STEM Education International,
Vol. 1, No. 2, December 2025, Page 80-86
p-ISSN: xxxx-xxx, e-ISSN: xxx-xxx

STEM-EI
STEM Education International

Development of Four-Tier Diagnostic Instruments to Identify Students' Understanding of Electrolyte and Non-Electrolyte Solutions

Septiana Ayuningrum Nofinadya*[1], Rima Nuraini[1], Isnanik Juni Fitriyah[2]

[1]Department of Chemistry, Universitas Negeri Malang, Indonesia 65145
[2]Department of Science Education, Universitas Negeri Malang, Indonesia 65145

**Corresponding author: nadyasepti13@gmail.com*

Abstract: This paper outlines the level of validity and reliability of a four-tier diagnostic instrument to identify the level of understanding of students on electrolyte and non-electrolyte solutions. The instrument development follows the stages proposed by Habiddin & Page (2019), namely mapping concepts, testing and interviewing, defining students' unscientific ideas, developing a four-tier prototype, validating the four-tier prototype, and refining the final four-tier instrument. The product eligibility percentage is 89.08%, which falls within the very feasible category.

Keywords: development, four-tier, electrolyte and non-electrolyte solutions, students' conception, unscientific understanding

INTRODUCTION

Difficulties and misconceptions among students regarding electrolyte and non-electrolyte solutions have been reported in the literature. A study in Padang found that 22.5% students from a public secondary school experienced misconceptions in this topic (Fany & Ulianas, 2021). Another study in Kalimantan also found that many students experienced misconceptions in the colligative properties of solutions (Pratiwi et al., 2023). While in another study, 64.70% of students experienced misconceptions about ionic compounds in electrolyte and non-electrolyte solutions (Siswaningsih et al., 2015). Misconceptions can be interpreted as an understanding of concepts that are incompatible with the concept accepted by the scientific community (Habiddin & Nofinadya, 2021). Students cannot explain the theoretical basis of electrolyte and non-electrolyte solutions (Rahmadhany et al., 2023). Or, in other words, students only memorise, remember, and guess to answer questions about related concepts.

Students' misconceptions can affect the learning process. In general, students still maintain misconceptions even after experiencing a different learning process from the previous one (Kirik & Boz, 2012). To identify these misconceptions, an appropriate instrument is required. Multiple-choice and description questions, which are generally used as learning evaluation tools, can be used to measure students' level of understanding. However, both types of instruments have drawbacks. The weakness of the description questions is that there is a subjective tendency when analysing students' answers, and the questions contained in the description questions cannot cover all the material that has been conveyed in learning. Multiple-choice questions, although considered more efficient for identifying students' understanding, also have weaknesses. Multiple-choice questions provide a greater opportunity for students to guess the correct answer than do description questions (Simkin & Kuechler, 2005). Another weakness is that this type of instrument can only evaluate students' content knowledge without considering the reasons behind their choice of answers (Chandrasegaran et al., 2007).

Given these limitations, other instruments are used to assess students' understanding of concepts. The two-tier multiple-choice instrument was among the first to be developed. The instrument consists of two levels: the first is multiple-choice questions, and the second is the reason for choosing an answer at the first level. The weakness of this type of instrument is that it can only identify students' mistakes when they experience misconceptions (Pesman & Eryilmaz, 2010). Then, the form is developed by adding confidence to each item or by using a three-tier multiple-choice instrument. This instrument also has a weakness: students are allowed to choose only a single confidence level. The level of confidence of these learners applies to the answers and reasons for each item. Thus, it cannot be distinguished between students who are sure of the answers and the reasons they choose, and students who are only sure of the answers and not of the reasons, or vice versa. This results in difficulties in assessing and analysing students' answers (Arslan et al., 2012). Based on this description, the researcher wants to develop a four-tier multiple-choice diagnostic instrument. This results in difficulties in assessing and analysing students' answers. Based on this description, the researcher wants to develop a four-tier multiple-choice diagnostic instrument. This results in challenges in determining and analysing students' answers (Arslan et al., 2012). Based on this description, the researcher wants to develop a four-tier multiple-choice diagnostic instrument.

The four-tier diagnostic test is a four-level test, with each question accompanied by a question about the student's level of confidence in each answer and the reason they chose that answer. This instrument allows students to select different levels of confidence in their answers and reasons, so that researchers can clearly determine each student's level of understanding (Habiddin & Page, 2019). This diagnostic test has several advantages, among others. Through a diagnostic test, the teacher can gain a deeper understanding of students' misconceptions. The teacher can also emphasise certain parts of the material that require this when they are explained, and then plan the steps, better learning steps to reduce student misconceptions. Several studies have employed this type of instrument to identify students' understanding and misunderstanding in chemistry (Ardina & Habiddin, 2023; Habiddin et al., 2020, 2022; Husniah et al., 2019).

METHOD

This research and development uses a procedure developed by Habiddin & Page (2019) with six main steps, namely (1) mapping concept, (2) testing and interviewing, (3) defining students' unscientific ideas, (4) developing the prototype of a four-tier, (5) validating the prototype four-tier, (6) refining the final four-tier. At the mapping concept, concept identification, question grid arrangement, and open-reason multiple-choice instruments were prepared. The concept identification stage aims to identify, describe and compile concepts that will identify possible misconceptions. The lecturers and chemistry teachers validated the open-ended multiple-choice instrument developed in the previous step. The feasibility assessment is based on nine criteria/components. The suggestions and comments given are used as the basis for improving or revising the open-reason multiple-choice instrument.

The testing and interviewing stage is a preliminary test conducted using a validated and revised multiple-choice instrument for open-ended responses. The initial data collection was conducted with students at the SMA/MA level who had taken the subject of electrolyte and non-electrolyte solutions. This initial data collection aims to identify and collect students' non-scientific understanding, so that this understanding can be used as an alternative concept in the development of a four-tier instrument.

Based on the data obtained from the preliminary test in the previous stage, an analysis of the answers and reasons was conducted during the stage of defining students' unscientific ideas. All collected reasons were classified into four categories, namely true concepts, false concepts, random or guesswork errors, and no reason.

The development of the prototype four-tier stage begins with preparing and developing a four-tier diagnostic instrument. Then analyse the items used in the initial data collection. This is done to know and correct invalid items. The four-tier diagnostic instrument has been validated to determine its feasibility. The validators referred to in this study were chemistry lecturers of FMIPA UM and chemistry teachers. The validation sheet at this stage differs from the one for multiple-choice open-ended questions because it consists of 10 assessment criteria. The validation results are used as a reference to improve the four-tier diagnostic instrument. The results of the validation, namely the validator's suggestions and comments as qualitative data, while quantitative data is a score given by the validator for each item. After the prototype of the four-tier diagnostic instrument has been validated by the validator, the next step is to validate the empirical (validating the prototype). The parameters used in the empirical validation are reliability and validity, Item difficulty level, item difference power and distractor effectiveness. Based on the results of empirical validation, improvements/revisions were made to the four-tier diagnostic instrument to identify students' level of understanding of electrolyte and non-electrolyte solutions, or the final four-tier refinement stage.

RESULTS AND DISCUSSION

Research and Development Results

The development of a four-tier diagnostic instrument began with an open-ended multiple-choice instrument. The development of an open-reason multiple-choice instrument began with mapping the concepts of electrolyte and non-electrolyte

solutions, which served as the basis for formulating 12 indicators of competency achievement (GPA). The competency attainment indicators developed do not measure algorithmic aspects but focus more on conceptual elements. Each predetermined competency achievement indicator is then developed into question indicators. Thus, the number of questions for the multiple-choice open-ended instrument being developed also amounted to 22 items. The open-reason multiple-choice instrument contains a question and four answer choices, but there are also items with three or two answer choices. This is because in these questions, it is complicated to make four homogeneous answer choices. The multiple-choice instrument for the reasons was then validated by two validators: a lecturer in the Department of Chemistry, Mathematics and Natural Sciences at UM and a High School Chemistry Teacher. Meanwhile, the classification results from the initial data collection are used to select reasons or tier reasons for the prototype of the four-tier diagnostic instrument that will be developed. The choice of reasons consisted of one genuine concept and three false concepts, with considerations: concepts that contained misconceptions, logically incorrect concepts, concepts with high frequency, and concepts with a confidence rating. Then, the questions on the multiple-choice instrument that allowed open-ended responses were analysed to determine which items were feasible to develop into a prototype four-tier diagnostic instrument. The results of the analysis of the items in the multiple-choice instrument for open-ended responses were based on several parameters, namely validity, reliability, difficulty level, and item differences.

Table 1. Example of Development Results for a Four-tier Diagnostic Instrument

Indicator of Competence: Identify the type of solution based on its conductivity.		
Indicator	**Question**	**Answer**
Students can identify non-electrolyte solutions from the brief information given.	*FIRST TIER* Sucrose, with the molecular formula $C_{12}H_{22}O_{11}$, is dissolved in water. If viewed from the electrical conductivity, the solution is ... A. Strong electrolyte solution B. Weak electrolyte solution C. Non-electrolyte solution *SECOND TIER* The level of confidence in the selected answer: 1) Just guessing 2) Not sure 3) Moderate 4) Sure 5) Very sure	C

The product of this research and development is a four-tier diagnostic instrument to assess students' conceptual understanding of electrolyte and non-electrolyte solutions, consisting of 13 items. The resulting product specifications are (1) a four-tier diagnostic instrument developed based on KD. 3.8; (2) the developed four-tier diagnostic instrument is equipped with instructions on how to work on questions and a grid of questions consisting of basic competencies, the intended material, indicators of competency achievement, indicators of questions, items; (3) the developed four-tier diagnostic instrument consists of four levels, the first tier is in the form of questions with several answer choices, the second tier is the level of students' confidence in the answers they choose on a scale of 1-5 (1 = only guessing, 2 = not sure, 3 = moderate, 4 = sure and 5 = very sure), the third tier is a choice of four reasons for the first tier, and the fourth tier is the level of confidence of students for reasons on a scale of 1-5. An example of the results of developing a four-tier diagnostic instrument prototype is shown in the Table 2.

Table 2. Example of Development Results for a Four-tier Diagnostic Instrument

Indicators of Competence: Identify the type of solution based on its conductivity					
Question					**Answer**
THIRD TIER Which is the correct reason for the answer chosen? A. $C_{12}H_{22}O_{11}$ cannot be ionised in water, so it cannot conduct electricity. B. $C_{12}H_{22}O_{11}$ has a small number of ions in the solution. C. $C_{12}H_{22}O_{11}$ has ions that can move freely in the solution. D. $C_{12}H_{22}O_{11}$ ionises completely in water so that it can conduct electricity.					A
FOURTH TIER The level of confidence in the selected answer:					
1) Just guessing	2) Not sure	3) Moderate	4) Sure	5) Very sure	

Only 13 of the previous 22 items were developed. This occurs because the reasons derived from respondents'/students' answers are insufficient to serve as the third tier of a four-tier diagnostic instrument prototype. The four-tier diagnostic instrument that had been developed was then corrected based on the content validation assessment by two validators: a lecturer in Chemistry at FMIPA UM and a high school chemistry teacher. This was done with the aim that the developed four-tier diagnostic instrument was in accordance with the realms of the material, the realm of construction, and the realm of language. Several items need to be corrected in response to the validator's comments and suggestions.

The revised four-tier diagnostic instrument was subsequently empirically validated, involving 62 respondents from two Class X classes at one of the public high schools in Ponorogo. Before analysing the items based on the empirical validation results, students' scores were calculated for each item. The data were divided into three: the scoring results for the selection of answers (tier answer), the scoring results for the selection of reasons (tier reason), and the scoring results for the second selection (both tiers). The guidelines for scoring the answers and reasons are presented in Table 3.

Table 3. Scoring Guidelines (Habiddin & Page, 2019)

Tier Answer **(A tier)**	*Tier Reason* **(R tier)**	**A combination of A tier and R tier**
Yes (1)	Yes (1)	1
Yes (1)	False (0)	0
False (0)	Yes (1)	0
False (0)	False (0)	0

Data collection was conducted online using Google Forms due to conditions that prevented in-person collection in classrooms because of the spread of the COVID-19 virus. The four-tier diagnostic instrument for identifying students' conceptual understanding of the developed electrolyte and non-electrolyte solution has advantages and disadvantages. The benefits of these four-tier diagnostic instruments include (1) four-tier diagnostic instruments are still not widely developed in chemistry education, especially in electrolyte and non-electrolyte solutions, (2) four-tier diagnostic instruments can be used as tools to identify students' conceptual understanding of electrolyte and non-electrolyte solution material. The weaknesses of the developed four-tier diagnostic instrument include the number of answer choices and the reason for choosing four tiers rather than five, such as the number of answer choices at the high school level.

Results of Content Validation and Empirical Validation

Before developing a four-tier diagnostic instrument as a final product, an open-ended multiple-choice instrument was developed, involving 153 respondents from two X classes at SMAN 3 Malang and three X classes at SMAN 8 Malang, with the criterion that they had taken electrolyte and non-solution material. Electrolyte. The instrument's reliability was in the high category, and all items were valid. In the difficulty level test, there were 5 questions in the easy category, 16 in the medium category, and 1 in the difficult category. Meanwhile, the difference calculation showed that 20 items were in the good category and 2 were in the enough category.

For the developed four-tier diagnostic instrument, both content and empirical validation were carried out. The content validation results showed that 13 items were valid, including those in the very feasible and feasible categories. Meanwhile, the empirical validation of the four-tier diagnostic instrument involved 62 respondents from two X classes at SMAN 2 Ponorogo who had studied electrolyte and non-electrolyte solution materials. The calculation of the parameters used in the empirical validation was divided into scoring results for the selection of answers (tier answer), for the selection of reasons (tier reason), and for the second selection (both tiers). The reliability test of the 13 items showed that the three tiers fell within the high category. The validity of the items on all three tiers was categorised as valid. In the analysis of the item difficulty level across the three tiers, questions were categorised as easy, medium, or difficult. Meanwhile, for the test item difference in the three tiers, it was categorised as sufficient, good and very good. The results of the calculation of the effectiveness of the distractor for scoring the choice of answers (tier answer) and the selection of reasons (tier reason) fell within the same percentage range, namely 6.45% - 72.58%.

CONCLUSIONS

The four-tier diagnostic instrument developed consisted of 13 items with electrolyte and non-electrolyte solutions. Among other advantages, this four-tier diagnostic instrument can serve as an evaluation tool to identify students' conceptual understanding of electrolyte and non-electrolyte solutions. The weakness of the developed four-tier diagnostic instrument, among others, is that the number of answer choices and reasons for the four-tier diagnostic instrument is less than five, such as the number of answer choices at the high school level, so that it can be used as a tool for evaluating learning outcomes based on the school curriculum; adjustments need to be made. Further research is required to disseminate this four-tier diagnostic instrument more widely. This was done to obtain a large number of respondents to assess the product's effectiveness in identifying students' understanding of concepts in electrolyte and non-electrolyte solutions.

REFERENCES

Ardina, D., & Habiddin, H. (2023). Acid-base properties of salt solution: Study at a secondary school in Banyuwangi. *AIP Conference Proceedings*, *2569*(1), 30018. https://doi.org/10.1063/5.0112074

Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A Three-Tier Diagnostic Test to Assess

Pre-Service Teachers' Misconceptions about Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain. *International Journal of Science Education*, *34*(11), 1667–1686.

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice*, *8*(3), 293–307.

Fany, L. A., & Ulianas, A. (2021). Analysis of Students Misconception using Two-Tier Multiple Choice Diagnostic Test on Electrolyte and Nonelectrolyte Topic in SMAN 2 Padang. *International Journal of Progressive Sciences and Technologies; Vol 29, No 1 (2021)*. https://doi.org/10.52155/ijpsat.v29.1.3403

Habiddin, H., Akbar, D. F. K., Husniah, I., & Luna, P. (2022). Uncovering Students' Understanding: Evidence for the Teaching of Acid-Base Properties of Salt Solution. *Educacion Quimica*, *33*(1), 64–76. https://doi.org/10.22201/fq.18708404e.2022.1.79488

Habiddin, H., Ameliana, D. N., & Su'aidy, M. (2020). Development of a Four-Tier Instrument of Acid-Base Properties of Salt Solution. *JCER (Journal of Chemistry Education Research)*, *4*(1), 51–57. https://doi.org/10.26740/jcer.v4n1.p51-57

Habiddin, H., & Nofinadya, S. A. (2021). The Multi-Tier Instrument in the Area of Chemistry and Science. In *Insights Into Global Engineering Education After the Birth of Industry 5.0* (pp. 1–17). IntechOpen. https://doi.org/10.5772/INTECHOPEN.100098

Habiddin, H., & Page, E. M. (2019). Development and validation of a four-tier diagnostic instrument for chemical kinetics (FTDICK). *Indonesian Journal of Chemistry*, *19*(3), 720–736. https://doi.org/10.22146/ijc.39218

Husniah, I., Habiddin, H., Sua'idy, M., & Nuryono, N. (2019). Validating an instrument to investigate students' conception of Salt hydrolysis. *Journal of Disruptive Learning Innovation (JODLI)*, *1*(1), 1–6.

Kirik, O. T., & Boz, Y. (2012). Cooperative learning instruction for conceptual change in the concepts of chemical kinetics. *Chemistry Education Research and Practice*, *13*(3), 221–236.

Pesman, H., & Eryilmaz, A. (2010). Development of a Three-Tier Test to Assess Misconceptions About Simple Electric Circuits. *Journal of Educational Research*, *103*(3), 208–222.

Pratiwi, A. N., Erlina, E., Lestari, I., Masriani, M., & Rasmawan, R. (2023). Identification of Students' Misconceptions Using a Four-Tier Multiple Choice Diagnostic Test on Colligative Properties of Solutions. *Jurnal Penelitian Pendidikan IPA*, *9*(11 SE-Research Articles), 10033–10042. https://doi.org/10.29303/jppipa.v9i11.4018

Rahmadhany, D. N., Putri, H. Y., Raudah, N. A., & Arianisa, S. (2023). Analisis Pemahaman Siswa SMA Pada Materi Larutan Elektrolit Dan Non Elektrolit. *Jurnal Sadewa : Publikasi Ilmu Pendidikan, Pembelajaran Dan Ilmu Sosial*, *2*(1 SE-Articles), 209–216. https://doi.org/10.61132/sadewa.v2i1.497

Simkin, M. G., & Kuechler, W. L. (2005). Multiple-Choice Tests and Student Understanding: What Is the Connection? *Decision Sciences Journal of Innovative Education*, *3*(1), 73–98. https://doi.org/https://doi.org/10.1111/j.1540-4609.2005.00053.x

Siswaningsih, W., Firman, H., & Rofifah, R. (2015). Pengembangan Tes Diagnostik Two-Tier Berbasis Piktorial Untuk Mengidentifikasi Miskonsepsi Siswa Pada Materi Larutan Elektrolit dan Nonelektrolit. *Jurnal Pengajaran Matematika Dan Ilmu Pengetahuan Alam (JPMIPA)*, *20*(2), 144–149. https://doi.org/https://doi.org/10.18269/jpmipa.v20i2.36236