



A Validated Instrument to Assess Representational Interpretation Skills in Basic Chemistry Laws: A Development and Validation Study

Lisa Tania & Andrian Saputra*

Department of Chemical Education, Universitas Lampung, Indonesia

Abstract: The abstract and symbolic nature of chemistry presents significant learning challenges for students, particularly in the understanding of foundational concepts through a variety of representations. It's tougher to teach well, especially in Indonesia, because there aren't many trustworthy instruments to measure students' interpretative skills, which are very important for representational competence. The goal of this study was to fill this gap by creating and testing a reliable assessment tool called the Assessment of Representational Competence in Fundamental Chemical Laws (ARC-FCL), which was meant to measure the interpretative skills of Indonesian senior high school students. The Define, Design, and Develop stages of the 4D model were employed in the Research and Development (R&D) process of this study design. The Kozma-Russell framework and the Three-Dimensional Learning Assessment Protocol (3D-LAP) helped shape the instrument's development. The ARC-FCL test has two levels of multiple-choice and essay questions about the laws of Lavoisier, Dalton, and Gay-Lussac. Two chemistry education professors and one experienced instructor reviewed the content to ensure its validity. Researchers got real-world data for psychometric analysis from 30 11th-grade students at a senior high school in Lampung, Indonesia. The study looked at item validity using Pearson correlation, item discrimination, and internal consistency reliability using Cronbach's alpha. All three designed items had great validity, with correlation values over the key threshold of 0.361 and good discrimination indices between 0.50 and 0.75. The instrument was somewhat reliable, with a Cronbach's Alpha coefficient of 0.572. This is fine for a formative diagnostic tool at this early stage of development. Item analysis also showed that students had the most trouble with the question about the Law of Multiple Proportions. In conclusion, this study developed a valuable and valid instrument for identifying the specific challenges faced by students when attempting to comprehend sub-microscopic representations of fundamental chemical principles. As a formative assessment instrument that fits with the Kurikulum Merdeka, the ARC-FCL is quite useful for Indonesian teachers. It encourages a move from memorizing things by heart to comprehending them better and learning how to do science. Future work should be focused on adding more items to the bank to make it more reliable and cover more ground.

Keywords: representational competence, assessment instrument, interpretive skills, basic chemical laws, 4D model, instrument validation.

▪ INTRODUCTION

Chemistry is a basic science that opens the door to many other scientific and technical domains. However, learning chemistry is challenging for students worldwide, primarily because it is abstract and relies on a specialized symbolic language (Gilbert & Treagust, 2009). One of the most important ideas in modern chemistry teaching is that to really comprehend something, you need to be able to move between different ways of showing chemical events. Johnstone (1991) notably put these representations into three groups: the macroscopic (things that can be seen and touched), the sub-microscopic (the level of atoms, molecules, and ions), and the symbolic (equations, formulas, and other arithmetic expressions). Treagust, Chittleborough, and Mamiala (2003) say that students

learn chemistry meaningfully when they can easily move between and connect these three levels. Kozma (2003) suggests that students' misunderstandings and difficulties with learning the subject stem from their inability to work effectively in various representational domains. As a result, it has become very important in chemistry education research and practice to help students learn how to deal with these representations and test their skills.

Representational competence is the ability to use, understand, and create various many kinds of representations well (Kozma & Russell, 2002). This talent goes beyond just recognizing symbols; it requires a complex collection of cognitive skills that let a student think with and through representations. The Kozma-Russell framework (2002) gives us a basic cognitive paradigm for understanding this kind of expertise. It states that competence encompasses numerous important traits, such as the ability to recognize and utilize key elements of a representation, understand the rules and conventions that govern its use, and, most importantly, be able to transition smoothly between different types of representation. For instance, a skilled scientist can read a chemical equation (symbolic), picture the interactions between the molecules (sub-microscopic), and guess what will happen in a test tube (macroscopic). This paradigm has helped change the focus of chemistry education from just memorizing facts to learning useful scientific abilities that may be used in other areas (Chandrasegaran, Treagust, & Mocerino, 2007). Students sometimes only memorize algorithms and symbols when they don't have a strong representational competency, which means they don't really understand the basic principles of chemistry (Nakhleh, 1992).

The ability to understand is a key part of representational competence as a whole. Interpretation is the mental process of understanding a representation by explaining what it means, getting its main information, and connecting it to a scientific principle that is already known (diSessa, 2004). Before a learner can turn a chemical equation into a painting of particles or make their own drawing of a phenomenon, they must first be able to correctly understand the information that is conveyed in each individual drawing. For example, to comprehend the symbolic notation " $2\text{H}_2(\text{g}) + \text{O}_2(\text{g}) \rightarrow 2\text{H}_2\text{O}(\text{l})$," you need to know not only what the symbols mean for the elements, but also what the stoichiometric coefficients are, what the states of matter are, and how the reactants change into products. If you don't get this first step of interpretation right, you'll make mistakes in more complicated jobs like translating or solving problems (Schnotz, 2002). Because of this, the ability to appropriately interpret representations is very important and is a major obstacle to developing chemical understanding. This is why it is so important to test this skill in order to find out where learning hurdles are coming from.

Given the significance of testing representational competency, particularly its interpretative capacity, it is essential that the methods employed are meticulously designed tested. Educators need good assessment tools to find out what problems students are having, curriculum designers need them to make focused interventions, and researchers need them to test out new teaching methods (Pellegrino, Chudowsky, & Glaser, 2001). However, it is challenging to create high-quality items that truly assess a student's understanding, rather than just their ability to memorize information. These things need to focus on the exact mental processes that go into interpretive skills. The Three-Dimensional Learning Assessment Protocol (3D-LAP) by Laverty et al. (2016) is an example of a structured design protocol that has been used in science assessment recently. It gives a systematic framework. The 3D-LAP makes sure that test questions clearly include three dimensions: Science and Engineering Practices (such making and

utilizing models), Disciplinary Core Ideas (basic ideas), and Crosscutting Concepts (themes that show up in many subjects). The 3D-LAP lets you make stuff on purpose to see how well pupils use what they know in certain situations, which is in line with the goal of testing "competence." This study will use the 3D-LAP as a guide to create objects that focus on the practice of "interpreting representations" within the main notions of basic chemical laws.

Although everyone recognizes the importance of representational competence, few evaluation methods are specifically created for its different components, especially in high school chemistry classes focused on basic concepts. A lot of the tests that are out there are too broad and mix different talents like translation, interpretation, and invention, which makes it hard to find out what specific students need to work on (Stieff, Hegarty, & Deslongchamps, 2011). Also, not many tools focus on the most important and often hard topics, including basic chemical laws (such conservation of mass, definite proportions, and stoichiometry), where being able to understand symbolic and sub-microscopic representations is very important for doing well in chemistry in the future. This gap makes it hard for instructors and researchers to get a clear picture of how well pupils understand things during these early learning phases. The worldwide chemistry education community really needs a specific tool that focuses on interpretive skills in these areas.

This problem is especially important in the Indonesian school system. The national science curriculum focuses on scientific skills and thinking, however there aren't any high-quality, psychometrically validated tests that are appropriate for Indonesian students in terms of culture and language (Rahmawati & Koul, 2019). Most contemporary examinations still rely on old-fashioned tests that focus more on solving problems using algorithms than on understanding concepts and scientific skills (Effendy, 2017). It's not enough to just translate Western-made tools because differences in the way the curriculum is set up, how teachers teach, and how language works can make them less legitimate and reliable (Adlim, 2021). So, it is very important to make a custom evaluation tool from scratch using a methodical development method like the 4D model (Define, Design, Develop, Distribute) that Thiagarajan, Semmel, and Semmel (1974) suggested. This study intends to do just that by focusing on senior high school students in Lampung Province and serving as a model for similar efforts across Indonesia.

This study aims to address existing gaps by developing and assessing a robust evaluation instrument that concentrates on the interpretative aspect of representational competency in fundamental chemical laws. The item development will use the 3D-LAP technique (Lavery et al., 2016) and the whole process will follow the 4D development model (Thiagarajan et al., 1974), all based on the theoretical framework of Kozma and Russell (2002). The goal is to make a tool that Indonesian teachers and researchers can trust and use to figure out what students are having trouble with, help them learn chemistry better, and increase their comprehension of the most important chemistry ideas. This work adds to the global conversation on science assessments by looking at the situation in Indonesia and providing a useful tool for solving local educational problems. This study is based on the following research questions:

1. What are the characteristics of a assessment instrument designed to measure the interpretive dimension of representational competence in basic chemical laws for Indonesian senior high school students?

2. What are validity and reliability of the assessment instrument developed through the 4D model?

▪ **METHOD**

Participants

The study took place in Lampung Province, Indonesia, to ensure the instrument's relevance to the local cultural and educational context. The research targeted all tenth-grade students at SMAN 1 Katibung during the 2023/2024 school year, totaling 238 students. These students were chosen because the tenth-grade national curriculum in Indonesia introduces fundamental chemistry laws, making them ideal for assessing this topic. Their initial understanding and potential challenges in interpreting chemical representations made their responses suitable for validating the new instrument.

From this group, a sample of 68 students was selected for the main study using cluster random sampling. This method was chosen for its practicality in a school setting and its ability to prevent contamination between groups. Two intact tenth-grade classes were randomly selected, with one assigned as the experimental group ($n = 34$) and the other as the control group ($n = 34$). This kept the social and learning environments within each classroom undisturbed, supporting the validity of the results. Both groups had similar academic achievement and socioeconomic backgrounds, and neither had received detailed instruction on fundamental chemistry laws before the study began.

Research Design and Procedures

This study used a rigorous research and development process to systematically create and validate the assessment tool. It adopted the 4-D model, a well-known procedural framework by Thiagarajan, Semmel, and Semmel (1974), which includes four stages: Define, Design, Develop, and Disseminate. The entire process spanned eight months, allowing careful execution of each phase. A survey design was integrated into this framework to gather validation data from experts and empirical data from students.

The Define stage concentrated on establishing the conceptual and pedagogical basis of the instrument. This involved a thorough front-end analysis, reviewing existing literature on representational competence, common misconceptions about the laws of Lavoisier, Dalton, and Gay-Lussac, and the pedagogical challenges of teaching these topics. A learner analysis was conducted to understand the cognitive traits and prior knowledge of Indonesian tenth-grade students. This was followed by a task analysis to break down the skills needed to interpret macroscopic, sub-microscopic, and symbolic representations of basic chemical laws. Finally, a concept analysis identified and sequenced the core chemical principles, leading to clear, measurable learning objectives for the assessment.

In the Design stage, these objectives were translated into a practical plan for the assessment instrument. Decisions were made regarding the test format and structure, opting for a two-tiered multiple-choice format combined with essay questions. The first tier tests content knowledge, while the second, which requires written explanations, probes students' reasoning and interpretations—helpful for identifying misconceptions. The essay questions aim to capture a comprehensive understanding. During this phase, the initial item pool was developed, and scoring rubrics were drafted based on indicators of representational interpretation skills.

The Develop stage involved creating the instrument and conducting initial validation. Test items and related sub-microscopic diagrams were developed, with some

adapted from prior research to ensure validity (Gkitzia, Salta, & Tzougraki, 2019; Loh, Subramaniam, & Tan, 2014; Treagust, 1986). A draft of the instrument, including questions and scoring rubrics, was then reviewed by a panel of experts: two senior chemistry education lecturers from a university and one experienced high school chemistry teacher from Lampung. They carried out content and face validation and provided both qualitative feedback on clarity, accuracy, and relevance, as well as quantitative ratings for formal validity. After revisions based on feedback, a pilot test with students from a nearby school was conducted to assess item clarity, timing, and any logistical issues.

The Disseminate stage involved the formal implementation of the refined instrument. It was used as a pre-test for both experimental and control groups to establish baseline representational competence. After instruction—where the experimental group experienced a representation-focused teaching strategy, and the control group received traditional methods the same test was administered as a post-test. Data from these tests were analyzed to evaluate the instrument's psychometric properties and the intervention's effectiveness. While these results are beyond this paper's scope, they provided the data needed for the final validation.

Instruments

The main tool developed and validated in this study is the "Assessment of Representational Competence in Fundamental Chemical Laws" (ARC-FCL). It is specifically designed to evaluate high school students' skills in interpreting chemical representations related to core chemical principles. The assessment has two main parts: a multiple-choice section and an essay section. It focuses on three key chemistry laws: the Law of Conservation of Mass, the Law of Multiple Proportions, and the Law of Combining Volumes.

The first part includes multiple-choice questions using a two-tiered diagnostic format. Students must choose the correct answer from four options and then justify their choice in writing. This approach provides a detailed assessment, differentiating between guesses and true understanding. These questions mainly assess the skill of recognizing features and patterns in chemical representations. Visual prompts include sub-microscopic diagrams of chemical reactions, where each particle or molecule is clearly labeled with its chemical identity and associated mass or volume.

The second part is an essay where students analyze descriptive scenarios of chemical experiments. They interpret the quantitative data and explain how a particular fundamental law is demonstrated. This section aims to gauge higher-level thinking, especially skills in understanding the meaning behind various chemical representations and connecting macroscopic data with chemical principles. A detailed scoring rubric was created to evaluate responses, awarding points based on the inclusion of essential concepts and logical coherence. The items and styles of representation were adapted from influential chemical education research to ensure pedagogical relevance and support robust methodology (Gkitzia, Salta, & Tzougraki, 2019; Loh, Subramaniam, & Tan, 2014; Treagust, 1986).

Data Analysis

To ensure the quality and trustworthiness of the newly developed ARC-FCL instrument, a comprehensive data analysis plan was implemented, focusing on establishing its validity and reliability. The analysis was conducted in three primary

phases: content validity analysis, construct validity analysis, and internal consistency reliability analysis. All quantitative analyses were executed using specialized statistical software.

Content validity was established through expert judgment. Feedback from a panel comprising three experts (two in chemistry education and one experienced educator) was systematically analyzed. Construct validity, which pertains to the extent to which the instrument measures the intended underlying theoretical construct of "representational competence," was evaluated using data from the main study. An item analysis was conducted on student responses from both the pre-test and post-test. This included calculating the item difficulty index (p) and the item discrimination index (d) for each question. Items with poor discrimination or those considered too easy or too difficult were flagged for revision or removal. This empirical approach helps to verify that the items function as intended and effectively differentiate between students with varying levels of ability, thus providing evidence that the instrument accurately measures the targeted construct.

Finally, the reliability of the ARC-FCL was assessed through the analysis of internal consistency. Internal consistency indicates the extent to which the items within the test are correlated, reflecting whether they measure the same underlying construct. This was quantitatively evaluated by calculating Cronbach's alpha coefficient for the dataset obtained during the main administration of the test. According to established psychometric standards, a Cronbach's alpha value of 0.70 or higher is considered acceptable for a newly developed instrument, indicating satisfactory reliability for group-level assessments (Nunnally, 1978).

▪ **RESULT AND DISCUSSION**

This section outlines the systematic process of creating and validating an assessment tool centered on representational competence in basic chemical laws. The development adhered to the 4D model stages Define, Design, and Develop excluding the Disseminate phase. It combines the results from each stage, offering a clear narrative of how the instrument was developed, its theoretical foundation, and its psychometric properties, as demonstrated by a trial with high school students in Lampung Province, Indonesia.

Define Stage: Foundational Analysis

The initial phase of this research, the Define stage, focused on laying a strong foundation for developing the assessment instrument. This involved a multi-dimensional analysis to ensure the tool would be relevant, appropriate, and effective for its purpose. A comprehensive front-end analysis revealed a significant gap in chemistry education, especially in Indonesian secondary schools, where students often struggle to connect observable macroscopic phenomena, microscopic particle interactions, and symbolic chemical representations. This challenge hampers their understanding of fundamental principles like basic chemical laws. Therefore, creating an assessment that specifically measures students' ability to interpret chemical representations is essential. This tool aims not only to evaluate competence but also to give educators insights into specific difficulties, guiding teaching strategies to bridge these gaps. Developing such an instrument is vital for shifting focus from memorization to a deeper understanding of the particulate nature of matter.

After recognizing this need, a learning analysis was conducted to align the instrument with the Indonesian Kurikulum Merdeka for Phase F (typically Grade XI or XII). This curriculum emphasizes conceptual understanding and practical application of knowledge. The learning objectives centered on basic chemical laws, as this topic offers an ideal context for testing representational skills. Students are required to move seamlessly between different levels of representation to grasp concepts like conservation of mass and fixed or multiple proportions. Key topics included the Law of Conservation of Mass (Lavoisier), the Law of Definite Proportions (Proust), the Law of Multiple Proportions (Dalton), the Law of Combining Volumes (Gay-Lussac), and Avogadro's Hypothesis. These are fundamental to stoichiometry and often involve visual and symbolic representations that can be difficult for beginners to interpret accurately.

Alongside, a learner analysis was carried out to understand the target students senior high schoolers in Lampung, Indonesia. This considered their cognitive development, prior chemistry knowledge, and typical learning styles. Recognizing that students are transitioning from concrete to abstract thinking, yet still rely heavily on visuals, the assessment was designed to be visually engaging and scaffolded to support interpretation. It was also tailored to address common misconceptions, like confusing atoms and molecules or misinterpreting subscripts and coefficients in formulas. The analysis emphasized the importance of questions that ask not only "what" but also "why" and "how," encouraging students to articulate reasoning and reveal their mental models.

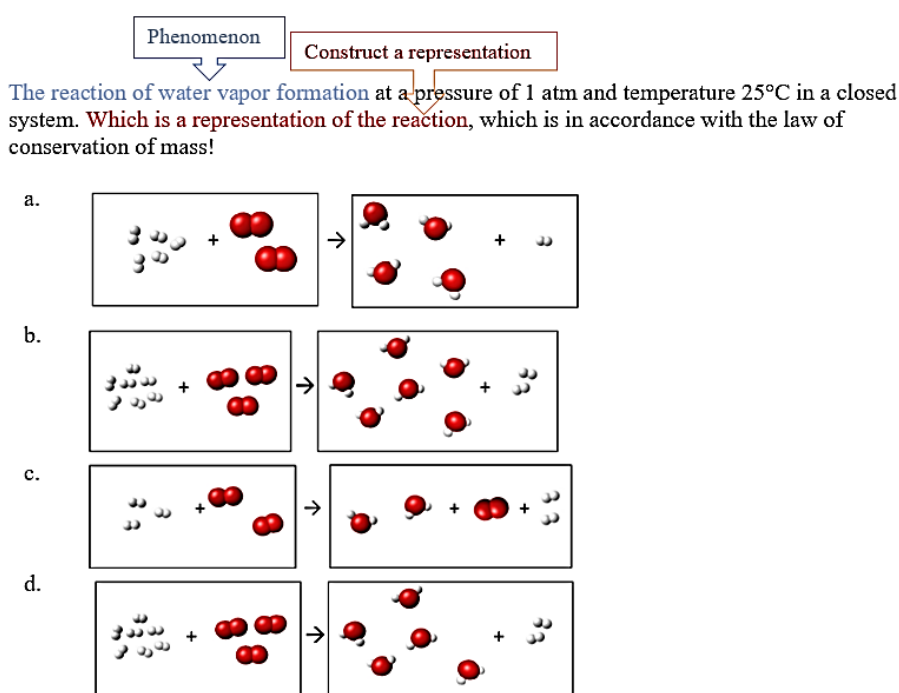
The concept analysis broke down the broad topic of basic chemical laws into specific concepts and skills. The main focus was on measuring representational competence specifically, the ability to interpret the meaning of chemical representations. This was operationalized using Kozma and Russell's (2005) framework, which highlights two key indicators for interpretation: recognizing features and patterns in representations, and explaining their meaning. For example, for Lavoisier's Law, students should identify that the number and type of atoms remain the same before and after a reaction. For Dalton's Law, they should recognize and explain the pattern of atoms combining in whole-number ratios across different compounds. This detailed mapping ensured each assessment item targeted a specific interpretative skill within a relevant chemical context. Finally, instructional objectives were defined to establish clear, measurable goals aligned with Kurikulum Merdeka. The main goal was for students to apply their understanding of chemical laws to analyze and interpret microscopic representations of reactions, moving beyond simple problem-solving.

Design Stage: Structuring the Assessment

In the Design phase, the conceptual foundation from the Define stage was transformed into a tangible assessment structure. This process involved systematically specifying test items, selecting media, designing formats, and creating an initial prototype. The design was heavily influenced by the Three-Dimensional Learning Assessment Protocol (3D-LAP), developed by Laverty et al. (2016), which aligns assessments with modern science education standards incorporating Science and Engineering Practices (SEPs), Crosscutting Concepts (CCCs), and Disciplinary Core Ideas (DCIs). This framework provided a strong blueprint for developing authentic assessment tasks that evaluate deep, integrated scientific understanding.

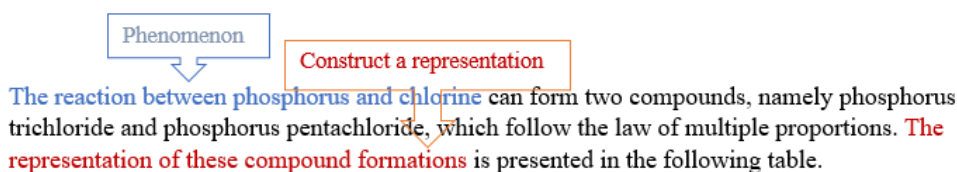
The first dimension, Science and Engineering Practices, was incorporated by requiring students to construct explanations and engage in argument from evidence. This approach goes beyond simple recall, encouraging students to articulate their

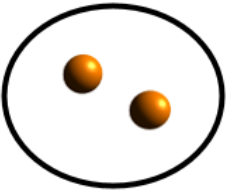
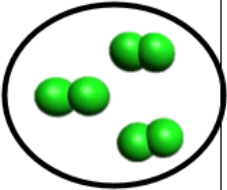
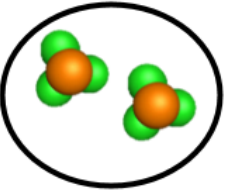
understanding. Each question was intentionally structured to prompt claim-making followed by reasoning. For example, Question 1, on the Law of Conservation of Mass, asks students to select the correct representation of a water formation reaction in a closed system and then justify their choice by connecting the visual evidence to the law. Question 2, on the Law of Multiple Proportions, asks students to determine the mass ratio of chlorine in phosphorus chlorides using particle diagrams, requiring them to explain their reasoning. Question 3, on Gay-Lussac's Law, asks for volume ratios and justifications based on gas representations. This consistent claim-and-reasoning format ensures assessment of students' ability to use evidence to support scientific explanations. The detail of Question 1, 2 and 3 can be seen as follow:

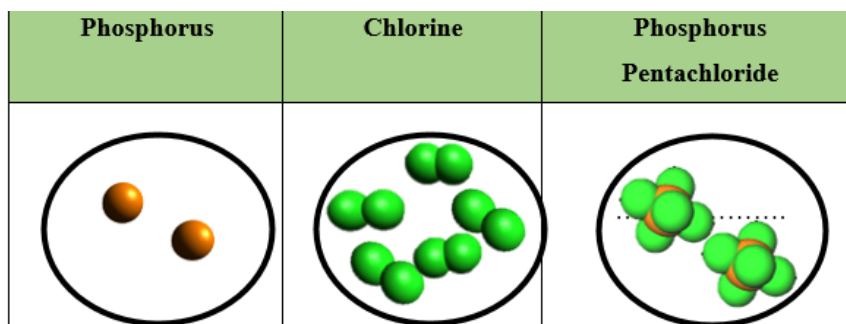


Give reasons why the image you chose is in accordance with the law of conservation of mass

Figure 1. Detail for question 1



Phosphorus	Chlorine	Phosphorus Trichloride
		



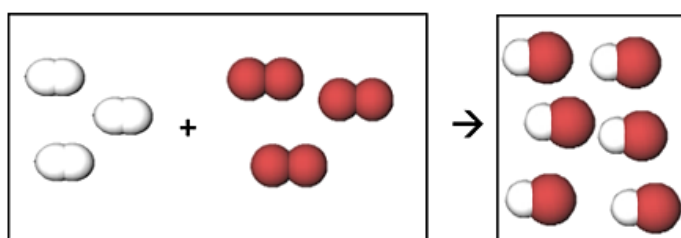
Based on the experimental data above, what is the ratio of the mass of the chlorine element in experiment 1 and experiment 2?

- 3: 1
- 3: 5
- 4: 6
- 6: 10

Figure 2. Details for question 2

Phenomenon

The following is a representation of the reaction of hydrogen gas with bromine, which produces hydrogen bromide gas, at the same temperature and pressure.



Based on this representation, what is the proper volume ratio of hydrogen gas, bromine gas and hydrogen bromide gas?

- 3: 3: 6
- 3: 4: 6
- 1: 1: 2
- 1: 1: 3

Provide Reasoning

Explain how you obtained these volume ratios

Make a claim

Figure 3. Details for question 3

The second dimension, Crosscutting Concepts, was incorporated through the principle that the structure of objects or systems determines their properties and functions. Here, the 'structure' refers to the sub-microscopic atom and molecule arrangements shown in diagrams, while the "function" relates to the chemical principles illustrated. Each question requires interpretation of visual structures to understand their chemical functions. For example, Question 1's reactant and product diagrams (showing atom arrangements) depict the principle of mass conservation. Question 2's structural differences in phosphorus chlorides (one with three chlorines, another with five) relate to their function under the Law of Multiple Proportions. Students decode these structural

features to grasp the underlying chemical laws. Question 3 uses simplified particle diagrams to help students deduce volume ratios, exemplifying Gay-Lussac's law. Linking visual structure to scientific function reinforces this crosscutting concept.

The third dimension, Disciplinary Core Ideas, underpins the content of the assessment. It focuses on core chemical laws, including Lavoisier's Law, Dalton's Law, and Gay-Lussac's Law, treated as principles explaining atomic rearrangements in reactions. Question 1 assesses that matter is conserved because atoms are conserved; students must recognize that the number of hydrogen and oxygen atoms remain unchanged through the reaction. Question 2 emphasizes that atoms combine in different ratios to form substances with varying properties, aligned with the Law of Multiple Proportions requiring students to count atoms and determine chlorine to phosphorus ratios from visual data. Question 3 explores how gas volumes relate in reactions, aligning with concepts of the mole and Avogadro's hypothesis. These questions grounded in fundamental DCIs ensure the assessment measures understanding of key chemical principles.

For media, interactive simulations from PhET Colorado and Javalab were chosen to depict dynamic, accurate molecular representations suitable for interpretation tasks. The format prioritized clarity and user-friendliness, presenting the phenomenon or diagram followed by a multiple-choice claim and a space for written reasoning. A confidence self-assessment was included to gather metacognitive data, asking students to rate their certainty on a five-point scale. The initial phase produced a prototype with three carefully designed items aligned with the 3D-LAP framework and the learning goals from the Define stage, ready for empirical testing in the Develop phase.

Develop Stage: Validation and Refinement

The Develop stage focused on empirically evaluating and refining the assessment instrument through pilot testing and psychometric analysis, confirming its validity, reliability, and item characteristics. Initially, expert validation was performed by a chemistry education evaluation specialist, ensuring content validity, clarity, and alignment with research goals. A subsequent empirical validation involved 30 Grade XI MIPA 3 students from SMAN 1 Katibung, with data analyzed using SPSS 27.0. Item validity analysis, using Pearson correlations, showed all items were valid, with coefficients well above the critical threshold of 0.361, indicating they effectively measure representational competence in basic chemical laws. The instrument's reliability, with a Cronbach's Alpha of 0.572, is moderate but acceptable for formative assessment, with potential for enhancement. Item analysis revealed moderate difficulty levels: Item 1 (mass conservation law) at 0.59, Item 2 (ratio interpretation) at 0.30, and Item 3 (Gay-Lussac's law) at 0.50. Discrimination indices ranged from 0.50 to 0.75, demonstrating good capacity to differentiate between students of varying achievement levels. Distractor analysis indicated all options functioned well, with plausible distractors chosen by more than 5% of respondents, reflecting common misconceptions. Overall, the instrument shows valid, discriminative, and adequately constructed items, suitable for assessing interpretative skills in chemistry, despite room for reliability improvement.

▪ CONCLUSION

This study successfully developed and validated the Assessment of Representational Competence in Fundamental Chemical Laws (ARC-FCL), an instrument designed to measure the interpretive skills of Indonesian senior high school

students regarding basic chemical laws. The results from the development process confirm the instrument's potential as a high-quality assessment tool. Psychometric analysis showed that all three test items have strong validity, reflected in their high correlation with the total test score ($r > 0.361$) and excellent discrimination indices, demonstrating their ability to distinguish between high- and low-achieving students. Additionally, the distractors for each item were found to be functional, aligning with common student misconceptions. Although the instrument's reliability was moderate (Cronbach's Alpha = 0.572), this is acceptable for an initial-stage diagnostic tool. The main contribution of this research is providing a specialized tool that fills a documented gap in chemistry education. By focusing on interpretation, the instrument helps educators diagnose students' difficulties in connecting sub-microscopic representations with abstract chemical principles, a key step in fostering deep understanding and moving beyond rote memorization.

The practical implications for education are significant. In Indonesia, the ARC-FCL can serve as a valuable formative assessment to identify students' specific weaknesses in interpreting particulate diagrams related to the laws of Lavoisier, Dalton, and Gay-Lussac. The finding that students struggled most with the Law of Multiple Proportions offers direct feedback for teachers, indicating the need for targeted instructional strategies and support. This tool can promote a shift from algorithmic problem-solving to nurturing authentic scientific skills, aligning with modern frameworks like Kurikulum Merdeka. However, the study has limitations. The instrument's moderate reliability suggests that future refinement is needed; adding more high-quality items could enhance its internal consistency. Also, validation was conducted on a limited sample from a single school in Lampung Province, which restricts the applicability of results to the broader Indonesian student population. Moreover, with only three items, the current instrument provides a focused but narrow assessment. Future efforts should aim to expand the item bank across different chemical laws and contexts to develop a more comprehensive diagnostic tool.

▪ REFERENCES

- Adadan, E. (2013). A rubric for assessing high school students' conceptual understanding of chemical bonding. *Educational Assessment*, 18(1), 1-27.
- Adlim, A., & Saminan, S. (2021). The need for assessment instruments for chemical literacy in the context of socio-scientific issues of the COVID-19 pandemic. *Jurnal Pendidikan IPA Indonesia*, 10(1), 12-23. <https://doi.org/10.15294/jpii.v10i1.26121>
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198.
- Briggs, D. C., & Alonzo, A. C. (2012). The psychometric modeling of learning progressions. *Journal of Educational Measurement*, 49(4), 427-447.
- 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. <https://doi.org/10.1039/B7RP90006F>
- Cooper, M. M., Stieff, M., & Dori, Y. J. (2016). The language of chemistry. In *The Wiley handbook of science teaching and learning* (pp. 371-396). John Wiley & Sons.
- Devetak, I., & Glazar, S. A. (2014). The influence of animations on students' understanding of chemical concepts at the submicroscopic level. *Chemistry Education Research and Practice*, 15(4), 603-614.

- diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, 22(3), 293–331. https://doi.org/10.1207/s1532690xci2203_2
- Effendy. (2017). The paradoxical situation of chemistry education in Indonesia and the need for a new direction. *Acta Chimica Asiana*, 1(1), 27–36. <https://doi.org/10.24198/aca.v1i1.13745>
- Gilbert, J. K., & Treagust, D. F. (Eds.). (2009). *Multiple representations in chemical education*. Springer. <https://doi.org/10.1007/978-1-4020-8872-8>
- Gkitzia, V., Salta, K., & Tzougraki, C. (2019). Development and application of a diagnostic instrument to assess Greek students' understanding of the basic concepts of the particulate nature of matter. *Chemistry Education Research and Practice*, 20(2), 433–447. <https://doi.org/10.1039/C8RP00267A>
- Gkitzia, V., Salta, K., & Tzougraki, C. (2019). Student understanding of the structure of matter in the three states: A study of the effect of an intervention based on a historical and a simulation approach. *Chemistry Education Research and Practice*, 20(4), 804–820.
- Hadenfeldt, J. C., Bernholt, S., Liu, X., & Parchmann, I. (2013). Using a Rasch-based learning progression for matter to facilitate the interpretation of a large-scale assessment in chemistry. *Journal of Chemical Education*, 90(6), 700–708.
- Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, 7(2), 75–83. <https://doi.org/10.1111/j.1365-2729.1991.tb00230.x>
- Keig, P. F., & Rubba, P. A. (1993). Translation of representations of the structure of matter and its relationship to reasoning, gender, and prior knowledge. *Journal of Research in Science Teaching*, 30(8), 883–901.
- Kozma, R. B. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, 13(2), 205–226. [https://doi.org/10.1016/S0959-4752\(02\)00021-X](https://doi.org/10.1016/S0959-4752(02)00021-X)
- Kozma, R., & Russell, J. (2002). Students becoming chemists: Developing representational competence. In J. K. Gilbert, M. de Jong, S. Justi, D. F. Treagust, & J. H. Van Driel (Eds.), *Chemical education: Towards research-based practice* (pp. 121–145). Kluwer Academic Publishers.
- Kozma, R. B., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 121–145). Springer.
- Laverty, J. T., Underwood, S. M., Matz, R. L., Posey, L. A., Carmel, J. H., Caballero, M. D., Fata-Hartley, C. L., Ebert-May, D., & Cooper, M. M. (2016). Characterizing college science assessments: The Three-Dimensional Learning Assessment Protocol. *PLoS ONE*, 11(9), e0162333. <https://doi.org/10.1371/journal.pone.0162333>
- Laverty, J. T., Underwood, S. M., Matz, R. L., Posey, L. A., Carmel, J. H., Caballero, M. D., Fata-Hartley, C. L., Ebert-May, D., & Couch, B. A. (2016). A computer-based assessment for measuring students' foundational knowledge of the chemistry of biology. *CBE—Life Sciences Education*, 15(2), ar17.
- Loh, A. S. L., Subramaniam, R., & Tan, K. C. D. (2014). Exploring students' understanding of the states of matter and solution chemistry using a three-tier diagnostic instrument. *International Journal of Science Education*, 36(15), 2475–2499. <https://doi.org/10.1080/09500693.2014.920551>

- Loh, Y. L., Subramaniam, R., & Tan, S. K. (2014). Using a three-tier diagnostic instrument to assess secondary school students' understanding of some chemical concepts. *International Journal of Science Education*, 36(5), 783-806.
- Nakhleh, M. B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education*, 69(3), 191. <https://doi.org/10.1021/ed069p191>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. National Academy Press.
- Rahmawati, Y., & Koul, R. (2019). The challenges of implementing a new science curriculum in Indonesian secondary schools. In D. F. Treagust, R. Koul, & W. F. Lin (Eds.), *Science Teacher Education in the Asia-Pacific* (pp. 81-94). Routledge.
- Schnotz, W. (2002). Towards an integrated view of learning from text and pictures. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 385-416). Lawrence Erlbaum Associates Publishers.
- Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction*, 29(1), 123-145. <https://doi.org/10.1080/07370008.2010.507318>
- Sun, D., & Furtak, E. M. (2016). A systematic review of the literature on the formative assessment of scientific argumentation. *Educational Research Review*, 18, 50-67.
- Taber, K. S. (2013). Three levels of chemistry educational research. *Chemistry Education Research and Practice*, 14(2), 151-155.
- Tang, K. S. (2016). The interplay of representations and patterns of interaction in a science classroom. *Science Education*, 100(4), 611-641.
- Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Indiana University.
- Treagust, D. F. (1986). Evaluating students' misconceptions by means of diagnostic multiple choice questions. *Research in Science Education*, 16(1), 199-207. <https://doi.org/10.1007/BF02356835>
- Treagust, D. F., Chittleborough, G. D., & Mamiala, T. L. (2003). The role of submicroscopic and symbolic representations in chemical explanations. *International Journal of Science Education*, 25(11), 1353-1368. <https://doi.org/10.1080/0950069032000075426>
- Tytler, R., Prain, V., Hubber, P., & Waldrup, B. (Eds.). (2013). *Constructing representations to learn in science*. Sense Publishers.
- Wu, H. K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education*, 88(3), 465-492.