], [15], [16].

Based on the aforementioned issues, the main research problem of this study is formulated as follows: "How can a Predict–Observe–Explain (POE)-based mental model diagnostic test on the concept of corrosion be developed to identify students' mental model profiles?", this general problem is further elaborated into the following research questions:

1) How is a POE-based mental model diagnostic test for the concept of corrosion designed?

2) How is the content validity of the POE-based mental model diagnostic test for the concept of corrosion established based on expert judgment?
3) What is the readability level of the developed POE-based mental model diagnostic test for the concept of corrosion?
4) What is the reliability of the POE-based mental model diagnostic test for the concept of corrosion?

METHOD

This study employs the Design and Development Research (DDR) method, which consists of the design, development, and limited evaluation stages, focusing on validation, readability testing, and instrument trials. The study participants consisted of five expert validators, including four chemistry education lecturers and one chemistry lecturer. In addition, fifteen high school students and four teachers were involved in the readability testing. A further fifteen high school students participated in the limited testing of the instrument.

The instrument developed in this study was a POE-based mental model diagnostic test constructed based on three assessment indicators. These indicators were formulated through a curriculum analysis of the Phase F Learning Outcomes in Chemistry. The established indicators were then elaborated into test items. Each item was designed in accordance with the POE stages, namely predict, observe, and explain, and was supported by stimuli such as images, videos, and observation tables. The test items and corresponding responses were developed by integrating the three levels of chemical representation, namely macroscopic, submicroscopic, and symbolic, based on references from several general chemistry textbooks.

The data were collected through expert validation, readability testing, and limited testing. The expert validation was conducted using three instruments, as presented in Table 1.

Table 1. Content Validation Instrument

No	Instrument	Valid	Invalid
1.	Alignment of item indicators with learning outcomes		
2.	Alignment of test items with item indicators		
3.	Alignment of answer keys with test items		

The readability testing included several components, as presented in Table 2.

Table 2. Readability Assessment Aspects of the POE-Based Mental Model Diagnostic Test

No	Component	Score			
		1	2	3	4
1.	Language				
2.	Test instructions				
3.	Images				
4.	Videos				
5.	Overall content of the test items				

The language aspect covered sentence clarity, absence of ambiguity, and appropriateness to students' level. The instruction aspect focused on the clarity of the test instructions, while the visual aspect, including images and videos, was evaluated based on quality, relevance, and ease of understanding. In addition, the content of the test items was assessed in terms of clarity and comprehensibility. The data were collected using an online questionnaire and rated on a scale of 1 to 4.

Next, a limited testing was conducted to analyze the instrument's reliability using the Cronbach's alpha coefficient, which is based on the principle of internal consistency. This test aimed to determine the level of consistency among the instrument items in measuring the same construct. The obtained coefficient values were then interpreted according to the Guilford reliability classification to determine the instrument's reliability level, as presented in Table 3

Table 3. Classification of Reliability Levels

Reliability Level	Criteria
$r_{11} \leq 0{,}20$	Very low
$0{,}20 < r_{11} < 0{,}40$	Low
$0{,}40 < r_{11} < 0{,}60$	Moderate
$0{,}60 < r_{11} < 0{,}80$	High
$0{,}80 < r_{11} < 1{,}00$	Very high

RESULT AND DISCUSSION

TDMM–POE Design

The design of the POE-based Mental Model Diagnostic Test instrument in this study was developed based on previous research, particularly through an analysis of the Phase F Learning Outcomes, which served as the basis for formulating the test item indicators [14]. The analysis of the Phase F Learning Outcomes (Grades XI–XII) related to the concept of corrosion identified three main concepts that students are expected to master, namely:

1) The process of corrosion;
2) Corrosion prevention methods; and
3) Factors influencing corrosion.

To develop an understanding of these concepts, students are required to first master the prerequisite concepts, namely redox reactions and voltaic cells.

Subsequently, a cognitive level analysis was conducted for the test item indicators based on the taxonomy of Anderson and Krathwohl (2001), as relevant to the topic of corrosion. The selected cognitive level was analyzing, as it reflects students' ability to break down a concept into its constituent parts and determine the relationships among those parts within an overall structure. At this level, the cognitive processes involve differentiating, organizing, and attributing. Therefore, this study adopted the analyzing level as the basis for developing the indicators. The initial indicators were formulated as follows:

1) Analyzing the process of corrosion
2) Analyzing corrosion prevention methods
3) Analyzing factors influencing corrosion

After the initial indicators were determined, the next step was to develop the test items and answer keys based on content analysis and multiple representations derived from the literature review, including general chemistry textbooks and other relevant sources. The concepts obtained were categorized into macroscopic–symbolic and submicroscopic–symbolic levels of representation. The following presents the characteristics of the developed TDMM–POE instrument:

1) Indicator 1: Analyzing the corrosion process

Students observe a phenomenon involving three iron nails under different conditions: partially submerged in water and exposed to air, exposed to air with silica gel, and fully submerged in water covered with oil.

In the prediction stage, students determine which container will undergo corrosion and provide their reasoning.

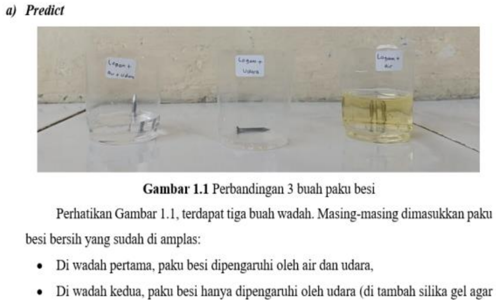


Figure 1. Design of the TDMM–POE instrument: prediction stage of the first indicator

In the observation stage, students watch an experimental video and complete an observation table.

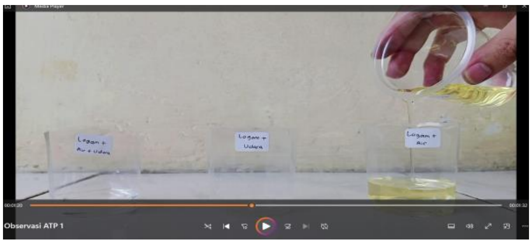
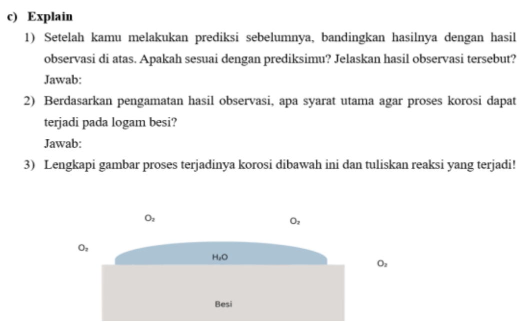


Figure 2. Design of the TDMM–POE instrument: observation stage of the first indicator

In the explanation stage, students compare their predictions with the observed results, draw conclusions about the main conditions required for corrosion, describe the process at the submicroscopic level, write the corresponding chemical reactions, and provide explanations.

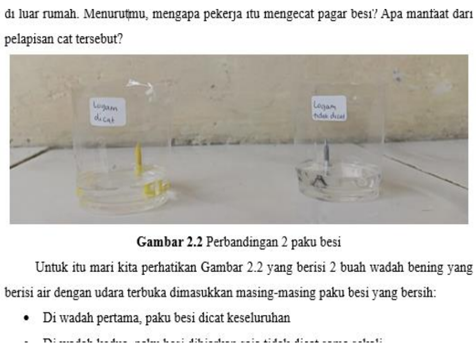

c) **Explain**

1) Setelah kamu melakukan prediksi sebelumnya, bandingkan hasilnya dengan hasil observasi di atas. Apakah sesuai dengan prediksimu? Jelaskan hasil observasi tersebut?
Jawab:
2) Berdasarkan pengamatan hasil observasi, apa syarat utama agar proses korosi dapat terjadi pada logam besi?
Jawab:
3) Lengkapi gambar proses terjadinya korosi dibawah ini dan tuliskan reaksi yang terjadi!

Figure 3. Design of the TDMM–POE instrument: explanation stage of the first indicator

2) Indicator 2: Analyzing corrosion prevention methods

Students observe two iron nails: one fully painted and one unpainted. In the prediction stage, students determine which nail will undergo corrosion and provide their reasoning


di luar rumah. Menurutmu, mengapa pekerja itu mengecat pagar besi? Apa manfaat dari pelapisan cat tersebut?

Gambar 2.2 Perbandingan 2 paku besi

Untuk itu mari kita perhatikan Gambar 2.2 yang berisi 2 buah wadah bening yang berisi air dengan udara terbuka dimasukkan masing-masing paku besi yang bersih:

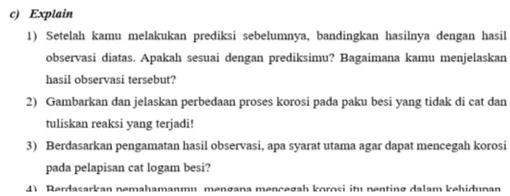
- Di wadah pertama, paku besi dicat keseluruhan

Figure 4. Design of the TDMM–POE instrument: prediction stage of the second indicator

In the observation stage, students watch an experimental video, complete an observation table, and explain the observed differences.

Figure 5. Design of the TDMM–POE instrument: observation stage of the second indicator

In the explanation stage, students summarize the main factors influencing corrosion prevention, describe the process at the submicroscopic level, write the corresponding chemical reactions, and watch a video presenting other corrosion prevention methods

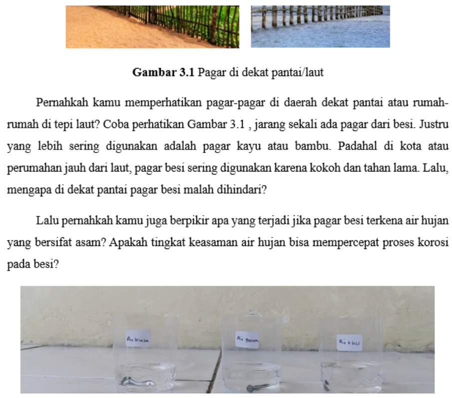

c) ***Explain***

1) Setelah kamu melakukan prediksi sebelumnya, bandingkan hasilnya dengan hasil observasi diatas. Apakah sesuai dengan prediksimu? Bagaimana kamu menjelaskan hasil observasi tersebut?
2) Gambarkan dan jelaskan perbedaan proses korosi pada paku besi yang tidak di cat dan tuliskan reaksi yang terjadi!
3) Berdasarkan pengamatan hasil observasi, apa syarat utama agar dapat mencegah korosi pada pelapisan cat logam besi?

Figure 6. Design of the TDMM–POE instrument: explanation stage of the second indicator

3) Indicator 3: Analyzing factors influencing corrosion

Students observe three iron nails under different conditions: submerged in salt water, in plain water, and in HCl solution. In the prediction stage, students determine which container will undergo corrosion the fastest and provide their reasoning.


Gambar 3.1 Pagar di dekat pantai/laut

Pernahkah kamu memperhatikan pagar-pagar di daerah dekat pantai atau rumah-rumah di tepi laut? Coba perhatikan Gambar 3.1 , jarang sekali ada pagar dari besi. Justru yang lebih sering digunakan adalah pagar kayu atau bambu. Padahal di kota atau perumahan jauh dari laut, pagar besi sering digunakan karena kokoh dan tahan lama. Lalu, mengapa di dekat pantai pagar besi malah dihindari?

Lalu pernahkah kamu juga berpikir apa yang terjadi jika pagar besi terkena air hujan yang bersifat asam? Apakah tingkat keasaman air hujan bisa mempercepat proses korosi pada besi?

Figure 7. Design of the TDMM–POE instrument: prediction stage of the third indicator

In the observation stage, students watch an experimental video, complete an observation table, and explain the differences among the containers.

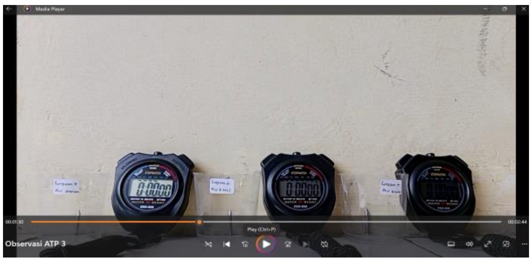
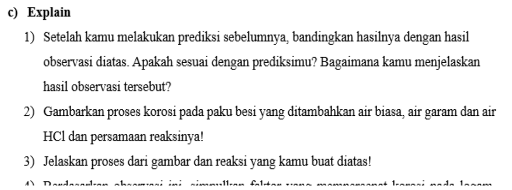


Figure 8. Design of the TDMM–POE instrument: observation stage of the third indicator

In the explanation stage, students deduce the factors that accelerate corrosion, describe the process at the submicroscopic level, write the corresponding chemical reactions, and provide explanations.



c) Explain

1) Setelah kamu melakukan prediksi sebelumnya, bandingkan hasilnya dengan hasil observasi diatas. Apakah sesuai dengan prediksimu? Bagaimana kamu menjelaskan hasil observasi tersebut?
2) Gambarkan proses korosi pada paku besi yang ditambahkan air biasa, air garam dan air HCl dan persamaan reaksinya!
3) Jelaskan proses dari gambar dan reaksi yang kamu buat diatas!

Figure 9. Design of the TDMM–POE instrument: explanation stage of the third indicator

Content Validation

The developed test item indicators were then validated for their alignment with the Learning Outcomes by four experts who were Chemistry Education lecturers.

Table 4. Summary of Validation Results on the Alignment of Test Item Indicators with Phase F Learning Outcomes

No	Test Item Indicator	Alignment with Learning Outcomes		
		Valid	Invalid	Suggestion
1)	Analyzing the process of corrosion	√		
2)	Analyzing corrosion prevention methods	√		
3)	Analyzing factors influencing corrosion		√	Suggested revision: “accelerate”

Based on Table 4, the validation results show that test item indicators 1 and 2 are in accordance with the Learning Outcomes. However, several suggestions for improvement were provided for indicator 3. The original formulation, “analyzing factors that influence corrosion,” was suggested to be revised into “analyzing factors that accelerate corrosion,” as this indicator is more closely related to the concept of reaction rate. Based on this input, the indicator was revised to focus on the aspect of corrosion kinetics; thus, the term “accelerate” was considered more appropriate than “influence.”

Furthermore, the alignment of the test items with the test item indicators was validated by four Chemistry Education lecturers and one Pure Chemistry expert lecturer.

Table 5. Summary of Validation Results on the Alignment of Test Items with Test Item Indicators

No	Test Item Indicator	Alignment with Test Items		
		Valid	Invalid	Suggestion
1)	Analyzing the process of corrosion	√		
2)	Analyzing corrosion prevention methods	√		
3)	Analyzing factors influencing corrosion		√	Adjusted to the revised indicators

Based on Table 5, the validation results show that the test items for indicators 1 and 2 are appropriate. However, the test items for indicator 3 need to be revised to better align with the revised indicator.

The next validation results, namely the suitability of the answers with the questions, were evaluated by four validators from Chemistry Education lecturers and one expert lecturer in Pure Chemistry. According to the validators, the answers still required substantial improvement in terms of language use and multiple levels of representation. The validators then provided several suggestions for improvement for each Test Item Indicator as follows:

1) Indicator 1

The validators recommended re-examining the conditions for rust formation and adding the overall reaction equation so that the answers are more appropriate to the question. Improvements in grammatical consistency were also suggested, particularly in the observation table. The statement “no significant change” was suggested to be clarified, and the answers needed to be supported with factual evidence so that the relationship between levels of representation (macroscopic, submicroscopic, and symbolic) becomes clearer

2) Indicator 2

The validators assessed that the answer to the question "What are the main requirements for preventing corrosion in iron metals?" needs to be revised, as corrosion prevention is not only carried out through surface coating but also by preventing iron from contacting oxygen and water. In addition, video observations of other prevention methods should be developed into separate questions so that students' understanding of various corrosion prevention strategies can be better explored.

3) Indicator 3

The validators assessed the wording of the question "In which container will an iron nail rust the fastest?" as inappropriate because "rusting" is not fully identical to "corrosion." The changes occurring in iron in HCl solution differ from those in plain water; therefore, the measurement indicators need to be clarified. The criteria for stopping the reaction and measuring time during the observation stage were also considered unclear. In the HCl solution, rust does not form on the nail, resulting in different observation outcomes. The validators suggested using more measurable criteria, such as measuring the mass of the nail before and after the process within the same time duration but under different conditions, so that the comparison of corrosion rates can be determined more clearly.

Next, the validators' feedback was reviewed and analyzed as the basis for revising and refining the developed POE-based mental model diagnostic test, so that the resulting instrument met the validity criteria [14].

Readability Test

A readability test is conducted to assess the extent to which respondents can understand the statements contained in an instrument. This test aims to ensure that the language, sentence structure, instructions, and media presented in the instrument are clearly understood, enabling respondents to provide appropriate responses. The readability test was carried out using a Likert-scale instrument with statement categories that were structured clearly and concisely [17].

Table 6. Criteria for Average Readability Score Categories

Score Range	Category
3,25 – 4,00	Highly comprehensible
2,50 – 3,24	Moderately comprehensible
1,75 – 2,49	Slightly comprehensible
1,00 – 1,74	Not comprehensible

The assessment was conducted using a 1–4 scale with category levels as shown in Table 6.

Table 7. Readability Test Results of the POE Mental Model Diagnostic Test Instrument

Component	Mean Score	Kategori
Language	3,86	Highly comprehensible
Test instructions	3,83	Highly comprehensible
Images	3,56	Highly comprehensible
Videos	3,92	Highly comprehensible
Overall content of the test items	3,67	Highly comprehensible
Mean	3,77	Highly comprehensible

Based on the readability test results presented in Table 7, all assessed aspects including language, question instructions, images, videos, and overall question content were categorized as very easy to understand. The video aspect obtained the highest average score of 3.92, while the image aspect obtained the lowest score of 3.56, but both remained in the same category. Overall, the average readability score of 3.77 or 94,25% indicates that the developed instrument has a very good level of readability.

This finding supports the theory proposed by [17], which states that readability test results are analyzed based on the percentage or average score obtained for each assessed aspect. If more than 30% of the instrument items still require revision, the readability test should be conducted again. Conversely, if the number of items requiring revision is below this threshold, the instrument can be considered to have an adequate level of readability. This shows that the POE (Predict Observe Explain) mental model diagnostic test instrument uses clear language, easy to understand instructions, and relevant stimuli that support student understanding. Therefore, the developed POE-based mental model diagnostic test is considered suitable for use, particularly in chemistry learning [14], [15], [16].

Reliability Test

Reliability testing was conducted to determine the level of consistency among the instrument items in measuring the same construct. Instrument reliability was analyzed using the Cronbach's alpha coefficient, which is based on the principle of internal consistency. The selection of the Cronbach's alpha coefficient as a measure of reliability was based on its suitability for assessing questionnaire-based instruments with more than one response alternative. In addition, the Cronbach's alpha test is considered appropriate for essay- and questionnaire-based instruments [18]. The results of the instrument reliability analysis are presented in Tables 8 and 9.

Table 8. Item Reliability Statistics Result

Item	Item-rest correlation	If item dropped
		Cronbach's α
1)	0.1543	0.918
2)	0.6153	0.909
3)	0.5198	0.911
4)	0.0152	0.925
5)	0.5851	0.910
6)	0.6895	0.907
7)	0.7872	0.905
8)	0.6790	0.907
9)	0.0827	0.922
10)	0.5766	0.910
11)	0.8474	0.902
12)	0.7872	0.905
13)	0.7894	0.904
14)	0.6166	0.909
15)	0.6885	0.907
16)	0.5766	0.910
17)	0.6608	0.908
18)	0.8826	0.901

Based on Table 8, the Cronbach's Alpha if item deleted values for each item ranged from 0.901 to 0.922. This indicates that no item significantly reduces the instrument reliability when deleted. In addition, the item-rest correlation values for most test items were in the moderate to high category, indicating that each item consistently measures the same construct. Therefore, all test items can be retained because they contribute positively to the overall reliability of the instrument.

Table 9. Scale Reliabilitiy Statistics Result

	Cronbach's α
scale	0.914

Based on Table 9, the Cronbach's alpha value was 0.914, which falls into the very high category according to Guilford's reliability classification [18]. This indicates that the instrument has a very good level of internal consistency and can therefore be considered reliable for measuring students' mental models of the corrosion concept. This high reliability value also indicates that the items in the instrument were consistently constructed and were able to measure the same construct in a stable manner.

The high reliability value strengthens the findings of previous studies that the POE-based mental model diagnostic test is suitable for analyzing students' mental model profiles, particularly in chemistry learning, which emphasizes the relationship among the three levels of representation: macroscopic, submicroscopic, and symbolic [14], [15], [16].

CONCLUSIONS AND SUGGESTIONS

Conclusions

The POE (Predict Observe Explain) mental model diagnostic test on the concept of corrosion was successfully developed through the Design and Development Research (DDR) stages, which include design, development, and limited evaluation. The developed instrument integrates the POE stages (predict, observe, explain), is supported by stimuli such as images, videos, and observation tables, and accommodates three levels of chemical representation, namely macroscopic, submicroscopic, and symbolic.

The validation results show that the instrument has good content validity based on expert judgment, with several revisions, particularly on the third indicator as well as improvements to test items and answers. The readability test results indicate that all aspects of the instrument are categorized as very easy to understand with an average score of 3.77. In addition, the reliability test results show a Cronbach's Alpha value of 0.914, which is categorized as very high. Therefore, the developed instrument is considered valid, reliable, and suitable for identifying students' mental model profiles on the concept of corrosion.

Suggestions

1) The developed diagnostic test instrument can be used by teachers as a tool to identify students' mental models and misconceptions on the concept of corrosion, thereby providing a basis for designing more effective learning strategies.
2) This study is limited to a small-scale trial; therefore, future researchers are suggested to conduct the test on a larger sample in order to obtain a more comprehensive overview of students' mental model profiles.
3) The development of similar instruments can be extended to other chemistry concepts with abstract characteristics, such as chemical equilibrium or acid–base solutions, in order to broaden the application of POE-based diagnostic tests in chemistry learning.
4) Further development is needed in analyzing students' mental model profiles in greater depth, so that the test results are not only descriptive but can also be used to design appropriate instructional interventions.

ACKNOWLEDGEMENT

The author would like to express sincere gratitude to four validators from Chemistry Education and one validator from Pure Chemistry for their time, attention, and constructive suggestions, which have contributed to improving the quality of this study. Appreciation is also extended to the teachers and students who willingly assisted in the data collection process.

In addition, the author also sincerely thanks the supervisor for providing guidance, motivation, and support throughout the writing process of this article. All forms of assistance and support from various parties have played an important role in the completion of this work.

REFERENCES

[1] Asih, F. E., Suyono, Ibnu, S., & Suharti. 2019. Student s ' Misconceptions on Understanding Corrosion Topic by and without Analogy. *Atlantis Highlights in Chemistry and Pharmaceutical Sciences*, Vol. 1, hal. 130–134.

[2] Azizah, N. T. S., Wiji, W., & Yuliani, G. 2025. Pengembangan e-modul berbasis intertekstual pada materi korosi untuk meningkatkan kemampuan representasional peserta didik. *Jurnal Riset Dan Praktik Pendidikan Kimia*, Vol. 13, No. 1, hal. 29–46.

[3] Muslikhah, N., Mulyani, S., & Utomo, S. B. 2020. Identifikasi miskonsepsi siswa pada materi konsep reaksi redoks dan reduksinya dengan menggunakan model REACT. *Jurnal Penelitian Pendidikan*, Vol. 23, No. 1, hal. 24–36.

[4] Praisri, A., & Faikhamta, C. 2020. Enhancing students' mental models of chemical equilibrium through argumentation within model-based learning. *International Journal of Learning, Teaching and Educational Research*, 19(7), Vol. 19, No. 7, hal. 121–142.

[5] Bilir, V., & Karaçam, S. 2021. Evaluation of mental models of prospective science teachers on chemical reactions. *Journal of Pedagogical Research*, Vol. 5, No. 1, hal. 258–274.

[6] Salilas, E. C., Muñoz, J. P., Jiménez, M., & Álvarez, y F. J. F. 2022. Analysis of the mental model about the atom concept in Spanish 15- to 18- years old students. *Educación Química*, Vol. 33, No. 2, hal. 181–193.

[7] Varela, B., Sesto, V., & García-rodeja, I. 2020. An Investigation of Secondary Students ' Mental Models of Climate Change and the Greenhouse Effect. *Springer: Research in Science Education*, Vol. 50, No. 2, hal. 599–624.

[8] Çalış, S. 2024. Turkish science high school students' mental models of the electron cloud. *Chemistry Education Research and Practice*, Vol. 24, No. 4, hal. 1105–1121.

[9] Volfson, A., Eshach. H., & Ben-Abu, Y. 2021. Preliminary development of a simple statistical tool for estimating mental model states from a diagnostic test. *Physical Review Physics Education Research*. Vol. 17, No. 2, hal. 023105.

[10] Umar, F. A., Samsudin, A., Kapıcı, H. Ö., Ramlis, T. R., Aminudin, A.H.,Mufida, S. N., Kunaedi, J. Astuti, I. R. W., & Dewi, F. A. 2024. Developing dispersion and polarization conceptual inventory (DiPolCI) to identify students' mental model. *Journal of Science Learning*, Vol. 7, No. 3, hal. 238–247.

[11] Rusilowati, A. 2015. Pengembangan Tes Diagnostik Sebagai Alat Evaluasi Kesulitan Belajar Fisika. *Prosiding Seminar Nasional Fisika Dan Pendidikan Fisika*, Vol. 6, No. 1, hal. 1–10.

[12] Harta, J., Listyarini, R. V., Pamenang, F. D. N., Wijayanti, L. W., & Lee, W. 2020. Developing Small Scale Chemistry Practicum Module to Identify Students' Ability in Predict-Observe-Explain (POE) Implementation. *Jurnal Kimia dan Pendidikan Kimia*, Vol. 5, No. 1, hal. 91–99.

[13] Sesen, B. A. 2013. Diagnosing pre-service science teachers' understanding of chemistry concepts by using computer-mediated predict–observe–explain tasks. *Chemistry Education Research and Practice*, Vol. 14, hal. 239-246.

[14] Widhiyanti, T., Luviani, S. D., Wiji, W., & Mulyani, S. 2022. Revealing Preservice Teachers ' Conceptions , Troublesome Knowledge and Threshold Concept of Chemical Equilibrium through Predict Observe Explain Mental Model Diagnostic Test (POE-MMDT) The Mental Model Represents The Understanding of Multiple Represent. *Asia Pacific Journal Of Educators and Education*, Vol. 37, No. 2, hal. 35–49.

[15] Katmiati, S., & Rahmi, C. 2021. Tes Diagnostik Prediksi-Observasi-Eksplanasi (POE) Reaksi Kimia Untuk Menggali Model Mental Siswa. *Jurnal Zarah*, Vol. 9, No. 2, hal. 97–104.

[16] Kholidanata, F., Wiji, W., & Yuliani, G. 2023. Profil Model Mental Siswa Pada Materi Hidrolisis Garam Berdasarkan Strategi Evaluasi Model Predict-Observe-Explain (POE). *Jurnal Riset dan Praktik Pendidikan Kimia*, Vol. 11, No. 2, hal. 142–154.

[17] Rohmad, D., Sarah, S. 2021. *Pengembangan Instrumen Angket*. Penerbit K-Media:Yogyakarta

[18] Putri, E. H., Wahyudy, M. A., Yulianto, A., Nuraeni, F. 2020. Development of Instruments to Measure Mathematical Anxiety of Elementary School Students. *International Journal of Learning, Teaching and Educational Research*, Vol. 19, No. 6, hal. 282-302.