Journal of Educational Sciences
Journal homepage: https://jes.ejournal.unri.ac.id/index.php/JES

P-ISSN 2581-1657
E-ISSN 2581-2203

Development of Multiple Representation Based Assessment Instruments to Measure Students' Conceptual Understanding of Buffer Solution Material

Rehamita. S*, Roza Linda, Dedi Futra
Postgraduate of Chemistry Education, University of Riau, Pekanbaru 28293, Indonesia.

ARTICLE INFO

Article history:

Received: 25 June 2024
Revised: 26 Sept 2024
Accepted: 02 Oct 2024
Published online: 26 Oct 2024

Keywords:

Assessment Instrument;
Buffer Solution;
Multiple Representation;
Concept Understanding

A B S T R A C T

Buffer solution material was chosen in the process of developing an instrument based on multiple representations, because the buffer solution material contains many representations. Buffer solution material requires explanations in various forms of representation that can visualize the buffer solution material so that students are expected to be able to observe the symptoms that occur and analyze and draw more comprehensive conclusions. This study aims to measure students' conceptual understanding of buffer solution material based on aspects of material, language, construction, and conceptual understanding. The method used is Research and Development with the ADDIE model, involving two schools, SMAN 1 and SMAN 7 Pekanbaru. The validation results by three expert lecturers showed Aiken's (V) values ranging from 0.88-0.97, with instrument reliability stated as good (Cronbach's alpha = 0.815). The level of difficulty of the questions consisted of 4% easy, 84% moderate, and 12% difficult. Measurement of conceptual understanding in 50 students showed a significant difference with a significance value of 0.012 <0.05, so H0 was rejected. The assessment instrument was declared valid and reliable for buffer solution material.

1. Introduction

National Education in Indonesia is organized based on Pancasila and the 1945 Constitution which functions to develop the potential and character of students to become people who are faithful, pious, have noble character, are independent, knowledgeable, think critically, are diverse, work together and are creative (RUU Sisdiknas, 2022). The Education System in Indonesia continues to be improved, one of which is the curriculum. Curriculum changes aim to improve the quality of the learning process and learning designs in schools (Masykur, 2019). Indonesia

* Corresponding author.
E-mail: rehamitas@gmail.com
Doi: https://doi.org/10.31258/jes.8.4.p.671-689

has experienced several changes in the education curriculum, one of the goals of which is to adapt education to the development and progress of the times so that Indonesian education can compete internationally and develop in accordance with the times. Currently, the government has issued a new policy on education in Indonesia, namely the Independent Learning Curriculum (Wulan, 2023).

The independent learning curriculum is intended to improve the quality of education in Indonesia so that it can create human resources who are not only superior in academics, but also have good character. The independent learning curriculum can be applied in every subject, including Chemistry. In chemistry, the things that are studied are materials, properties of materials, how and why substances combine or separate to form other substances, and the energy that accompanies their changes (Ministry of Education and Culture, 2022). One of the objectives of learning, especially in chemistry, is that students can understand concepts, principles, laws, and theories well and are interrelated (Wasonowati, 2014). Education is the most important factor in a person's life, because it can distinguish a person's ability to think (Bella, 2023).

Conceptual understanding is the ability to master learning materials, in understanding concepts students not only know but are also able to re-express concepts in language that is easier to understand and apply. Students who are able to master a concept that has been conveyed by the teacher can certainly re-represent the concept in various formats or what is called multiple representations. *Multiple Representations* are models that present material concepts with various forms of representation such as macroscopic, submicroscopic and symbolic. Macroscopic representation is a concrete picture where students observe phenomena that occur in everyday life. Submicroscopic representation is an abstract picture that explains macroscopic phenomena while symbolic representation is used to represent macroscopic phenomena using chemical equations, mathematical equations, graphs, reaction mechanisms, and analogies (Anugraheni, 2018).

Conceptual understanding is not only developed in learning, but must also be supported by test tools that reflect the ability to understand concepts, because tests or assessments are an integral part of classroom learning. Assessment functions as a measuring tool for competencies achieved from learning that can be used to improve the learning process and also to determine student achievement (Yuliany et al., 2022). Assessment instruments can be made in various representations, but in their application, the test instruments made still use a single representation to interpret a concept. The application of a single representation in making assessment instruments can result in the measurement process of students' mastery of concepts having limited interpretation (Aulia, 2014b). Assessment instruments that apply only one representation are less able to measure students' abilities, because each student has very diverse intelligence (Ellianawati et al., 2020).

Buffer solution material was chosen in the process of developing instruments based on multiple representations, because the buffer solution material contains many representations. Buffer solution material requires explanations in various

forms of representation that can visualize the buffer solution material so that students are expected to be able to observe the symptoms that occur and analyze and draw more comprehensive conclusions (Alighiri & Drastisianti, 2018). Through this test instrument, the level of students' representational abilities will be known starting from macroscopic, submicroscopic and symbolic representation abilities. Multiple representation abilities are very necessary for students in solving chemistry problems that contain various representations. Students' critical thinking abilities are very necessary to shape students' cognitive strengths (Agustia, 2024).

The results of interviews with subject teachers at SMAN 1 and SMAN 7 Pekanbaru obtained information that in the assessment of the chemistry learning process, teachers use multiple-choice and descriptive questions that are not yet based on *multiple representations* to measure students' conceptual understanding of the buffer solution material. The questions given by teachers on the buffer solution material are more representative of the macroscopic and symbolic. Meanwhile, the results of interviews with students at SMAN 1 and SMAN 7 Pekanbaru obtained information that students have difficulty answering questions on the buffer material in the form of calculations, symbols in the buffer solution material that are still foreign to students and understanding abstract questions. The development of assessment instruments based on *multiple representations* is very much needed to minimize conceptual errors in students and reduce the habit of students who are only able to work on questions with one representation (Anugraheni, 2018) . The implementation of evaluation activities in schools is usually only limited to measuring student learning outcomes, has never measured the level of student conceptual understanding, so that an evaluation instrument is developed to measure the level of student conceptual understanding (Hairunnissa, 2023).

Based on the results of research conducted by (Kurniasari, 2021) which analyzed the ability of *multiple representations* with conceptual understanding, it states that when students are given a *multiple representation approach* , students' multiple representation abilities will increase which will have an impact on increasing students' conceptual understanding. According to research by Ishak (2022), students' ability to work on submicroscopic representation questions is lower when compared to symbolic and macroscopic representations. This shows that students are better able to master knowledge related to real phenomena that occur in students' daily lives and students' ability to solve questions related to formulas and diagrams. This is in line with research (Habibah, 2023) the score for submicroscopic representation ability questions is lower when compared to macroscopic and symbolic representations. Students find it difficult to solve submicroscopic representation questions because students find it difficult to explain phenomena that cannot be observed because they consist of particular levels to be able to explain the state of particles.

Based on research by Alighiri & Drastisianti (2018) highest conceptual understanding profile at the macroscopic level. This can be seen from the questions adapted from practical questions about the effect of dilution on the pH

of buffer solutions. The macroscopic aspect is clearly visible to students in the form of changes in the color of the universal indicator used to test a buffer solution. Based on research conducted by Anugraheni (2018), students' ability to work on symbolic representation questions is higher than other representation abilities. This is in line with students' daily abilities in working on questions that tend to focus on symbolic representation. Based on several previous studies, it can be concluded that students' ability to solve questions with submicroscopic representations is lower than macroscopic and symbolic representations. Therefore, questions based on multiple representations are very important to reduce conceptual errors and reduce students' habits of working on questions with only one representation.

Based on research by Fadhilah & Widarti (2018) regarding the development of an assessment instrument based on multiple representation interconnections on buffer solution material , the results obtained were that assessment instruments based on multiple representation interconnections are feasible to use. Research (Saputra et al., 2024) related to the development of multi-representation test instruments on thermodynamics material, obtained results that the developed test instruments are valid and practical. Research (Ellianawati et al., 2020) examined multi-representation assessments based on 21st century skills on straight motion material, from the results of the study obtained multi-representation assessment instrument products developed are valid and reliable. In line with research by Aulia (2014b) which found that multi-representation-based test products based on the Introductory Solid State Physics course were valid, practical, and reliable. Based on the background that has been described, the author conducted a study entitled Development of Assessment Instruments to Measure Students' Conceptual Understanding of Buffer Solution Material.

This study aims to measure students' conceptual understanding of buffer solutions based on aspects of material, language, construction, and conceptual understanding. It seeks to identify the characteristics (validity, reliability, difficulty level, and discriminative power) of the instrument items developed and to understand user responses to the assessment instrument in terms of readability and time adequacy of the developed instrument items. Additionally, the study aims to determine the results of measuring students' conceptual understanding of buffer solutions.

2. Methodology

The research was conducted at the postgraduate program of the Faculty of Teacher Training and Education, University of Riau, Chemistry Education study program, SMA Negeri 1 Pekanbaru and SMA Negeri 7 Pekanbaru. The implementation period is the even semester of the 2024/2025 academic year. The population in this study were students in two state senior high schools in Pekanbaru, namely State Senior High School 1 Pekanbaru and State Senior High School 7 Pekanbaru who had studied buffer solutions. The characteristic trial involved 30 students. In the one-on-one trial, 3 students of grade XI were

involved. who had high, medium and low academic levels. The limited trial response involved 3 teachers and 18 students of grade XI. In the large-scale trial, the research sample consisted of students of grade XI of State Senior High School 1 Pekanbaru and State Senior High School 7 Pekanbaru. The sampling technique in this study was the *cluster sampling technique*.

The approach model used is a development model that refers to the type of research and development *Research and Development* (R&D). The development model used is the ADDIE model which consists of several steps, namely *Analysis* , *Design* , Development , Implementation *and* Evaluation.

a) The analysis stage is the initial stage in research and development using the ADDIE model. In the initial stage of problem analysis, the problem is analyzed by interviewing teachers and students of grade XI. The needs analysis aims to determine the assessment instruments needed by students and teachers in understanding chemical concepts. Analysis of the buffer solution material domain in the chemistry education curriculum based on the regulations of the Ministry of Education and Analysis of indicators of conceptual understanding of the opinions of experts regarding conceptual understanding listed in scientific articles.
b) The design stage is carried out to design the form of the assessment instrument to be developed. The instrument design determines the indicators. The indicators will be used as a reference in the process of developing the assessment instrument.
c) The development stage is the implementation of a previously designed product design, where the instrument is validated by three material experts using the Aiken formula to calculate content validity. The validated instrument will be tested on 30 students to determine characteristics such as construct validity, reliability, level of difficulty, and discriminatory power, with construct validity measured using Pearson Correlation and reliability with Cronbach Alpha (values > 0.6 are considered reliable). The level of difficulty of the questions indicates their quality, and a limited trial will be conducted on 18 students and 3 chemistry teachers to obtain responses regarding readability, time adequacy, and feasibility of the assessment instrument.
d) The implementation stage is a large-scale trial stage conducted on class XI IPA students in the even semester of the 2023/2024 academic year at SMA Negeri 1 Pekanbaru and SMA Negeri 7 Pekanbaru. The data from the measurement test results were analyzed to determine the students' conceptual understanding abilities. At the evaluation stage, qualitative data analysis and quantitative data analysis were carried out with the aim of product feasibility. Qualitative data analysis is used to process data in the form of input, criticism and suggestions from experts and field tests for further product revision. While quantitative analysis is obtained from respondents and student score results.

The evaluation stage is carried out qualitative data analysis and quantitative data analysis aimed at product feasibility. Qualitative data analysis is used to process

data in the form of input, criticism and suggestions from experts and field tests for further product revision. While quantitative analysis is obtained from respondents and student score results.

3. Results and Discussion

The data obtained in the study in accordance with the product development research method of the assessment instrument to measure conceptual understanding is a development model that refers to the type of educational research *and development* (R&D) by adapting the ADDIE development model, namely a development model consisting of five stages including analysis , design , development , implementation *and* evaluation . Results and discussion in each phase *The development of the* ADDIE model will be explained further.

Analysis Phase

Preliminary research was conducted with several analyses, namely problem analysis, needs analysis and material analysis. Based on the results of interviews with several teachers, information was obtained that in assessing learning outcomes, questions are usually adopted from various sources without analyzing indicators for achieving the learning objectives of the independent curriculum. Teachers have not implemented assessment instruments that can assess conceptual understanding . The results of the analysis of the level of thinking on daily assessment questions obtained information that questions with cognitive levels C1 = 20%, C2 = 20%, C3 = 60% and C4 = 20%. The buffer solution material assessment instruments used by teachers mostly still measure aspects of memorization (C1), understanding (C2), and short calculations (C3). Then students are not used to solving factual problems, the form of questions given by teachers is limited to the ability to understand and remember. From the analysis of students, it was found that students in grade XI MIPA SMA /MA or equivalent were on average 16 -17 years old. Jean Piaget in his theory of cognitive development stated that in adolescence, a person is in the formal operational thinking stage. This means that adolescents are able to think logically and draw conclusions from formal theoretical reasoning based on ratios and hypotheses. At this stage, adolescents can think flexibly, effectively and can handle complex problems. SMA students according to their level of cognitive development require assessment instruments that can measure conceptual understanding.

Analysis of the domain of buffer solution material in the chemistry education curriculum based on the regulations of the Ministry of Education and Culture. Analysis of the buffer solution material curriculum based on learning achievements and learning objectives. Based on the analysis of the development of learning indicators, the form of conceptual understanding questions that can be made is using cognitive levels C3-C5 with a focus on question indicators in accordance with the indicators of conceptual understanding shown by students. It has been compiled by BSNP (2006) showing that there are seven indicators, namely: restating a concept, classifying objects according to certain properties

according to their concepts, giving examples and non-examples of concepts, presenting concepts in various forms of mathematical representation, developing necessary or sufficient conditions of a concept, using, utilizing, and choosing certain procedures, applying concepts or algorithms for conceptual understanding, with buffer solution material in grade XI which can help students achieve meaningful learning.

Design Phase

At the planning stage there are aspects of the assessment instrument design that will be developed. The assessment instrument developed is a grid of conceptual understanding questions in the form of multiple-choice questions arranged based on learning achievements and learning objectives according to conceptual understanding indicators and based on multiple representations. The learning objectives of buffer solution material are derived from learning achievements for SMA buffer solution material according to the independent curriculum.

Development Phase

This stage is the stage of implementing the development of the product design that has been designed in the previous stage. Concept understanding assessment instrument compiled by researchers contains chemical material on buffer solutions for research subjects in grade XI of high school in Pekanbaru. The instrument product developed will be continued with validation by 3 expert validators. The validation carried out in this study is content validity. The validation results were analyzed using V aiken. If the value of the aiken V analysis results is smaller than the minimum value of the aiken V, then the question item is considered invalid. Question items that are declared valid are corrected based on the notes provided by the validator, then tested on high school students.

The validator gives an assessment score on the validation sheet using a Likert scale of 1-5 and the concept understanding indicator uses a Guttman scale of 0-1. Furthermore, the values given by the 3 validators are analyzed using the Aiken V formula to validate all question items. Quantitative data are obtained from calculating the average value of the validation sheet which includes 4 aspects of the criteria, namely: 1) Material aspect, 2) Language aspect, 3) Construction aspect, 4) Concept understanding aspect, as shown in table 1.

Based on Aiken validation data according to the material, language, and construction aspects, 25 questions experienced an increase in score from 70 in the first validation, which was then revised according to expert input for improvement so that in the second validation assessment the score increased to 96. The data obtained from the validation results through the validation sheet were analyzed descriptively by percentage. The results of this study are in line with the theory put forward by Ngalim Purwanto (2013) which explains that a test is said to have a high level of item validity if the test is able to provide results that are in accordance with what is to be measured. Valid items can be included in the

question instrument, while invalid items should be discarded or not reused in the next test.

Table 1. Validation Criteria Aspects

Validation Criteria Aspects
Material Aspect
a. The relationship between learning outcomes and learning objectives
b. Suitability of learning objectives with assessment instruments
c. Conformity between assessment instruments and Bloom's Taxonomy levels
d. Conformity between material and instrument
e. The content of the material asked is according to the level, type of school or class level.
Language Aspects
a. Presentation of cases and writing in assessment instruments
b. Presentation of language and writing in assessment instruments
Construction Aspects
a. Presentation of images, tables, graphs or similar on assessment instruments
b. Readability of assessment instruments
c. The suitability of the breadth of questions to the development of high school students
d. Suitability of the assessment instrument with the answers from the instrument
e. Completeness of assessment instruments (question grid, question items and answers)

The first stage Aiken's (V) value ranges from 0.736-0.840, included in the moderate category, but there are still some that need to be revised. Question number 16 has a low validation value compared to the validation values of other numbers. There is still writing that does not match the content, still uses inappropriate language, and the presentation of images is not clear enough so that it requires further revision. For some question numbers, it must be corrected based on expert revisions. Valid questions can be included in the question instrument and removed in the next trial test, while invalid questions should be discarded or not reused in the next test.

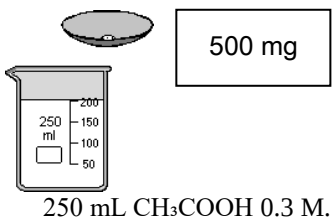
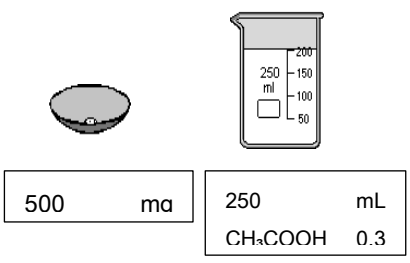
The second stage Aiken's (V) value ranges from 0.88-0.97, declared valid and meets the minimum validation requirements accepted. This value meets the minimum requirements for the accepted $V_{\text{table value}}$, which is accepted for 3 experts/validators with a scale category of 1-5, the instrument developed is declared valid because the Aiken V value is greater than 0.86. Question items with a validation index below 0.86 are not used (Ad'hiya Laksono, 2018). Based on research (Aulia, 2014a) the results of the validation of the question instrument obtained were very high, namely 0.91. Based on this value, it shows that the question instrument developed can measure students' multi-representation abilities in solid state physics courses. After validation, revisions/improvements are made to the questions and question grids according to suggestions or input from the validator and the revised results are provided to the validator to re-validate. The suggestions/input from the validators and follow-up/improvements from each validator can be seen in Table 2**.**

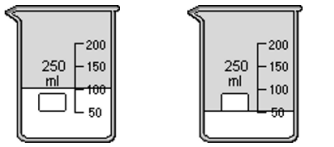
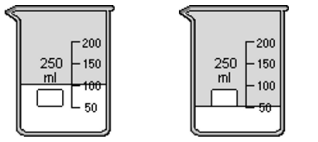
Table 2. Suggestions and Input for Improvement of Each Validator

Validators	**Suggestions/Feedback**	**Follow-up**
Validator 1	a. The presentation of the image in the question is not clear b. The position of the image is not very attractive	a. Replacing the image in the question b. Tidying up the position of the image in the question
Validator 2	a. Consistency in writing, there are "buffer" solutions and there are "support" solutions b. Some of the writing in the picture is not clear c. Consistency in mathematical writing . (dot) or , (comma), for example 10.0 or 10.0 grams	a. Improvements in writing buffer solutions b. Correction of writing in images c. Correction of mathematical writing to , (comma)
Validators 3	a. Complete the answer key and adjust the answers to the questions again b. Include image source c. Includes multiple levels of representation d. Includes distribution of conceptual understanding indicators	a. Complete the answer key and adjust the answers to the questions again b. Include image source c. Includes multiple levels of representation d. Includes distribution of conceptual understanding indicators

Improvements to the question items resulting from validation by the validator that have been obtained are used to improve the questions that are developed. Examples of questions that have been improved can be seen in table 3.

Table 3. Corrected Example Questions

Feedback and Suggestions	**Before Revision**	**After Revision**
Validators Question No. 13: c. Adjust image position	**Learning Objectives:** 11.19.5 calculate the pH of buffer solutions **Concept understanding indicators:** • Applying problem-solving concepts or algorithms **Question :** Pay attention to the experiment below!  250 mL CH_3COOH 0.3 M. If the two substances above are mixed, what is the pH of the mixture? (Ka acetic acid = 1.8 x 10-5 and log 5 = 0.95).	**Learning Objective:** 11.19.5 calculate the pH of buffer solutions. **Concept understanding indicators:** • Applying problem-solving concepts or algorithms **Question :**  If the two substances above are mixed, what is the pH of the mixture? (Ka acetic acid = 1.8 x 10-5 and log 5 = 0.95).

Feedback and Suggestions	Before Revision	After Revision
	A. 3,954 B. 4,046 C. 4,745 D. 5,046 E. 5,745	A. 3,954 B. 4,046 C. 4,745 D. 5,046 E. 5,745
Validators Question No. 20: d. Complete the answer key and adjust the answers to the questions again	**Learning Objectives:** 11.19.5 calculate the pH of buffer solutions **Concept understanding indicators:** • Applying problem-solving concepts or algorithms **Question :** A student will make a buffer solution with the following composition:  Larutan NH_3 0,2 M — larutan H_2SO_4 0,1 M Image source: (Wati, 2021) If the two solutions above are mixed, the pH of the mixture is... K b NH3 = 1 x 10^{-5} A. 9.2 B. 9.3 C. 9.5 D. 9.6 E. 9.7	**Learning Objectives:** 11.19.5 calculate the pH of buffer solutions **Concept understanding indicators:** • Applying problem-solving concepts or algorithms **Question :** A student will make a buffer solution with the following composition:  Larutan NH_3 0,2 M — larutan H_2SO_4 0,1 M Image source: (Wati, 2021) If the two solutions above are mixed, the pH of the mixture is... K b NH3 = 1 x 10^{-5} A. 9.2 B. 9.0 C. 6.0 D. 5.0 E. 4.8

After the instrument is revised in the first validation, the instrument is then continued with a second validation to obtain a valid instrument in terms of material/content. The validated assessment instrument will be tested on 30 students of SMA Negeri 1 Pekanbaru and SMA Negeri 7 Pekanbaru class XI IPA in the 2023/2024 academic year. Students will work on conceptual understanding questions that have been declared valid in terms of content/material, totaling 25 questions. The results of the initial trial were analyzed to determine the characteristics of the assessment instrument product to measure conceptual understanding. students , namely: construct validity, reliability, level of difficulty and differentiating power.

The construct validity in this study is by using *Pearson Correlation* . This test was conducted using *SPSS for Windows software* version 24.0. After the validation coefficient value of each question item was obtained, then the results were compared with the r value from the table at a significance level of 5% and a significance level of 1% with df = N-2. If r count> r table then the validity of the

significance used (Suharsimi Arikunto, 2010). Table 3 above shows the results obtained from 25 construct validity questions, there are 22 valid questions and 3 invalid questions. Invalid questions cannot be used in the study, so invalid data is deleted. The *Pearson correlation value* of each item is > 0.361 if n = 30, therefore the data that can be used is 22 questions. Valid questions can be included in the question instrument and removed from the next trial test, while invalid questions should be discarded or not used again in the next test.

The questions that were declared invalid were (1). Question number 5, namely the indicator explaining the types of buffer solutions; (2). Questions number 22 and 25 are in the learning indicator explaining the role of buffer solutions in everyday life. Three questions were declared invalid because the analysis of the questions by students was still lacking, based on the aspect of understanding the concept that was measured, students were asked to be able to analyze the problems contained in the questions so that they could prove the facts according to the answers chosen in solving the problems presented in the question instrument, here the answers chosen by students were less able to analyze in solving problems by proving the facts presented because there were still many students who answered incorrectly on the question items.

Reliability is measuring the extent to which the measuring instrument used for respondents' answers in the study can reach the answers to research problems objectively or the extent to which the truth of the research objectives can be answered by respondents correctly through or with the use of the tools or instruments used. In other words, reliability shows the consistency of a measuring instrument in measuring the same symptoms (Wahyuni, 2014). The results of the research reliability test show that the variables tested produce a *Cronbach alpha coefficient value* of 0.815, namely greater than 0.6, thus it can be concluded that the instrument is reliable and suitable for use. Based on research (Tafryda, 2022), the online instrument for measuring science communication skills based on multiple representations has a reliability coefficient value of 0.837 so that the developed question instrument is reliable.

According to Nana Sudjana (2011), the level of difficulty of the questions is seen from the ability or capability of students in answering a question. Analysis of the level of difficulty was carried out for 25 valid questions. The results of calculating the level of difficulty of the questions after validation. 4% of the questions were in the easy category, 84% of the questions were in the medium category and 12% of the questions were in the difficult category, with a comparison of 3 questions in the easy category, 21 questions in the medium category, and 1 question in the difficult category. Based on research (Tafryda, 2022), the results of the analysis of the level of difficulty of the questions obtained 80% of the questions in the medium category, 12% in the easy category, and 8% in the difficult category. From these results, the online instrument for measuring scientific communication skills based on multiple representations is classified as proportional as seen from the percentage of the number of questions in the medium category which is greater than the easy and difficult categories.

According to Bagiyono (2017) Analysis of the level of difficulty of questions means examining test questions in terms of their difficulty so that questions can be obtained which are easy, medium and difficult. The assumption used to obtain good quality questions, in addition to meeting validity and reliability, is the balance of the level of difficulty of the questions. Analysis of the level of difficulty is used to predict the measuring instrument itself (questions) and the ability of students to understand the material taught by the teacher. For example, one question is included in the easy category, then the prediction of this information is that most students answer the question correctly, meaning that most students have understood the material asked (Djanuarsih, 2012).

Based on research (Ellianawati et al., 2020), several factors that influence the results of the analysis of the level of difficulty of questions are:

a) Students are less able to grasp the meaning of the question
b) Students forget the formula that will be used
c) Students are not familiar with multi-representation questions

Discriminating power is the ability of questions to differentiate between high-ability students and low-ability students. The discriminating power value is expressed through the discriminating power index. The higher or greater the discriminating power index of the question, the greater the question can differentiate between the upper group and the lower group. The discriminating power of the 25 questions analyzed obtained 15 questions in the very good category (60%), 5 questions in the good category (20%), 3 questions in the sufficient category (12%), and 2 questions in the less good category (8%). Acceptable or good questions mean that the questions have the ability to differentiate between students who have understood the material and students who have not understood the material. Questions that are not accepted/cannot be used must be discarded because they cannot differentiate between students who understand the material and those who have not understood the material, possibly because the answer key is not correct, the competency being measured is not clear and the material being asked is too difficult. The criteria for good questions are if they have a discriminating power value > 0.30, while questions that have a discriminating power ≤ 0.30 need to be discarded or replaced. The results of the analysis of the discriminating power of conceptual understanding questions declared good and meets the requirements because it has a discriminatory power of ≥ 0.30 without revision (Ramadhan et al., 2019).

According to Suharsimi Arikunto (2010), a test is said to be good if it meets the following requirements:

1) Validity. An assessment instrument is said to be good if it has a construct validity value stated by a Pearson correlation > 0.36 for n = 30.
2) Reliability. The assessment instrument is said to be feasible or reliable if it has a Cronbach alpha value > 0.60.
3) Regarding meeting the difficulty level, it is not too difficult and not too easy. The higher the difficulty level, the easier the question and the lower the difficulty level, the more difficult the question.

4) The question has good discriminating power. The higher the discriminating power of the question, the better the question and the lower the discriminating power of the question, the worse the question.

The results of the characteristics of the conceptual understanding assessment instrument in the table above, it can be concluded that 25 valid assessment instrument items were analyzed, it turns out that 18 questions meet the requirements for use in limited scale tests and large scale tests, namely questions on numbers 1, 3, 4, 6, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 23 and 24. There are 7 questions that do not meet the requirements, namely questions on numbers 2, 5, 7, 13, 20, 22, and 25. This is because based on the results of the characteristics that have been carried out, where the questions obtain validation, reliability, difficulty level, and discrimination values that meet the requirements. Questions that do not meet the requirements are influenced by the different cognitive levels of students, so they have not been able to complete the conceptual understanding questions. Based on the researcher's considerations, the number of questions tested on students was 15, this was seen from the achievement of learning indicators in the material and the allocation of the test implementation time plan.

In the one-on-one test, the conceptual understanding ability assessment instrument consisting of 15 questions was tested on 3 grade XI students face-to-face (offline), the students selected were students with high, medium and low abilities. The one-on-one test aims to identify and reduce errors in the developed product. These three students provided responses according to the questions asked by the researcher in the interview after completing all the questions (Zulaiha, 2012). Some information obtained from the one-on-one test was that students with higher abilities worked on the questions faster than students with low and medium abilities. Test participants stated that the conceptual understanding ability questions they worked on required time to digest and analyze the purpose of the questions. If test participants can understand and sort out the known and asked parameters and the purpose of the questions, then students can find the answers to the questions.

After the test participants had completed all the questions, a direct interview was conducted between the researcher and the test participants. The results of the researcher's interview with the test participants obtained the following information: The message/information conveyed in the instrument is easy to understand and the sentences in the questions are effective and have a good structure/construction. The display of images on the instrument is very clear, attractive and related to the questions in the instrument. The grammar and spelling used in the questions are correct. Test participants stated that the Conceptual Understanding Ability Assessment Instrument is very suitable for measuring cognitive aspects and is very useful for training conceptual understanding abilities. Based on the results of the responses of the one-on-one test participants to the product, the assessment instrument was declared good to use with slight revisions for the next stage.

The assessment instrument that was declared to meet the requirements of validity, reliability, level of difficulty and discriminatory power, was continued with a limited test conducted on 18 grade XI students in the 2023/2024 academic year and 3 chemistry teachers who teach in grade XI at SMA Negeri 1 Pekanbaru and SMA Negeri 7 Pekanbaru to find out their responses to the assessment instrument that was developed. The results of the user response questionnaire were used to determine the readability of the questions, the adequacy of time and the feasibility of the instrument. The student response questionnaire to the concept understanding assessment instrument in the limited scale trial showed that the instrument had a question readability aspect consisting of 18 statement items in the very good category. Instrument users stated that the sentences in the instrument were effective and easy to understand, the sentence structure used was correct, the sentences used were simple, using standard language, the messages/information conveyed were easy to understand, the use of terms and symbols was consistent and the image display could be read. In the aspect of time adequacy consisting of 3 statement items, it showed a good category, this was related to the time given to work on the questions being stated as sufficient by the instrument users and there was no need to reduce the time in working on the instrument. The instrument user response questionnaire was also given to 3 chemistry teachers.

The teacher response questionnaire contains 18 statement items aimed at determining the feasibility of the conceptual understanding ability assessment questions, where the results obtained related to the feasibility of the instrument were 88.3%, indicating results in the very good category. The teacher stated that the material presented in the questions was in accordance with the description of the substance of the material contained in the learning achievements and learning objectives, the concepts presented in the questions did not give rise to many interpretations, the concepts were in accordance with the definitions that apply in their fields of science correctly. Instructions for working on the questions, question sentences, pictures/tables/question graphs were very clear. The questions developed did not depend on the next questions, the questions used language that was in accordance with Indonesian language rules and were communicative. The instrument had a stimulus and could stimulate test participants to solve problems and the conceptual understanding assessment instrument developed made a major contribution to students. Based on the results of the limited trial analysis, it can be concluded that the conceptual understanding assessment instrument can be accepted without improvement as a final product before being used in a large-scale test.

The implementation stage is the stage of product usage activities on a large scale. The implementation stage in this study consists of learning observation activities and ends with daily assessments using conceptual understanding assessment instruments on buffer solution material. A large-scale test is a test to measure students' conceptual understanding of buffer solution material which is carried out directly. The instrument used in this large-scale test is an assessment instrument that has been tested on students and the quality of the instrument has been known, both validity, reliability, level of difficulty and distinguishing power. The

assessment instrument as the final product of this development research functions as *an Assessment of Learning* , which is an assessment carried out after the learning process is complete. This assessment is intended to determine the achievement of learning outcomes after students have completed the learning process. The sample in this measurement test involved 25 students each at SMA Negeri 1 Pekanbaru and SMA Negeri 7 Pekanbaru from class XI IPA in the even semester of the 2023/2024 school year. The selection of samples was based on the criteria for high and medium schools obtained from the UTBK ranking in 2023. The results of the calculation of the average percentage of the Concept Understanding Indicator can be seen in Table 4.

Table 4. Average Concept Understanding Indicator

Conceptual Understanding Indicator	**Mark %**	**Category**
Restating a concept	70	High
Classifying objects according to certain properties (according to their concept)	74	High
Providing examples and non-examples of a concept	62	High
Presenting concepts in various forms of mathematical representation	67	High
Developing necessary or sufficient conditions of a concept	64	High
Using, utilizing, and selecting certain procedures or operations	62	High
Applying concepts or understanding concepts	58	Medium

Students who have conceptual understanding tend to be quicker in defining problems , formulating problem solving , determining and implementing problem solving strategies , conducting evaluations (Hidayah et al., 2017) . This is in line with Piaget's cognitive development stages which state that high school students aged 15-18 years are at the formal operational stage. At this stage, high school students can think abstractly, logically, deductively, inductively, and can draw conclusions from available information (Farah , 2020).

The reason for the lack of understanding of student concepts is because in the learning process at school, understanding concepts has not become the focus of the skills developed. Schools are still focused on achieving mastery of the material only. (Nugraha et al., 2017) . Science learning tends to memorize rather than develop thinking power so that students are weak in conveying ideas, weak in analyzing, and dependent on others. (Patonah, 2014) . There are still many students who find it difficult to apply the knowledge and concepts they know to solve problems, this indicates that students must practice a lot in applying the concepts or knowledge they have (Susilowati & Ramli, 2017) . In fact, understanding concepts is a thinking process that occurs in a person who aims to make rational decisions about something that can be believed to be true (Tawil & Liliasari, 2013).

The percentage of achievement results show that most students have a high and medium conceptual understanding. While no students have a very high conceptual understanding. This shows that very few students have a very good conceptual understanding. The indicator of questions with low achievement is because

students with low achievement are not used to working on questions that require conceptual understanding. Most students have difficulty remembering the material taught in school. The results of this measurement are expected to be an evaluation for teachers to improve classroom learning and train students with questions that require students' understanding skills.

The results of the large-scale test or further conceptual understanding measurement test to determine the differences in the conceptual understanding profiles of students in high, medium and low schools using the One Way Anova difference test. Before the measurement results are analyzed for difference tests, there are several prerequisite tests that must be met. Data must be normally distributed and homogeneously distributed. The final stage of this research is the evaluation stage. This evaluation stage is to determine the suitability of each step of the activity and the product that has been developed with its specifications (M. Hidayah, 2017) . The results of data analysis from 5 stages of development obtained that an assessment instrument for understanding concepts in the material of buffer solutions had been developed.

Research on the application of assessment instruments for students' conceptual understanding of buffer solution material with the help of multiple representation questions shows significant results in measuring student understanding. This study aims to evaluate the extent to which students understand the concept of buffer solutions through questions that use various forms of representation, such as text, images, graphs, and diagrams. The results of the study showed that in schools with a high category, students achieved an understanding percentage of 70.1%, which is included in the high category. This shows that most students are able to answer questions correctly and understand basic concepts well. Meanwhile, in schools with a medium category, the percentage of student understanding is 59.2%, which is included in the medium category. This shows that although students have a fairly good understanding, there are still some concepts that they do not fully understand. The use of multiple representations has been shown to help students see concepts from various perspectives, making it easier for them to integrate information and understand the material more deeply.

Overall, the application of assessment instruments with multiple representation questions in buffer solution material is effective in measuring and improving students' conceptual understanding. These results can be the basis for the development of more effective and inclusive educational strategies, ensuring that all students have the opportunity to understand the material well. By using various forms of representation, students can build a more comprehensive understanding of the concept of buffer solutions. This is in line with the theory of multiple representations which states that presenting information in various formats helps students process and integrate knowledge better.

This study supports the theory of conceptual understanding and multiple representations, showing that presenting information in multiple forms can improve students' conceptual understanding. Thus, the use of multiple

representations not only helps in assessment, but also in the learning process itself, making it an important approach in science education.

4. Conclusion

This study successfully developed a valid and reliable conceptual understanding assessment instrument for students. Through two stages of validation by three material experts, the developed instrument showed increased validity, with 22 out of 25 questions declared valid. This instrument also met the requirements for good reliability. The level of difficulty of the questions was well distributed between easy, medium, and difficult categories, and the discriminatory power of the questions showed satisfactory results. In addition, the study found significant differences in conceptual understanding between students from high and medium school categories. Thus, this instrument can be used effectively to measure students' conceptual understanding in learning.

Acknowledgement

The authors would like to thank the faculty of the Chemistry Education Program at Riau University for their support and guidance throughout this research. This study was conducted with the approval of the school authorities at SMA Negeri 1 Pekanbaru and SMA Negeri 7 Pekanbaru. Additionally, we appreciate the contributions of the participating students who made this research possible.

References

Agustia, Z., Yennita, Y., & Fakhruddin, F. (2024). Profile of Critical Thinking Ability of Class VIII State Middle School Students in Science Learning. *Journal of Education and Learning Research, 2*(1), 46-53.

Alighiri, D., & Drastisianti, A. (2018). *Multiple Representation Learning . 12* (2), 2192–2200.

Anugraheni, NS (2018). *Analysis of Students' Representation Ability Using Multi-Representation Based Test Instruments* . 258–262.

Aulia, L. R. (2014a). *Development of Multiple Representation Based Test Instruments in Introductory Solid State Physics Courses* . 1–7.

Aulia, L. R. (2014b). *Development of Multirepresentation-Based Test Instruments in the Introductory Solid State Physics Course Lusi* . 1–7.

Bella, S., Azhar, A., & Islami, N. (2023). Development of Think-Pair-Share (Tps) Model Based Learning Tools for Global Warming Materials. *Journal of Education and Learning Research, 1*(1), 20-29.

Djanuarsih, E. (2012). Validitas dan reliabilitas butir soal. *E-Jurnal Dinas Pendidikan Kota Surabaya, 1*(1), 1-12.

Ellianawati, E., Mufiatunnikmah, S., Setyaningsih, N. E., & Subali, B. (2020). Multi-Representation Assessment Based on 21st Century Skills on Straight Motion Material. *Physics Education Research Journal, 2* (1), 19. Https://Doi.Org/10.21580/Perj.2020.2.1.5038

Fadhilah, S. N., & Widarti, H. R. (2018). *Development of Assessment Instrument Based on Multiple Representation Interconnection on Buffer Solution Material . November* , 110–119.

Habibah. (2023). *Improving Student Learning Outcomes on Buffer Solution Topics Through Multi-Level Representation Learning . 02* (01), 1–10.

Hairunnissa, R. (2023). *Development of Two-Tier Multiple Choice Evaluation Instrument Using Personal Computer to Measure Students' Conceptual Understanding of Buffer Solution Material* . Https://Doi.Org/10.29303/Cep.V6i1.3372

Hidayah, M. (2017). Development of Psychomotor Assessment Instruments in Outdoor Biology Practicum for Senior High School Students. *Journal of Biology Education Research , 1* (2).

Hidayah, R., Salimi, M., & Susiani, T. S. (2017). Critical Thinking Skill: Concepts and Assessment Indicators. *Sebelas Maret University Scholar Park Journal , 1* (2).

Ishak, C. (2022). *Description of Students' Conceptual Understanding in Solving Electrolyte and Non-Electrolyte Solution Problems Based on Macroscopic, Submicroscopic and Symbolic Representations . 4* .

Ministry of Education and Culture. (2022). *Learning Outcomes in Early Childhood Education, Elementary Education Level, and Secondary Education Level in the Independent Curriculum* (Issue 021).

Kurniasari, L. Y. (2021). *Analysis of Multi-Representation Ability and Its Relation to Understanding of Physics Concepts Analysis . 16* (2), 142–150. Https://Doi.Org/10.29303/Jpm.V16i2.2404

Nugraha, A. J., Suyitno, H., & Susilaningsih, E. (2017). Analysis of Critical Thinking Ability Reviewed from Science Process Skills and Learning Motivation Through PBL Model. *Journal of Primary Education Uns , 6* (1).

Patonah, S. (2014). Elements of Objective Reasoning in Science Learning Through the Metacognitive Approach of Junior High School Students. *Journal of Indonesian Science Education, Universitas Pgri Semarang , 3* (2).

Permendikbud. 2015. *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 53 Tahun 2015, tentang Penilaian Hasil Belajar Pendidik Pada Pendidikan Dasar dan Pendidikan Menengah*. Kemendikbud. Jakarta.

Ramadhan, S., Mardapi, D., Prasetyo, Z. K., & Utomo, H. B. (2019). The development of an instrument to measure the higher order thinking skill in physics. *European Journal of Educational Research, 8*(3), 743-751.

Susilowati, S., & Ramli, M. (2017). Analysis of Critical Thinking Skills of State Islamic Senior High School Students in Magetan Regency. *Proceedings of the National Seminar on Science Education, Uns* .

Sudjana, N. (2011). Metode Statistika Edisi ke 6. *Tarsito.*

Sudjana, N. (2013). Dasar-dasar Proses Belajar Mengajar. *Bandung: Sinar Baru Algesindo.*

Sujadi. (2003). Metodologi Penelitian Pendidikan. Jakarta: Rineka Cipta.

Sugiyono. (2019). Metode Penelitian Pendidikan. *Bandung: Alfabeta*.

Sunyono. (2015). *Model Pembelajaran Multiple Representasi.* Media Akademi.

Tafryda, S. (2022). *Development of Online Instruments to Measure Students' Multiple-Representation-Based Science Communication Skills in Straight Line Motion Material* . *11* (2), 92–102.

Tawil, M., & Liliasari. (2013). *Complex Thinking and Its Implementation in Science Learning* . Makasar State University.

Saputra, T. A., Effendi, & Sinensis, A. R. (2024). *Development of Multi-Representation Test Instruments on Thermodynamics Material to Identify Students' Representation Abilities* . *16* (1), 1–328.

Wati, F. (2021). Analysis of Students' Chemical Representation Ability in Understanding Buffer Solution Material Using the Dac Instrument (Definition, Algorithmic, Conceptual). *Uin Walisongo* , 1–227.

Wulan. (2023). *Analysis of Chemical Understanding and Process Skills in Independent Curriculum Phase F Topic of Acid-Base for Senior High School*. *3* (November), 37–43.

Yuliany, N., Hidayat, R., Nursalam., Suharti., & Baharuddin. (2022). Development of a Measurement Tool for Conceptual Understanding Ability on the Topic of Three-Variable Linear Equation Systems. *Al Asma: Journal Of Islamic Education*, *4* (2), 95–107. Https://Doi.Org/10.24252/Asma.V4i2.32238

Zulaiha, S. (2012). *Code Switchings Used by the Students of Zainul Hasanain Islamic Boarding School Pajarakan-Probolinggo* (Doctoral dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).

How to cite this article:

Rehamita. S, Linda, R., & Futra, D. (2024). Development of Multiple Representation Based Assessment Instruments to Measure Students' Conceptual Understanding of Buffer Solution Material. *Journal of Educational Sciences, 8(4),* 671-689.