Paedagoria : Jurnal Kajian, Penelitian dan Pengembangan Kependidikan
http://journal.ummat.ac.id/index.php/paedagoria
Prefix DOI: 10.31764

ISSN 2086-6356 (Print)
ISSN 2614-3674 (Online)
Vol. 16, No. 2 April 2025, Hal. 145-154

DEVELOPMENT OF PICTORIAL-BASED FIVE-TIER MULTIPLE CHOICE DIAGNOSTIC TEST TO IDENTIFY MISCONCEPTIONS OF ELECTROCHEMICAL CELL MATERIAL

Nur Afni Himmatul Ulya[1], Setia Rahmawan[1]
[1]Cemistry Education, Sunan Kalijaga State Islamic University Yogyakarta, Indonesia
afniasyakir2@gmail.com[1], setia.rahmawan@uin-suka.ac.id[1]

INFO ARTIKEL

Riwayat Artikel:

Diterima: 20-12-2024
Disetujui: 01-04-2025

Kata Kunci:

Five-tier Multiple Choice; Miskonsepsi; *Pictorial*; Sel Elektrokimia; Tes Diagnostik

ABSTRAK

Abstrak: Miskonsepsi dalam pembelajaran kimia, khususnya pada sel elektrokimia seperti sel volta dan elektrolisis, sering terjadi karena kompleksitas konsep redoks, perpindahan elektron, serta proses di anoda dan katoda. Penelitian ini bertujuan mengembangkan instrumen tes diagnostik *five-tier multiple choice* berbasis *pictorial* untuk mengidentifikasi miskonsepsi pada materi ini. Metode yang digunakan adalah Research and Development (R&D) model Tessmer, yang terdiri dari lima tahap: *preliminary*, *self-evaluation*, *expert review*, *one-to-one evaluation*, dan *small group evaluation*. Validitas diuji menggunakan *Content Validity Ratio* (CVR) dan *Content Validity Index* (CVI), sementara reliabilitas diuji dengan model Rasch. Hasil penelitian menunjukkan bahwa instrumen ini valid (CVI rata-rata 1,00) dan memiliki reliabilitas sedang (Cronbach's Alpha 0,71-0,76). Instrumen ini berhasil mengidentifikasi miskonsepsi terkait prinsip kerja sel volta, elektrolisis, dan arah aliran elektron. Mayoritas peserta didik mengandalkan internet sebagai sumber utama informasi, tetapi kurangnya verifikasi memperbesar potensi miskonsepsi. Instrumen ini diharapkan membantu guru dalam menyusun strategi pembelajaran berbasis data untuk meningkatkan pemahaman peserta didik.

Abstract: *Misconceptions in chemistry learning, particularly in electrochemical cells such as voltaic cells and electrolysis, often arise due to the complexity of redox concepts, electron transfer, and processes occurring at the anode and cathode. This study aims to develop a pictorial-based, five-tier multiple-choice diagnostic test to identify misconceptions in this subject. The method used follows the Research and Development (R&D) approach based on the Tessmer model, which consists of five stages: preliminary analysis, self-evaluation, expert review, one-to-one evaluation, and small group evaluation. Validity was assessed using the Content Validity Ratio (CVR) and Content Validity Index (CVI), while reliability was tested using the Rasch model. The results indicate that the instrument is valid (average CVI = 1.00) and has moderate reliability (Cronbach's Alpha = 0.71–0.76). It successfully identifies misconceptions related to the working principles of voltaic cells, electrolysis, and the direction of electron flow. The majority of students rely on the internet as their primary source of information; however, the lack of verification increases the risk of misconceptions. This instrument is expected to help teachers design data-driven learning strategies to enhance students' understanding.*

A. INTRODUCTION

Chemistry is an important subject that develops students' critical thinking skills and creativity (Abd.Rachman et al., 2017). However, chemistry is often considered a difficult and under-desirable subject due to its complex challenges (Muderawan et al., 2019). Previous research has shown that many students have difficulty understanding abstract chemistry concepts and require a deep understanding (Permana et al., 2019). A studi by Kristiana et al. (2020) it also revealed that students' understanding of electrochemical cell materials is still low, which indicates the need for an effort to improve student learning outcomes.

In previous research conducted by Rahayu et al. (2011) It found that many students had misconceptions in the concepts of electrolysis, electric flow, voltage cells, and electrode reactions. Only about 44% of Indonesian students understand the concepts in this material, while 84% still have difficulties with the redox reaction material. This difficulty is due to the need to understand three levels of chemical representation: macroscopic, symbolic, and submicroscopic (Ahmad & Lah, 2013). Imbalances in the understanding of these three representations often lead to misconceptions (Sari, 2021). Misconceptions in students can be an obstacle in the learning process and affect their ability to understand a concept thoroughly (Hermita et al., 2017). Misconceptions in chemistry learning have a profound impact and become an obstacle to understanding the following materials because of the close connection between various chemical concepts (Savira et al., 2019). Therefore, effective diagnostic tools are needed to identify and analyze the misconceptions experienced by students in order to improve their understanding of chemical concepts.

Diagnostic tests are designed to thoroughly identify individual students' difficulties in a single test session, with the aim of detecting specific weaknesses in their understanding (Putri & Rinaningsih, 2021). A variety of diagnostic test instruments have been developed to assess student comprehension, including interviews, open-ended questionnaires, concept maps, images, word associations, as well as multiple-choice tests (Kaltakci-gruel et al., 2017). Among these methods, multiple-choice diagnostic tests are considered more practical, but they have limitations in distinguishing correct answers based on actual understanding from answers that are chosen for the wrong reasons (Budiharti et al., 2018).

Effectively addressing misconceptions requires proper identification of the cause (Rosita et al., 2020). Previous research has shown that one-tier and two-tier diagnostic tests are still not effective enough in identifying misconceptions because they only assess students' final answers without considering the reasons behind those answers (Annisa et al., 2019). Meanwhile, the four-tier diagnostic test, as researched by Roghdah et al. (2021). and Nurlita (2022), proven to be able to better identify students' levels of understanding and misconceptions. However, the research of Wahyuni et al. (2022) highlighting that the four-level test is still not able to uncover the root cause of students' misconceptions thoroughly. In an effort to overcome these limitations, Simamora et al. (2023) Introducing a five-tier diagnostic test that adds open-ended questions at the fifth level to capture more deeply the student's conceptual reasoning. This approach makes diagnostic tests more effective in evaluating students' conceptual understanding and opens up opportunities for better learning intervention strategies. In addition, diagnostic tests in general are often presented in narrative form, which can cause boredom in students when working on the questions (Santi & Rahayu, 2022). As a solution to overcome these weaknesses, the use of image-based questions or pictorial-based diagnostic tests can be a more interesting and effective alternative solution in identifying student misconceptions.

Pictorial tests are assessment tools that utilize images or visual representations with specific meanings to illustrate a particular concept, object, or phenomenon (Virgiawan et al., 2020). Various forms of pictorial representation, such as pictures, tables, diagrams, graphs, and mind maps, can enhance students' understanding by providing concrete visualizations of abstract concepts (Tavassoli et al., 2013). Additionally, these images can depict submicroscopic aspects of chemistry and aid in identifying misconceptions that students may experience (Hanson, 2015). Despite these advantages, previous research has primarily focused on using pictorial representations in general learning contexts rather than incorporating them into structured diagnostic assessments. Moreover, conventional diagnostic instruments still rely heavily on textual explanations, which can be challenging for students with reading comprehension difficulties (Nuraisyah et al., 2022). Therefore, integrating pictorial tests into diagnostic assessments offers a more effective approach to identifying misconceptions while simultaneously improving students' understanding of chemistry concepts and enhancing their interest and motivation in the subject.

The urgency of this research lies in the increasing need for effective diagnostic tools to identify students' difficulties in understanding complex chemical concepts. With the increasing challenges in chemistry learning, this study aims to develop and implement a five-level image-based multiple-choice diagnostic test to improve student understanding. To date, no research has been found that develops an image-based five-level diagnostic instrument for electrochemical cell materials. Therefore, this study introduces an innovative diagnostic tool that not only identifies misconceptions but also engages students through visual representations. The aim of this study was to develop and test the effectiveness of an image-based five-level multiple-choice diagnostic test in identifying students' difficulties in understanding chemical concepts. Thus, this research is expected to make a real contribution to improving the quality of chemistry learning and its relevance in the development of science education.

B. RESEARCH METHODS

Development Model

This research uses the Research and Development (R&D) method to produce and evaluate products systematically (Sugiyono, 2013). This study uses an R&D method with the Tessmer model, which includes a preliminary stage, a formal evaluation, and product testing based on recommendations of Martin Tessmer (1993), which includes self-evaluation, prototyping (expert review, one-on-one, small group) with low resistance revisions, as well as field tests with high resistance revisions (Akker et al., 1999).

Development Procedure

This study requires a procedure to achieve the test package's final prototype as intended. This research consists of four main stages: preliminary, self-evaluation, prototyping, and field test:

1. Fase preliminary

In the preliminary stage, the researcher conducted several analyses, namely, student analysis, material analysis, and curriculum analysis. In the preliminary study, the researcher produced a pictorial-based five-tier diagnostic test design on electrochemical cell material to be used as an initial prototype draft.

2. Self-Evaluation - Researchers review and refine Prototype I based on language, question structure, and content suitability.
3. Fase prototyping – involves three evaluation stages:

1. Expert Review: Five validators assess language, construction, and content.

2. One-to-One: Testing with three students of different abilities for feedback.

3. Small Group: Testing with six students (two high, two moderate, two low ability) to refine Prototype III.

4. Field Test - Prototype III is tested with 12th-grade science students, incorporating feedback for final improvements.

1. Trial Questions

The expert test was conducted by two chemistry lecturers and three chemistry teachers, while a field practitioner test was carried out by a chemistry teacher to identify minor errors. The empirical test involved 60 12th-grade science students through limited and field trials. Data collection included qualitative (interviews, validator comments, and responses) and quantitative aspects (validity, reliability, difficulty level, incorrect answer patterns, and question effectiveness).

2. Question Item Analysis Techniques

Validity uses the Content Validity Ratio (CVR) to assess the suitability of question items to domains measured based on expert assessment (Lawshe, 1975). Scoring item answers uses the CVR method, which processes the score. The CVR formula is:

$$\mathrm{CVR} = \frac{Ne - \frac{N}{2}}{\frac{N}{2}}$$

Information:

CVR = Content validity ratio

Ne = Number of validators who gave a "valid" rating

N = Number of validators

Then, an evaluation of the suitability of the format and language of the question items was carried out using the CVI formula, which provided a further overview of the extent to which the question items were considered valid in the context of conformity with the criteria set. The following is the CVI formula:

$$\mathrm{CVI} = \frac{total\ Number\ of\ CVR}{number\ of\ questions}$$

After obtaining valid question items, a reliability test was carried out using the Rasch model. The processing of reliability values is carried out using Ministep software version 5.8.3.8 to determine reliability criteria. According to (Sugiyono, 2022), to obtain a distribution of measurement values close to normal, the Number of respondents for the test of questions with validity and reliability tests is at least 30. In this study, the test was carried out using two classes with different questions but still referring to the same indicators. Although the questions are different, the preparation and testing have met the main requirements of quantitative research calculations.

C. RESULTS AND DISCUSSION

1. Tes Validity

The diagnostic test instruments were analyzed using CVR and CVI based on validating five validators, covering aspects of the material, construction, language, and instructions for working on the questions. The results of the analysis of the validity of the diagnostic test instrument are as follows:

Table 1. Results of CVR and CVI

Question Items	V 1	V 2	V 3	V 4	V 5	CVR Value	CVI Value
1-30	1	1	1	1	1	1,00	1,00

Based on Table 1, the question items that have been assessed by the validator have shown that all question items have a CVR and CVI value of 1.00, which is in the very appropriate category. This indicates the instrument is valid and can be used for further trials. Some questions are revised to improve quality, such as correcting redactions, adding information, or presenting data to be more effective in measuring students' abilities.

Question Number	Suggestion
3	It is necessary to add a description of the anode and cathode used
4,5,6	The taxonomy of bloom does not match. In package A, it is necessary to add a description of the compound phase to the answer. In the typo answer key, the tier 1 and tier 2 C packages should be tier 1 and tier 3
13,14,15	Packages A, B, and C can be created for this purpose. The indicator of the question is changed to predict the cell diagram instead of compiling the cell diagram. The descriptive cell notation in the problem and image is inconsistent (the lowest Ag electrode in the Ast Ion solution and the Ni electrode are submerged in the Ni^{2+} Ion solution, but the solution is drawn as "the same."
18	It is better to provide in the form of story literacy to encourage students to use their imagination so that they can increase their reasoning skills
20	It is known that cathodes and anodes are presented in the question. In the tier 3 option, it does not make sense if the cell Eº data has been given; the solution can be given a reason for why or the meaning of cell Eº and how the cell Eº occurs.
22	Note: The reason can be shortened. A lower (negative) Eº can protect can be added so that it is more HOTS than just comparing the Eº of the cell.

The following table shows the revision process carried out on one of the question items.

Table 2. Valid Question Items for Revision Before Revision

<table>
<tr><th>Tier</th><th>Questions and answer options</th></tr>
<tr><td>Tier 1</td><td>Take a look at the picture below!
Protection Against Corrosion

Figure 3. Cathodic Protection
Source: tjdecho.com
An engineer named Ahmad plans to protect the iron structure from corrosion. He has standard electrode potential data for some metals that can be used in the cathodic protection process. In his planning, Ahmad had to choose a metal that could act as a cathode to protect the iron from rusting. Ahmad considers the standard electrode potential of the various metals available:
$Mg^{2+}(aq) + 2e^- \rightarrow Mg(s)$ Eº = -2,37 volt
$Pb^{2+}(aq) + 2e^- \rightarrow Pb(s)$ Eº = -0,13 volt
$Ni^{2+}(aq) + 2e^- \rightarrow Ni(s)$ Eº = -0,25 volt
$Fe^{2+}(aq) + 2e^- \rightarrow Fe(s)$ Eº = -0,44 volt
$Cu^{2+}(aq) + 2e^- \rightarrow Cu(s)$ Eº = +0,34 volt
$Ag^{+}(aq) + e^- \rightarrow Ag(s)$ Eº = +0,80 volt
Metals that protect the iron from corrosion using the cathodic protection method are...
A. Ag
B. Cu
C. Ni
D. Mg
E. Pb</td></tr>
<tr><td>Tier 2</td><td>Are you sure of your answer?
A. Sure
B. Not sure</td></tr>
<tr><td>Tier 3</td><td>What is your reason for choosing the answer?
A. Nickel (Ni) can protect iron from corrosion through cathodic protection methods because its reduction potential is higher than</td></tr>
</table>

	iron's. B. Magnesium (Mg) can protect iron from corrosion using the cathodic protection method because its reduction potential is lower than iron's. C. Copper (Cu) can protect iron from corrosion by cathodic protection methods because its reduction potential is higher than iron's. D. Silver (Ag) can protect iron from corrosion through cathodic protection because its reduction potential is lower than iron's. E. Lead (Pb) can protect iron from corrosion using the cathodic protection method because its reduction potential is higher than iron's.
Tier 4	Are you sure of your answer? A. Sure B. Not sure
Tier 5	Where did you get the information from? A. Package book B. Internet C. Teacher D. Friends E. Parents

Table 3. Valid Question Items for Revision After Revision

Tier	**Questions and answer options**
Tier 1	Take a look at the picture below! Protection Against Corrosion  **Figure 4.** Cathodic Protection Source: tjdecho.com An engineer named Ahmad is planning to protect the iron structure from corrosion. He has standard electrode potential data for some metals that can be used in the cathodic protection process. In his planning, Ahmad had to choose a metal that could act as a cathode to protect the iron from rusting. Ahmad considers the standard electrode potential of the various metals available: Mg^{2+} (aq) + 2e$^-$ → Mg (s) Eº = -2,37 volt Pb^{2+} (aq) + 2e$^-$ → Pb (s) Eº = -0,13 volt Ni^{2+} (aq) + 2e$^-$ → Ni (s) Eº = -0,25 volt Fe^{2+} (aq) + 2e$^-$ → Fe (s) Eº = -0,44 volt Cu^{2+} (aq) + 2e$^-$ → Cu (s) Eº = +0,34 volt Ag^{+} (aq) + e$^-$ → Ag (s) Eº = +0,80 volt Metals that protect the iron from corrosion using the cathodic protection method are... A. Ag B. Cu C. Ni D. Mg E. Pb
Tier 2	Are you sure of your answer? A. Sure B. Not sure
Tier 3	What is your reason for choosing the answer? A. Nickel (Ni) can protect iron from corrosion because its reduction potential (-0.25 V) is more favorable than iron (-0.44 V), so nickel is more easily oxidized and can act as an effective sacrificial anode. B. Magnesium (Mg) can protect iron from corrosion because its reduction potential (-2.37 V) is much more damaging than iron (-0.44 V), so magnesium is more easily oxidized and can act as an effective sacrificial anode. C. Copper (Cu) can protect iron from corrosion because its reduction potential (+0.34 V) is more favorable than iron (-0.44 V), so copper is more easily oxidized and can act as an effective sacrificial anode. D. Silver (Ag) can protect iron from corrosion because its reduction potential (+0.80 V) is much more positive than iron (-0.44 V), so silver oxidizes more easily and can act as an effective sacrificial anode. E. Lead (Pb) can protect iron from corrosion because its reduction potential (-0.13 V) is more favorable than iron (-0.44 V), so lead is more easily oxidized and can act as an effective sacrificial anode.
Tier 4	Are you sure of your answer? A. Sure B. Not sure
Tier 5	Where do you get the information from? A. Package book B. Internet C. Teacher D. Friends E. Parents

Tables 2 and 3 show the revision of questions related to cathodic protection, clarifying redactions, answer choices, and adding supporting data. One-to-one evaluations were carried out on three students with different abilities. The results of the interview revealed that several answer options in the reason tier (tier 3) felt similar, causing confusion. Some problems are also considered too difficult, especially those that involve complex calculations. The revised results of this evaluation are called prototype 3, which is then retested through a small group evaluation. Small group evaluations were carried out on six students with various levels of ability. The revision includes the addition of test filling instructions for each tier. The results of this evaluation are called prototype 4, which will then be tested in a field test.

2. Test Reliability

The reliability test was carried out based on the test results of questions involving 30 students in class XII B and 30 students in XII C at MAN. Data analysis using Ministep software version 5.8.3.8 produces a variety of reliability indicators that are important for evaluating the quality of test instruments. The results of the analysis are shown in the following figure.

```
    SUMMARY OF 30 MEASURED Person
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN       9.4      25.0      -.67      .49        .99   -.09   1.05    .03 |
|  SEM        .8        .0       .17      .02        .04    .23    .07    .23 |
| P.SD       4.1        .0       .92      .08        .22   1.23    .39   1.24 |
| S.SD       4.1        .0       .94      .09        .23   1.25    .39   1.26 |
| MAX.      17.0      25.0       .80      .76       1.42   2.32   2.25   2.62 |
| MIN.       2.0      25.0     -2.75      .43        .70  -2.26    .61  -2.11 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .52 TRUE SD    .77  SEPARATION  1.48  Person RELIABILITY  .69 |
|MODEL RMSE    .49 TRUE SD    .78  SEPARATION  1.58  Person RELIABILITY  .72 |
| S.E. OF Person MEAN = .17                                                   |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .71  SEM = 2.17
STANDARDIZED (50 ITEM) RELIABILITY = .83

-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      11.3      30.0       .00      .44        .99   -.06   1.05    .14 |
|  SEM       1.0        .0       .18      .01        .03    .18    .06    .21 |
| P.SD       4.9        .0       .87      .05        .15    .89    .32   1.01 |
| S.SD       5.0        .0       .89      .05        .16    .91    .32   1.03 |
| MAX.      22.0      30.0      1.82      .62       1.23   1.46   1.69   2.38 |
| MIN.       3.0      30.0     -1.85      .40        .66  -2.32    .54  -1.79 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .45 TRUE SD    .74  SEPARATION  1.65  Item   RELIABILITY  .73 |
|MODEL RMSE    .44 TRUE SD    .75  SEPARATION  1.71  Item   RELIABILITY  .75 |
| S.E. OF Item MEAN = .18                                                     |
-------------------------------------------------------------------------------
```

Figure 1. *Summary Statistics Person*

Based on Ministep's analysis, in class XII B the reliability of the participants was 0.69 and the reliability of the question items was 0.73, both of which were classified as moderate. Cronbach's Alpha score of 0.71 indicates sufficient internal consistency of the question, but not optimal. Person Separation of 1.48 is still inadequate in distinguishing students' abilities, while Item Separation of 1.65 is quite good in distinguishing the difficulty level of questions.

```
    SUMMARY OF 30 MEASURED Person
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      12.7      25.0       .01      .46        .99    .00   1.02    .06 |
|  SEM        .9        .0       .17      .01        .03    .16    .06    .18 |
| P.SD       4.6        .0       .91      .04        .15    .85    .33    .95 |
| S.SD       4.7        .0       .93      .04        .15    .87    .33    .97 |
| MAX.      20.0      25.0      1.55      .57       1.36   1.80   2.12   2.73 |
| MIN.       4.0      25.0     -1.85      .43        .77  -1.45    .60  -1.18 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .47 TRUE SD    .78  SEPARATION  1.64  Person RELIABILITY  .73 |
|MODEL RMSE    .46 TRUE SD    .79  SEPARATION  1.71  Person RELIABILITY  .74 |
| S.E. OF Person MEAN = .17                                                   |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .76  SEM = 2.24
STANDARDIZED (50 ITEM) RELIABILITY = .85
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      15.2      30.0       .00      .42       1.00   -.04   1.02   -.03 |
|  SEM        .9        .0       .16      .01        .04    .23    .06    .22 |
| P.SD       4.6        .0       .80      .03        .18   1.11    .31   1.05 |
| S.SD       4.7        .0       .82      .04        .18   1.13    .32   1.08 |
| MAX.      24.0      30.0      2.15      .55       1.48   2.67   2.11   2.53 |
| MIN.       4.0      30.0     -1.61      .40        .67  -2.61    .61  -2.25 |
|-----------------------------------------------------------------------------|
| REAL RMSE    .44 TRUE SD    .67  SEPARATION  1.54  Item   RELIABILITY  .70 |
|MODEL RMSE    .42 TRUE SD    .68  SEPARATION  1.62  Item   RELIABILITY  .72 |
| S.E. OF Item MEAN = .16                                                     |
-------------------------------------------------------------------------------
```

Figure 2. *Summary Statistics Person*

In class XII C, the reliability of the participants showed a value of 0.73, which was classified as moderate. This shows that the instrument has a moderate ability to distinguish students' level of ability. The reliability of the question item reached 0.70, which was also classified as moderate, indicating that the questions had a pretty good level of consistency but were not optimal. Cronbach's Alpha score of 0.76 indicates that the internal consistency of the questions is quite good but not optimal to achieve a higher standard. In addition, Person Separation has a score of 1.64, which shows that the test's ability to separate students based on their ability level is quite adequate. On the other hand, Item Separation of 1.54 is quite good, indicating that the problem can distinguish the difficulty level of the question item.

From the results of the reliability test in grades XII B and XII C, it can be concluded that the questions made have shown relatively good reliability, with reliability values ranging from 0.69 to 0.73, which shows the ability of the instrument to distinguish the ability of students moderately. The internal consistency of the questions, measured using Cronbach's Alpha, also showed a moderate result, with a score between 0.71 and 0.76, indicating that the questions were entirely consistent but could still be improved.

3. Misconceptions

Misconceptions are a major challenge in learning chemical concepts, particularly redox and electrochemical reactions, as they hinder problem-solving abilities. To assess the extent of misconceptions, an analysis of students' responses in classes B and C was conducted, categorizing them into understanding, misconceptions, and non-understanding. Table 4 highlights variations in comprehension between the two classes, helping identify areas that require further learning improvements. The following are the results of the percentage analysis of misconceptions in both classes:

Table 4. Percentage of Class B and C Misconceptions

Indicator	Understand		Misconceptions		Not Understood	
	B	C	B	C	B	C
Explain the working principle, principal components, and differences between voltaic and electrolysis cells.	3,33%	3,33%	70,00%	70,00%	26,67%	26,67%
Demonstrates the redox reaction products that occur in the anode and cathode during electrolysis of various solutions.	20,00%	20,00%	46,67%	46,67%	33,33%	33,33%
Understanding how the environment affects the rate of iron corrosion.	63,33%	63,33%	13,33%	13,33%	23,33%	23,33%
Interpreting the results of the cell's Eº value to predict whether or not the voltaic cell reaction will occur spontaneously.	13,33%	13,33%	33,33%	33,33%	53,33%	53,33%
Predicting the cell diagram that occurs in a series of voltaic cells.	3,33%	3,33%	60,00%	60,00%	36,67%	36,67%
Apply electrolysis concepts to explain the metal plating process, including ion migration and reactions occurring in the electrodes.	10,00%	10,00%	26,67%	26,67%	63,33%	63,33%
Calculate the rate of corrosion on a given material under specific environmental conditions.	20,00%	20,00%	23,33%	23,33%	56,67%	56,67%
Calculating the Standard Cell Potential (E° of the cell) using the Nernst Equation	30,00%	30,00%	40,00%	40,00%	30,00%	30,00%
Evaluate the standard electrode potential concept to select the right metals in cathodic protection and predict the outcome of electrochemical reactions between two metals in solution.	6,67%	6,67%	46,67%	46,67%	46,67%	46,67%
Analyze redox reactions occurring in anodes and cathodes in electrolysis cells with inert and non-inert electrodes	16,67%	16,67%	23,33%	23,33%	60,00%	60,00%
Faraday's law explains the relationship between the electric current used and the corrosion rate observed.	13,33%	13,33%	26,67%	26,67%	60,00%	60,00%
Analyzing the concept of volta cells in fuel cell circuits based on images.	20,00%	20,00%	40,00%	40,00%	40,00%	40,00%
Demonstrates the redox reaction products that occur in the anode and cathode during electrolysis of various solutions.	6,67%	6,67%	30,00%	30,00%	63,33%	63,33%
Interpreting the results of the cell's Eº value to predict whether or not the voltaic cell reaction will occur spontaneously.	16,67%	16,67%	33,33%	33,33%	50,00%	50,00%
Evaluate the standard electrode potential concept to select the right metals in cathodic protection and predict the outcome of electrochemical reactions between two metals in solution.	10,00%	10,00%	46,67%	46,67%	43,33%	43,33%

Table 4 reveals significant differences in students' understanding, misconceptions, and non-understanding between classes B and C. In class B, comprehension was low (16.89%), with high misconceptions (37.33%) and non-understanding (45.78%). The highest understanding was in "How the environment affects iron corrosion" (63.33%), while the lowest was in "Predicting cell diagrams in voltaic cells" (3.33%). In class C, understanding (33.11%), misconceptions (32.89%), and non-understanding (34.00%) were more balanced. Despite these differences, both classes struggle with misconceptions, emphasizing the need for a more effective learning approach and better analysis of students' information sources to improve electrochemistry understanding. Table 5 shows the percentage of information sources that students access for each question:

Table 5. Percentage of Class B and C Information Sources

No Question	Package Book		Internet		Teacher		Friends		Parents	
	B	C	B	C	B	C	B	C	B	C
1	16,67%	16,67%	60,00%	60,00%	20,00%	10,00%	3,33%	13,33%	0,00%	0,00%
2	6,67%	26,67%	50,00%	46,67%	16,67%	6,67%	26,67%	20,00%	0,00%	0,00%
3	3,33%	16,67%	26,67%	40,00%	6,67%	10,00%	63,33%	33,33%	0,00%	0,00%
4	10,00%	13,33%	83,33%	50,00%	3,33%	6,67%	3,33%	30,00%	0,00%	0,00%
5	10,00%	10,00%	46,67%	40,00%	6,67%	13,33%	36,67%	36,67%	0,00%	0,00%
6	16,67%	6,67%	33,33%	66,67%	6,67%	10,00%	43,33%	16,67%	0,00%	0,00%
7	10,00%	3,33%	20,00%	83,33%	0,00%	3,33%	70,00%	10,00%	0,00%	0,00%
8	23,33%	10,00%	60,00%	33,33%	6,67%	10,00%	10,00%	46,67%	0,00%	0,00%
9	3,33%	0,00%	30,00%	70,00%	3,33%	3,33%	63,33%	26,67%	0,00%	0,00%
10	13,33%	6,67%	56,67%	40,00%	0,00%	3,33%	30,00%	50,00%	0,00%	0,00%
11	3,33%	13,33%	50,00%	60,00%	26,67%	10,00%	20,00%	16,67%	0,00%	0,00%
12	6,67%	0,00%	40,00%	33,33%	20,00%	26,67%	33,33%	40,00%	0,00%	0,00%
13	3,33%	3,33%	43,33%	33,33%	6,67%	33,33%	46,67%	30,00%	0,00%	0,00%
14	6,67%	26,67%	53,33%	40,00%	16,67%	23,33%	23,33%	10,00%	0,00%	0,00%
15	26,67%	3,33%	50,00%	50,00%	13,33%	6,67%	10,00%	40,00%	0,00%	0,00%

Table 5 shows that students in classes B and C mainly rely on the internet for information, with usage ranging from 50% to 83%. Despite this, misconceptions remain high, indicating a lack of verification. Peer discussions are also common, particularly in questions 3 and 9 for class B (63.33%) and questions 3 and 8 for class C (33.33% and 46.67%). However, peer discussions can spread misinformation. Textbooks and teachers are less frequently used, though they provide accurate explanations. Parents are not used as a learning resource at all. While class C has a more balanced understanding, misconceptions persist in both classes. Strengthening the role of teachers and textbooks, guiding students in evaluating online sources, and monitoring peer discussions can help improve understanding and reduce misconceptions in electrochemical concepts.

D. CONCLUSIONS AND SUGGESTIONS

The conclusion of this study shows that a pictorial-based five-tier diagnostic test was developed with 30 valid and reliable questions to identify misconceptions in electrochemical cell matter. The average Content Validity Index (CVI) value of 1.00 indicates that this instrument is very suitable for use. In contrast, the reliability with Cronbach's Alpha value ranging from 0.71 to 0.76 shows quite good consistency, although it can still be improved. Most students have misconceptions about essential concepts, such as the working principle of voltaic cells and electrolysis, electron flow, and prediction of reaction spontaneity using the Eº value of the cell. The primary source of this misconception comes from misdirected information, especially from the internet and discussions with friends. Visual elements in diagnostic tests help understand abstract concepts, but their effectiveness depends on the instructions' clarity and the questions' context.

As a suggestion, Revision is needed on complex questions to improve clarity. Teachers are expected to provide structured explanations and integrate diagnostic data in learning. Students are advised to use package books and reliable sources. In addition, training on the use of diagnostic tests needs to be carried out so that misconceptions can be overcome.

ACKNOWLEDGMENTS

The author would like to thank all parties who have supported this research, especially Sunan Kalijaga State Islamic University Yogyakarta, supervisors, validators (chemistry lecturers and teachers), as well as students and teachers of MAN 3 Bantul. The award is also addressed to the family and colleagues for their support and motivation. Hopefully this research will be useful for the world of education, especially in improving the understanding of chemical concepts.

LIST OF REFERENCES

Abd.Rachman, F., Ahsanunnisa, R., & Nawawi, E. (2017). Development of LKPD Based on Critical Thinking of Solubility Materials and Solubility Results in Chemistry Subjects in High School. *ALKIMIA : Jurnal Ilmu Kimia Dan Terapan*, *1*(1), 16–25. https://doi.org/10.19109/alkimia.v1i1.1326

Ahmad, N. J., & Lah, Y. C. (2013). A Designed Teaching Sequence as a Tool to Improve Students ' Conceptual Understanding of the Conductivity in the Electrolytic Cell A Designed Teaching Sequence as a Tool to Improve Students ' Conceptual Understanding of the Conductivity in the Electrolyte. *Canadian Center of Science and Education*, *9*(2). https://doi.org/10.5539/ass.v9n2p298

Akker, J. an van den, Branch, R. M., Gustafson, K., Nieveen, N., & Plomp, T. (1999). Design Approaches and Tools in Education and Training. In *Stochastic Environmental Research and Risk Assessment* (Vol. 29, Issue 7). https://doi.org/10.1007/s00477-014-0937-9

Annisa, R., Astuti, B., & Mindyarto, B. N. (2019). Four Tier Diagnostic Test for Identification of Students' Understanding and Misconceptions on Regular Circular Motion Material. *Jurnal Pendidikan Fisika Dan Kelilmuan (JPFK)*, *5*(1), 25–32. https://doi.org/http://doi.org/10.25273/jpfk.v5i1.3546

Budiharti, R., Radiyono, Y., Rizky, N., Nuraini, A., Putri, H. V., Saputro, D. E., & Adhitama, E. (2018). Development of a Four-Stage Misconception Diagnostic Test on Kinematics. *Jurnal Ilmiah Pendidikan*,*2*,237–249. https://doi.org/http://dx.doi.org/10.21831/cp.v37i2.16491

Hanson, R. (2015). Identifying Students' Alternative Concepts in Basic Chemical Bonding – A Case Study of Teacher Trainees in the University of Education, Winneba. *International Journal of Innovative Research and Development*, *4*(1), 115–122.

Hermita, N., Suhandi, A., Syaodih, E., Samsudin, A., Isjoni, Johan, H., Rosa, F., Setyaningsih, R., Sapriadil, & Safitri, D. (2017). Constructing and Implementing a Four-Tier Test about Static Electricity to Diagnose Pre-service Elementary School Teacher Misconceptions. *International Conference on Mathematics and Science Education*. https://doi.org/10.1088/1742-6596/895/1/012167

Inggit, S. M., Liliawati, W., & Suryana, I. (2021). Identification of Misconceptions and Their Causes Using the Five-Tier Fluid Static Test (5TFST) Instrument in Grade XI Students of Senior High School. *Journal of Teaching and Learning Physics*, *6*(1), 49–68. https://doi.org/10.15575/jotalp.v6i1.11016

Kaltakci-gurel, D., Eryilmaz, A., & Mcdermott, L. C. (2017). Development and application of a four-tier test to assess pre-service physics teachers ' misconceptions about geometrical optics. *Research in Science & Technological Education*, *5143*(April), 1–23.

Kristiana, E., Sidauruk, S., & Meiliawati, R. (2020). Difficulties of Class X MIA State High School Students in Palangka Raya City for the 2018/2019 Academic Year in Understanding the Concept of Lewis Structure Using the Two-Tier Multiple Choiche Instrument. *Jurnal Ilmiah Kanderang Tingang*, *11*(1), 200–208. https://doi.org/10.37304/jikt.v11i1.87

Martin Tessmer. (1993). *Planning and Conducting Formative Evaluations*. Routledge Taylor & Francis Group.

Muderawan, I. W., Wiratma, I. G. L., & Nabila, M. Z. (2019). Analysis of the Factors Causing Students' Learning Difficulties in Solubility Materials and Solubility Results. *Jurnal Pendidikan Kimia Indonesia*, *3*(1), 17. https://doi.org/10.23887/jpk.v3i1.20944

Nuraisyah, S. A., Nahadi, & Hernan. (2022). Profile of Misconception of High School Students in the City of Bandung on Salt Hydrolysis Materials Using a Pictorial-Based Two-Level Multiple-Choice Diagnostic Test. *Jurnal Riset Dan Praktik Pendidikan Kimia*, *10*(1), 77–85.

Nurlita, S. (2022). Development of four-tier diagnostic tests based on science process skills to identify misconceptions in redox materials. *Journal of Natural Science Learning*, *02*(02), 90–103.

Permana, I., Permata, F., Latifah, U., Al-anshori, J., & Rachman, S. D. (2019). Efforts to Improve Understanding of Chemistry and Its Application Through Socialization and Practicum Methods. *Jurnal Pengabdian Kepada Masyarakat*, *4*(6), 129–132.

Putri, E. S., & Rinaningsih. (2021). Review: Diagnostic Tests as Formative Tests in Chemistry Learning. *Journal of Chemical Education*, *10*(1), 20–27.

Rahayu, S., Treagust, D. F., Chandrasegaran, A. L., Kita, M., & Ibnua, S. (2011). Assessment of electrochemical concepts: A comparative study involving senior high-school students in Indonesia and Japan. *Research in Science and Technological Education*, *29*(2), 169–188. https://doi.org/10.1080/02635143.2010.536949

Roghdah, S. J., Zammi, M., & Mardhiya, J. (2021). Development of Four-Tier Multiple Choice Diagnostic Test to Determine the Level of Students' Understanding of Concepts in Thermochemical Materials. *Phenomenon : Jurnal Pendidikan MIPA*, *11*(1), 57–74.

Rokhim, D. A., Rahayu, S., & Dasna, I. W. (2023). Analysis of Chemical Misconceptions and Their Diagnostic Instruments: Literature Review. *Jurnal Inovasi Pendidikan Kimia*, *17*(1), 17–28. https://doi.org/10.15294/jipk.v17i1.34245

Rosita, I., Liliawati, W., & Samsudin, A. (2020). Development of Five-Tier Newton's Laws Test (5TNLT) Instrument to Identify Students' Misconceptions and Causes of Misconceptions. *Jurnal Pendidikan Fisika Dan Teknologi (JPFT)*, *6*(2). https://doi.org/https://doi.org/10.29303/jpft.v6i2.2018

Santi, A. N. I., & Rahayu, M. (2022). Analysis of Students' Misconceptions on Electrolyte and Non-Electrolyte Solution Materials Using Pictorial-Based Four-Tier Diagnostic Tesr Multirepresentation Instrument. *Journal of Chemical Education*, *11*(3), 210–219.

Sari, D. A. (2021). Application of Inquiry-Based Learning to Electrochemistry Materials on Conceptual Understanding, Mental Models and Student Attitudes. *Orbital: Jurnal Pendidikan Kimia*, *5*(2), 137–150. https://doi.org/10.19109/ojpk.v5i2.9178

Savira, I., Wardani, S., Harjito, & Noorhayati, A. (2019). Design of Three Tiers Multiple Choice Test Instruments for the Analysis of Student Misconceptions Regarding Buffer Solutions. *Jurnal Inovasi Pendidikan Kimia*, *13*(1), 2277–2286. https://doi.org/https://doi.org/10.15294/jipk.v13i1.15924

Simamora, R., Maison, & Kurniawan, W. (2023). Identification of Misconceptions of Students Using Five-Tier Diagnostic Test on Static Fluid Materials at SMAN 7 Jambi City. *Jurnal Penelitian Pendidikan Fisika*, *8*(2), 139–144. https://doi.org/10.36709/jipfi.v8i2.18

Sugiyono. (2013). *Quantitative, Qualitative and R&D Research Methods* (19th). CV Alfabeta.

Sugiyono. (2022). *Quantitative, Qualitative, and R&D Research Methods*. Penerbit Alfabeta.

Tavassoli, A., Jahandar, S., & Khodabandehlou, M. (2013). The Effect of Pictorial Contexts on Reading Comprehension of Iranian High School Students : a Comparison Between Pre-Vs. During Reading Activities. *Indian Journal of Fundamental and Applied Life Sciences*, *3*(3), 553–565.

Virgiawan, W., Nahadi, Kusrijadi, A., Siswaningsih, W., & Rahmawati, T. (2020). Profile of Misconception of State High School Students in the City of Bandung on Mol Concept Material Using a Pictorial-Based Two-Tier Diagnostic Test. *Jurnal Riset Dan Praktik Pendidikan Kimia*, *8*(1).

Wahyuni, S., Maison, & Hidayat, M. (2022). Identification of Misconceptions of the Five Tier Diagnostic Test on Energy Matter and the Law of Energy Conservation. *Jurnal Metaedukasi: Jurnal Ilmiah Pendidikan*, *4*(1), 45–53.