



DEVELOPMENT AND VALIDATION OF THE VISUAL REPRESENTATIONS TEST IN CHEMISTRY FOR PRE-MEDICAL STUDENTS

**Jonathan M. Barcelo,
Precious Lady Gine N. Araneta**

Abstract. *The ability of pre-medical students to use visual representations in chemistry is crucial in understanding complex and abstract chemistry concepts.*

To measure students' knowledge and understanding of visual representations in chemistry, this study developed and validated the Visual Representations Test in Chemistry (VRTC). A hypothesized three-level progression of this construct was drafted and validated using Rasch analysis.

The first version of the VRTC comprised 26 multiple-choice items covering particle models, ball-and-stick models, skeletal structures, electrostatic potential maps, Fischer projections, and Newman projections. The test was administered to 827 pre-medical students across four institutions in the Philippines. Dichotomous

Rasch analysis provided evidence of reliability, unidimensionality, and item fit. Item-content validity indices indicated fair-to-excellent values, and the item hierarchy confirmed the hypothesized progression of knowledge about visual representations in chemistry. Analysis of variance of item difficulties revealed significant differences in mean item difficulty across the three hypothesized levels of knowledge, $F(2,23) = 17.273, p < .001$, supporting construct validity. However, Rasch analysis revealed student ability gaps, items with overlapping difficulty levels, and differential item functioning based on year level. Overall, the VRTC exhibited adequate psychometric properties for assessing pre-medical students' knowledge and understanding of visual representations in chemistry.

Keywords: *chemistry concepts, item content validity, pre-medical students, Rasch analysis, visual representations*

**Jonathan M. Barcelo,
Precious Lady Gine N. Araneta**
Saint Louis University, Philippines

Introduction

Learning chemistry meaningfully requires students to make sense of visual representations because chemical ideas are often presented through observable phenomena, particle-level models, and symbolic forms. Visual representations provide epistemic access to chemical phenomena that often require engagement with visual information and cannot be directly observed (Chi et al., 2018; Stull et al., 2016). In chemistry, visual representations such as equations, diagrams, molecular models, and other chemical representations are used to convey information, highlight essential features of concepts, and highlight the specific concepts, skills, and representational features that students are expected to learn (Schnotz, 2014; Stull et al., 2016; Talanquer, 2022). These representations support students in making sense of abstract chemical ideas, linking symbolic expressions with molecular-level processes, and building the analytical skills needed in science and health-related fields. Visual representations are commonly employed during problem solving (Cook et al., 2008; Höst et al., 2012; Rau, 2015) and scientific reasoning (Talanquer, 2022). Previous studies have also highlighted their importance in understanding molecular phenomena (Rodemer et al., 2021; Tsapalis et al., 2021), improving problem-solving abilities and content knowledge (Rau, 2015), and sustaining students' attention and motivation (Cook, 2006).

Chemistry teaching relies on multiple visual representations to convey the nature, properties, and transformations of chemical substances (Talanquer, 2022). However, each visual representation emphasizes a particular dimension of the concept being taught (Rau, 2017a). Thus, multiple visual representations in chemistry are used to present complementary information, develop mental models, and support chemical reasoning (Rau, 2016; Stull et al., 2016; Talanquer, 2022). Molecular models and representations, for instance, are used to visualize atomic and molecular structures (Zarkadis et al., 2017) and to explain structure-activity relationships (Carvalho et al., 2005). Other studies have also combined molecular structure and spectroscopic



analyses using molecular models (Fang et al., 2011; Wang, 2012). However, the use of multiple representations tends to benefit expert chemistry learners more because they can create meaningful clusters, whereas novice learners often focus on surface features and struggle to organize chemical information efficiently (Kozma et al., 2000; Rau, 2016). Therefore, most interventions aimed at improving learners' use of visual representations have focused on promoting multiple-representation competence (Höst et al., 2012; Mason et al., 2019; Rau, 2015; Rau et al., 2015).

Knowledge and understanding of visual representations in chemistry are rarely required in professional health sciences practice. However, effective learning of chemistry concepts requires knowledge of visual representations in chemistry (Rodemer et al., 2021). Since several chemistry concepts are required to understand health concepts and apply them to professional premedical courses (Armstrong & Poë, 2020; Barcelo & Ferido, 2025), instructional interventions that strengthen pre-medical students' competence in using visual representations in chemistry are needed to support the transfer of chemical knowledge to health-related concepts. However, there is a need for a validated assessment tool for measuring the target construct among pre-medical students to monitor changes as they progress to professional courses.

Visual Representation Competence in Pre-Medical Chemistry

In pre-medical programs, chemistry concepts serve as conceptual scaffolds for understanding the chemical basis of nutrition, disease processes, diagnostics, and pharmacology (Armstrong & Poë, 2020; Shulman, 2013; Barcelo & Ferido, 2023). Visual representations in chemistry provide epistemic access to these concepts by helping students make sense of molecular structures, models, and diagrams. They also support the learning of structure-property relationships, which encourages students to recognize emergent properties of substances, such as acidity, and emergent processes, such as diffusion (Luisi, 2002; Talanquer, 2007). Students who understand emergent properties may develop more sophisticated chemical thinking (Sevian & Talanquer, 2014), which is needed in decision-making. Since chemistry concepts are common in health contexts, a molecular understanding is needed to connect them to health decisions. Thus, the ability to interpret and use visual representations in chemistry is an important competency for pre-medical students.

Students need knowledge of representations (Rau, 2017b) if they are expected to use visual representations in chemistry to learn complex chemistry concepts. Ainsworth (2006) also noted that the complexity of a representation depends on the student's ability to understand its features and to represent an abstract concept. Visual representations may elicit high cognitive load because they require visual and spatial literacy (Prigozhin et al., 2014; Schönborn & Anderson, 2010). For novice chemistry learners, this load may be higher because their interpretations are often based on surface features or fragmented concepts (Cook, 2006). Hence, it is important to assess pre-medical students' knowledge of visual representations in chemistry.

However, studies have shown that high school and university students struggle to relate concepts embedded in visual representations to actual chemistry concepts (Cooper et al., 2013; Sim & Daniel, 2014; Stieff, 2011). Many students also rely on surface features and use heuristic reasoning when solving problems involving visual representations (Weinrich & Talanquer, 2016), and pre-medical students may experience similar difficulties. Learning an unfamiliar chemistry concept through an unfamiliar visual representation may further hinder the development of chemical knowledge (Rau, 2017a). Since biochemistry, the most relevant pre-requisite chemistry course in pre-medical programs, builds on prior knowledge from basic organic chemistry courses (Shulman, 2013), students need sufficient knowledge of visual representations before they can use them to learn chemistry concepts (Rau, 2017b). Therefore, assessing pre-medical students' knowledge and understanding of visual representations in chemistry is necessary to design learning activities that apply them in health-related or medical contexts.

Measurement Frameworks for Validating Visual Representation Assessments in Chemistry

Although only a limited number of assessment instruments have been developed to examine students' use of visual representations in chemistry, these tools have commonly been evaluated using classical test theory or item response theory. Classical test theory relies on raw test scores, which do not constitute linear measures of the underlying construct (Davis & Boone, 2021; Planinic et al., 2019). The Rasch model, which is anchored in item response theory, transforms raw scores into linear measures called log-odds units, or logits (Liu & Boone, 2023). Unlike other item response theory models, Rasch analysis predicts the probability of a response based on student ability and item difficulty (Tesio et al., 2023a). It also provides empirical evidence that an instrument measures the intended construct, while the item hierarchy can be compared with the predicted hierarchy based on a theoretical

framework (Boone, 2016; Boone & Noltemeyer, 2017; Zoechling et al., 2022). In the development of chemistry assessment instruments, the Rasch model can identify items that exhibit misfit and provide evidence to guide their revision or removal in subsequent iterations. Adequate item fit is therefore necessary to ensure that the instrument measures chemistry-specific abilities validly.

Classical test theory has been widely used to develop tools that assess knowledge of chemical representations. Sim and Daniel (2014) developed the Test of Representational Competence (TRC), which measures representational competence, and the Test of Chemical Representations (TCR), which measures understanding of chemical representations. Both tests involve particle models and classify students' scores as low, medium, or high based on quartiles of total scores. Al-Balushi and Al-Hajri (2014) developed the Organic Chemistry Visualization Test (OCVT), which measured visualization ability and content knowledge involving skeletal structures, molecular structures, and ball-and-stick models. The tool also assessed structure-property relationships. In this assessment, raw scores were used to measure visualization skill and organic chemistry content knowledge, although no categorical ability classifications were provided. Luxford and Bretz (2014) developed the Bonding Representations Inventory (BRI), a quantitative assessment designed to identify students' misconceptions about chemical bonding and its representations using Lewis structures. In this tool, the total score represents students' conceptual understanding of chemical bonding.

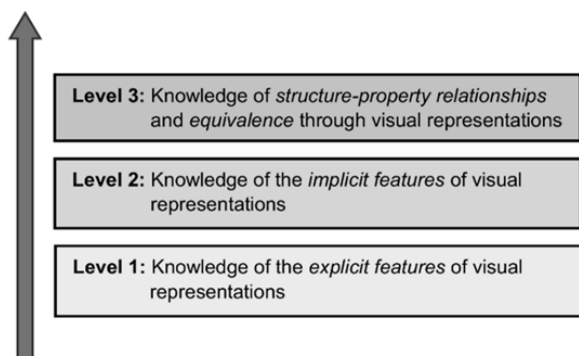
Other studies used the Rasch model to validate instruments that measure abilities involving chemical representations. Taskin et al. (2015) developed the Chemical Representation Inventory (CRI), which measures knowledge of chemical representations, including ball-and-stick models, skeletal structures, wedge-and-hash structures, and particle models. The CRI was validated using the partial credit Rasch model. Nickel et al. (2025) also developed the Chemical Representation Inventory (CRI), an assessment tool anchored on representational competence, or the ability to understand, use, and generate chemical representations for learning, problem-solving, and communication in chemistry. Specifically, the instrument focused on three lower-level representational skills: interpretation, translation, and construction. The CRI was evaluated using the dichotomous Rasch model for content knowledge items and the partial credit model for representational competence items.

The review of previous assessment tools that measure students' competencies in visual representations in chemistry highlights the importance of this topic in chemistry education. Despite this importance, none of these studies has used a progression-based model to classify students' knowledge and understanding of visual representations in chemistry or developed theoretically grounded items to objectively measure this construct. Rau (2017b) noted that conceptual knowledge of representations and representational competence are intertwined. Talanquer (2022) also emphasized the need to analyse the implicit and explicit information embedded in chemical representations. These insights suggest that chemistry learners should recognize not only the features of visual representations, but also the information they encode. To date, limited attention has been given to measuring the target construct among pre-medical students, necessitating the development of a new assessment tool.

Theoretical Framework

The hypothesized progression of knowledge and understanding of visual representations in chemistry was anchored on the Cognitive Theory of Multimedia Learning (Mayer, 2014) and Cognitive Load Theory (Sweller et al., 2011). As shown in Figure 1, increasing levels were hypothesized to reflect transitions from selection, or recognition of explicit features, to organization, or interpretation of implicit features, and finally to integration, which involves structure-property reasoning and representational equivalence. This framework is also supported by representational competence literature in chemistry (Ainsworth, 2006; Kozma et al., 2000), which emphasizes the ability to interpret, translate, and integrate multiple representations. These levels were grounded on the expected competencies of pre-medical students who have completed prerequisite chemistry courses such as general chemistry and general organic chemistry. The hypothesized levels do not indicate the visual complexity of a representation, but distinguish the type of cognitive processing required to answer an item correctly. These levels served as the basis for drafting the item types in the initial version of the Visual Representations Test in Chemistry (VRTC). The first version of VRTC is shown in [Supplementary Material A](#).



Figure 1*Hypothesized Progression of the Target Construct*

At Level 1, students should be able to identify, retrieve, or describe information that is explicitly displayed in a single visual representation. The required information is directly available from the model, structure, or diagram and can be accessed without coordinating multiple representations or making substantial chemical inferences. In the VRTC (please refer to Appendix B), this level includes tasks such as identifying the feature shown by an atomic representation (Item 1), counting carbon atoms in a skeletal structure (Item 7), identifying nonbonding electrons (Item 8), identifying atoms and functional groups in ball-and-stick models (Items 17 and 19), and identifying electron-rich regions in a representation (Item 24). These items reflect selection of explicit representational features because students must recognize what the visual representation presents.

At Level 2, students should link two or more visible features in a representation to infer information that is not directly labelled or explicitly shown. This level requires students to organize information such as atomic identity, bonding, geometry, charge distribution, molecular arrangement, or intermolecular interactions inside a coherent chemical interpretation. In the VRTC, Level 2 tasks include representing iodine sublimation (Item 2), describing water molecules during boiling (Item 3), predicting hydration of sodium ions (Item 4), inferring bond angles and hybridization (Items 5 and 6), interpreting three-dimensional molecular orientation from wedge-dash notation (Item 9), and interpreting electron-density information from electrostatic-potential maps (Items 22 and 23). These items require students to connect multiple explicit features and construct an interpretation of molecular structure, intermolecular behaviour, or chemical properties.

At Level 3, students must coordinate information across representations or integrate representational information with chemistry knowledge to make a structural or property-based judgment. These items require students to recognize equivalence or complementarity between different representations, mentally manipulate three-dimensional molecular arrangements, or use molecular structure to predict chemical characteristics. In the VRTC, Level 3 items include identifying a functional isomer (Item 11), comparing two molecular structures in terms of chemical properties (Items 12 and 16), identifying equivalent structures and products of bond rotation (Items 14 and 15), identifying a nucleophilic site using two representations of the same molecule (Item 21), and translating between Newman and Fischer projections (Items 25 and 26). These tasks represent integration because students must combine representational knowledge with conceptual knowledge of bonding, stereochemistry, molecular structure, reactivity, polarity, acidity, or intermolecular interactions. The proposed progression is therefore expected to move from identifying what is visibly represented to inferring what the representation implies to coordinating multiple representations and chemical concepts to explain or predict molecular properties and behaviour.

Research Problem

The scope and depth of chemistry concepts in pre-medical programs are shaped by the health contexts in which these concepts are applied. Since many chemistry concepts require access to molecular phenomena, it is important to assess students' knowledge and understanding of visual representations in chemistry, particularly when these representations are used to explain health-related concepts. Hence, there is a need for an assessment tool that can measure this construct among pre-medical students, diagnose possible misconceptions, and inform the design of pedagogically appropriate interventions. In this study, visual representations were limited to those commonly encountered in the chemistry textbooks and prerequisite chemistry courses of pre-medical students.



These include particle models, skeletal structures, ball-and-stick models, electrostatic potential map models, Newman projections, and Fischer projections. These representations were selected because they are frequently used to depict chemical entities and support the learning of chemistry concepts applied in health-related settings.

Research Focus

The objective assessment of knowledge and understanding of visual representations in chemistry is important in diagnosing premedical students' ability to access chemical phenomena at the sub-microscopic level. While previous studies have utilized the Rasch model in developing assessment tools related to conceptual knowledge of visualizations (Taskin et al., 2015) and representational competence (Nickel et al., 2025), this is the first attempt to develop a visual representation test that is theoretically grounded, progression-based, and designed for pre-medical programs. Furthermore, the items were developed to ensure that they are related to the most common types of visual representations in chemistry that are commonly used in the pre-requisite chemistry courses in pre-medical programs, particularly on chemical representations such as particle models, skeletal structures, ball and stick models, electrostatic potential maps, Newman and Fischer projections. Rasch analysis was used to examine the test's construct validity, reliability, and other psychometric properties.

Research Aim and Research Questions

This study aimed to develop and validate the Visual Representations Test in Chemistry (VRTC). Specifically, this study sought to answer the following research questions:

1. What is the item content validity of the VRTC?
2. What are the reliability, separation, item fit, and unidimensionality indices of the VRTC?
3. What evidence supports the construct validity of the VRTC based on the hypothesized progression of knowledge and understanding of visual representations in chemistry?
4. Do VRTC items exhibit differential item functioning based on year level?

Research Methodology

General Background

A quantitative approach using a survey design was employed in this study. The VRTC was designed to measure pre-medical students' knowledge and understanding of common visual representations in pre-medical chemistry courses. To determine the types of visual representations to include in the test, chemistry textbooks were reviewed to identify the most commonly used chemical representations. Interviews were also conducted to determine the most common visual representations used by teachers in their chemistry classes. A letter of communication was sent to various tertiary education institutions that offer pre-medical programs. Initial administration of the test was conducted in AY 2019–2020. Due to the COVID-19 pandemic, data gathering resumed in AY 2022–2023.

Participants

Following ethics approval, letters of coordination were sent to seven selected tertiary institutions, but only four agreed to participate. Pilot testing of the Visual Representations Test in Chemistry (VRTC) was conducted from May 2019 to October 2023, with a total of 827 students, comprising 404 first-year and 423 second-year students. A preliminary review of curricula, course syllabi, instructional materials, and chemistry course sequences was conducted to recruit students with varying levels of prior exposure to chemistry. The first-year cohort comprised 91 BS Biology and 92 BS Pharmacy students from an institution in Benguet, 24 BS Medical Technology students from an institution in Pampanga, and 197 BS Medical Technology students from a university in Manila. These students had completed general chemistry but had not yet taken organic chemistry. The second-year cohort included 423 Medical Laboratory Science students from an institution in Benguet who had completed general, organic, and analytical chemistry and were enrolled in biochemistry. Differences in course sequencing across institutions were expected to yield variation in students' knowledge and understanding of visual representations in chemistry.



Responses were scored dichotomously (1 = correct; 0 = incorrect), encoded in Microsoft Excel, and prepared for analysis. All participants participated voluntarily.

Instrument and Procedures

The first version of VRTC has 26 items: 4 on particle models, 12 on skeletal structures, 5 on ball-and-stick models, 3 on electrostatic potential maps, and 2 on three-dimensional projections (Newman projection and Fischer projection). The test was constructed in a multiple-choice format, with one correct answer and three distractors, following the recommendations of Towns (2014). The description and characteristics of hypothesized difficulty of the items are summarized in Appendix A. Easy items elicit low cognitive load and require low extraneous and essential processing, while difficult items elicit high cognitive load and high extraneous and essential processing. Item difficulty was hypothesized based on the proposed progression of knowledge of visual representations in chemistry. Since sophisticated knowledge of molecular features supports understanding of structure-property relationships (Cooper et al., 2013), the VRTC included common chemical representations found in pre-medical chemistry textbooks and learning materials, consistent with previous literature (Rau, 2015; Rau, 2017b; Sim & Daniel, 2014; Tsapalis et al., 2021). Since organic chemistry prepares pre-medical students for biochemistry (Shulman, 2013), several VRTC items used representations of organic molecules. The test was evaluated by four chemistry professors, two chemistry education professors, and one biochemistry professor to determine whether the test items measure the construct of conceptual understanding of visual representations. After evaluating each item, the item validity index was calculated to determine which items to retain or revise. The evaluators' suggestions were also used to revise the items. Item-content validity was determined using the Lawshe Content Validity Index. The procedure and interpretation of the Lawshe Content Validity Index were based on previous studies (Ayre & Scally, 2014; Polit & Beck, 2006).

Data Analysis

Dichotomous Rasch analysis was conducted using WINSTEPS version 5.9.0. Person and item reliability and separation indices were examined to evaluate measurement precision and the reproducibility of the person and item hierarchies. Reliability coefficients above .80 were interpreted as satisfactory, whereas values from .60 to .80 were considered acceptable for preliminary instrument development (Linacre, 2012; Mohamad et al., 2015). A person separation index of at least 1.50 was considered adequate, with values of 2.00 or higher indicating good discrimination among ability levels (Linacre, 2012). Item fit was evaluated using infit and outfit mean-square (MNSQ) statistics, with values from 0.70 to 1.30 considered acceptable for this low-stakes multiple-choice assessment (Bond & Fox, 2007; Boone et al., 2014). Although standardized fit statistics (ZSTD) were examined, MNSQ values were prioritized because ZSTD is sensitive to large sample sizes (Aryadoust et al., 2020). Infit MNSQ was given greater emphasis because it is less influenced by isolated unexpected responses (Neumann et al., 2013). Point-measure correlations were also inspected to determine whether items contributed positively to the intended construct; values of .30 or higher were considered desirable (Arnold et al., 2018; Li et al., 2016). Unidimensionality was examined through principal component analysis of standardized residuals. A first residual contrast below 2.0 eigenvalue units was taken as evidence that no substantial secondary dimension was present (Linacre & Tennant, 2009). Finally, the alignment between item difficulty and student ability was assessed using the person-item map or Wright map. Mean item difficulties across year levels were compared using analysis of variance with Fisher's least significant difference post hoc test in SPSS version 30.0.

Evaluation of Differential Item Functioning

Differential item functioning was evaluated to determine whether students who have completed general chemistry, organic chemistry, and biochemistry answer items differently compared to students who have not yet completed any chemistry course in premedical programs. Tesio et al. (2023b) remarked that the evaluation of differential item functioning is based on the idea that measures may differ across persons and within persons across time points, and any item that is flagged to exhibit differential item functioning indicates item bias. However, an item that exhibits differential item functioning need not be deleted immediately (Hagquist & Andrich, 2017; Tesio et al., 2023b) and must be evaluated carefully to determine whether the DIF affects person measures. Several studies emphasized that the differential item functioning should also be reported after evaluating the reliability,

construct validity, and unidimensionality of research instruments (Aryadoust et al., 2020; Hagquist & Andrich, 2017; Tesio et al., 2023b).

Research Results

Item-content Validity Indices of VRTC

Table 1 presents the item content validity indices (I-CVI) for VRTC. Based on the results, five items (SS6L2, SS9L1, SS10L1, SS11L3, SS12L3) had I-CVI values between 0.70 and 0.79 and kappa (*k*) values less than 0.74. These items were reviewed and revised based on the evaluators’ suggestions. The items were rephrased to improve clarity, and the options were shortened.

Table 1
Item Content Validity of VRTC

Item	Code	I-CVI	Chance agreement	Kappa (<i>k</i>) statistic	<i>k</i> interpretation and decision
1	PM1L1	1.000*	0.008	1.000	✓✓✓■
2	PM2L2	0.857*	0.055	0.849	✓✓✓■
3	BSM1L1	0.857*	0.055	0.849	✓✓✓■
4	PM3L2	0.857*	0.055	0.849	✓✓✓■
5	SS1L2	0.857*	0.055	0.849	✓✓✓■
6	SS2L1	1.000*	0.008	1.000	✓✓✓■
7	SS3L1	0.857*	0.055	0.849	✓✓✓■
8	SS4L1	1.000*	0.008	1.000	✓✓✓■
9	SS5L2	1.000*	0.008	1.000	✓✓✓■
10	SS6L2	0.571**	0.273	0.410	✓✓✓●
11	SS7L1	1.000*	0.008	1.000	✓✓✓■
12	SS8L1	1.000*	0.008	1.000	✓✓✓■
13	SS9L1	0.714**	0.164	0.658	✓✓✓●
14	SS10L1	0.714**	0.164	0.658	✓✓✓●
15	SS11L3	0.714**	0.164	0.658	✓✓✓●
16	SS12L3	0.714**	0.164	0.658	✓✓✓●
17	BSM2L1	0.857*	0.055	0.849	✓✓✓■
18	BSM3L2	0.857*	0.055	0.849	✓✓✓■
19	BSM4L1	0.857*	0.055	0.849	✓✓✓■
20	BSM5L1	0.857*	0.055	0.849	✓✓✓■
21	BSM6L3	1.000*	0.008	1.000	✓✓✓■
22	EPM1L2	0.857*	0.055	0.849	✓✓✓■
23	EPM2L2	0.857*	0.055	0.849	✓✓✓■
24	EPM3L1	1.000*	0.008	1.000	✓✓✓■
25	3DP1L3	0.857*	0.055	0.849	✓✓✓■
26	3DP2L3	0.857*	0.055	0.849	✓✓✓■

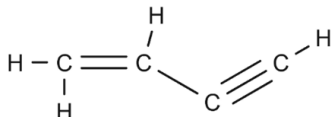
Note: *=appropriate; ** = for revision
✓ = fair; ✓✓ = good; ✓✓✓ = excellent
● = revised; ■ = retained



The “alkenyne” molecular structure (Item 22, EPM1L2) was also modified to depict the accurate bond angle in sp -hybridized carbon atoms. In addition, the molecular structure was converted into an EPM model to illustrate the electron density. The revisions are illustrated below.

Initial item (Item 22)

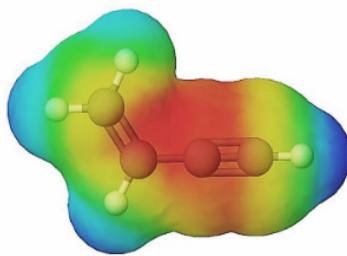
Based on the information presented by the representation below, which statement about the molecule is FALSE?



- A) The molecule has high electron density.
- B) The greatest bond angle in the molecule is 180° .
- C) The molecule has one sp^3 hybridized carbon atom.
- D) The molecular structure has three pi bonds.

Revised item (Item 22)

Based on the information presented by the representation below, which statement about the molecule is FALSE?



- A) Blue parts of the molecule have lower electron density.
- B) The greatest bond angle in the molecule is 180° .
- C) The molecule has one sp^3 hybridized carbon atom.
- D) The molecular structure has three pi bonds.

Reliability of VRTC

The initial 26-item VRTC showed satisfactory item reliability of 0.98 and item separation of 7.68. Person reliability (0.84) and person separation (2.25) also exceeded the recommended thresholds, supporting the test's ability to differentiate students by their knowledge and understanding of visual representations in chemistry and to establish a stable ordering of item difficulty (Linacre, 2012). The first residual contrast was 1.37 eigenvalue units, indicating that the test was sufficiently unidimensional and contained no meaningful secondary dimension (Planinic et al., 2019).

Item Fit

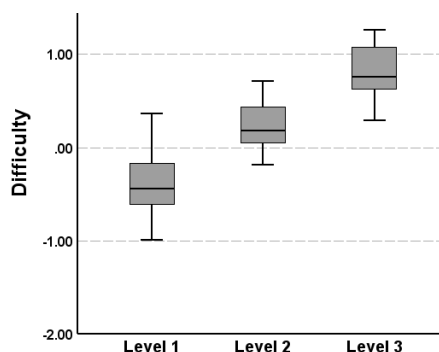
Table 2 shows that all items had acceptable infit and outfit MNSQ values. Although an MNSQ value of 1.00 reflects an ideal fit, values between 0.50 and 1.50 are generally considered appropriate for measurement (Wright & Linacre, 1994). ZSTD values were within the expected ± 2.0 range for all items except Items 7 (SS3L1) and 25 (3DP1L3). However, because MNSQ indices provide the primary basis for judging item fit (Souza et al., 2017), particularly when ZSTD values may be affected by sample size, both items were retained as there are no substantive concerns to support their deletion. In addition, all items had positive point-measure correlations, indicating that each item contributed positively to the measurement of students' knowledge and understanding of visual representations in chemistry.

Table 2
Item Fit Statistics of VRTC

Item	Code	Difficulty	SE	Infit		Outfit		PMC
				MNSQ	ZSTD	MNSQ	ZSTD	
Item 1	PM1L1	-0.17	0.08	0.97	-1.06	0.96	-0.67	0.49
Item 2	PM2L2	0.72	0.08	0.98	-0.62	0.95	-0.80	0.49
Item 3	BSM1L1	-0.44	0.08	1.00	0.09	0.97	-0.48	0.46
Item 4	PM3L2	0.28	0.08	0.96	-1.26	0.93	-1.22	0.50
Item 5	SS1L2	0.04	0.08	0.98	-0.61	0.96	-0.72	0.49
Item 6	SS2L1	-0.51	0.08	0.97	-0.91	0.91	-1.57	0.48
Item 7	SS3L1	-1.59	0.09	1.07	1.55	1.23	2.11	0.36
Item 8	SS4L1	0.37	0.08	1.03	0.84	1.03	0.49	0.46
Item 9	SS5L2	-0.18	0.08	1.05	1.71	1.06	1.08	0.43
Item 10	SS6L2	0.58	0.08	0.94	-1.73	0.90	-1.70	0.51
Item 11	SS7L1	-0.24	0.08	1.00	-0.09	1.05	0.97	0.47
Item 12	SS8L1	-0.14	0.08	1.01	0.18	1.01	0.22	0.46
Item 13	SS9L1	-0.99	0.08	1.03	0.71	1.00	-0.02	0.43
Item 14	SS10L1	-0.84	0.08	0.96	-1.19	0.90	-1.47	0.48
Item 15	SS11L3	0.29	0.08	0.97	-0.77	0.95	-0.95	0.49
Item 16	SS12L3	0.76	0.08	0.99	-0.27	0.97	-0.38	0.48
Item 17	BSM2L1	-0.61	0.08	0.99	-0.18	1.00	0.06	0.46
Item 18	BSM3L2	0.08	0.08	1.00	0.12	1.01	0.21	0.47
Item 19	BSM4L1	-0.32	0.08	1.03	0.82	1.05	0.83	0.45
Item 20	BSM5L1	0.13	0.08	0.99	-0.37	0.97	-0.55	0.48
Item 21	BSM6L3	0.63	0.08	1.01	0.34	1.01	0.24	0.46
Item 22	EPM1L2	0.30	0.08	1.00	0.10	0.97	-0.56	0.47
Item 23	EPM2L2	0.07	0.08	1.04	1.34	1.04	0.76	0.45
Item 24	EPM3L1	-0.54	0.08	0.99	-0.21	1.00	0.05	0.46
Item 25	3DP1L3	1.08	0.09	1.06	1.36	1.23	2.64	0.42
Item 26	3DP2L3	1.27	0.09	0.99	-0.21	1.11	1.17	0.46

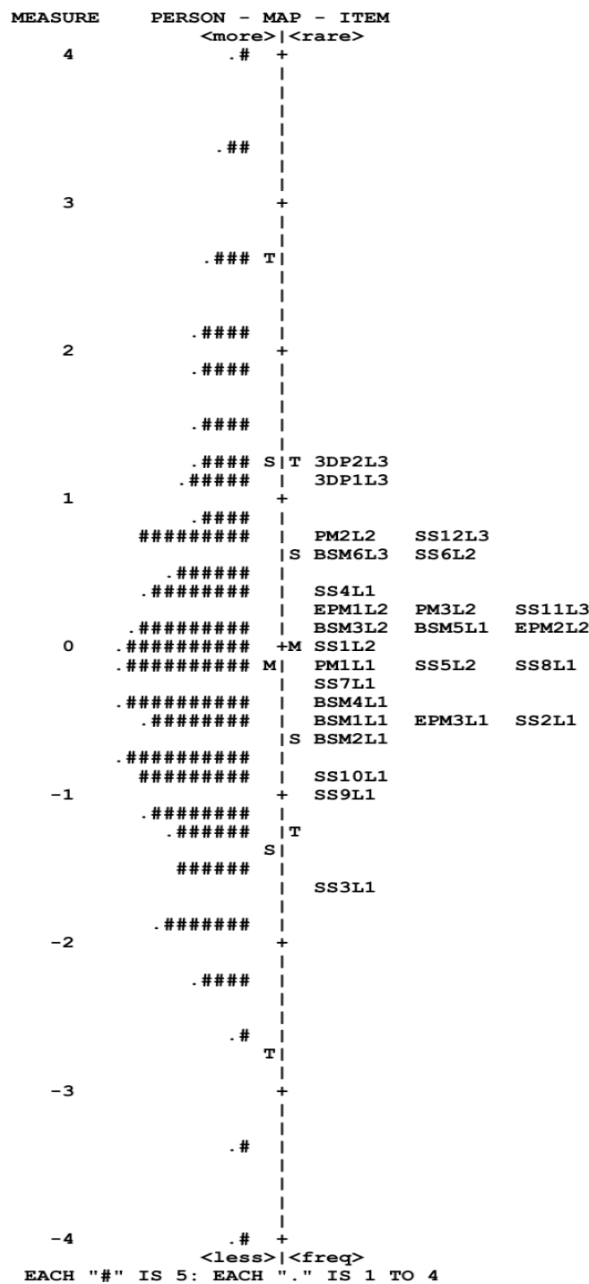
Item Hierarchy as a Measure of Construct Validity

Figure 2 shows a clear increase in mean item difficulty across the three VRTC levels. Level 1 items were the least difficult ($M = -0.45$, $SD = 0.50$), followed by Level 2 items ($M = 0.24$, $SD = 0.30$) and Level 3 items ($M = 0.81$, $SD = 0.38$). This ordering is consistent with the hypothesized progression of the construct among pre-medical students. A one-way ANOVA indicated significant differences in mean item measures across the three levels, $F(2, 23) = 17.273$, $p < .001$. Post hoc Fisher's LSD comparisons further showed that Level 1 items were significantly easier than Level 2 items, which were, in turn, significantly easier than Level 3 items.

Figure 2
Mean Difficulty of Items in VRTC

Person-Item Map or Wright Map

Figure 3 presents the person-item map (Wright map), which locates student ability and item difficulty on a common logit scale and indicates the likelihood of a correct response as a function of their relative positions (Fulmer, 2014). Student ability estimates ranged from -4.66 to 4.64 logits, exceeding the item-difficulty range of -1.59 to 1.27 logits. This pattern indicates that the VRTC item difficulty was within the ability of pre-medical students. Item 26 (3DP2L3) was the most difficult, followed by Item 25 (3DP1L3), whereas Items 07 (SS3L1), 13 (SS9L1), and 14 (SS10L1) were the easiest. The map also revealed gaps between some item locations and student ability levels, suggesting that additional items may be needed to improve measurement coverage across the full range of student abilities.

Figure 3*Wright Map of the Item Pool of VRTC*

Differential Item Functioning

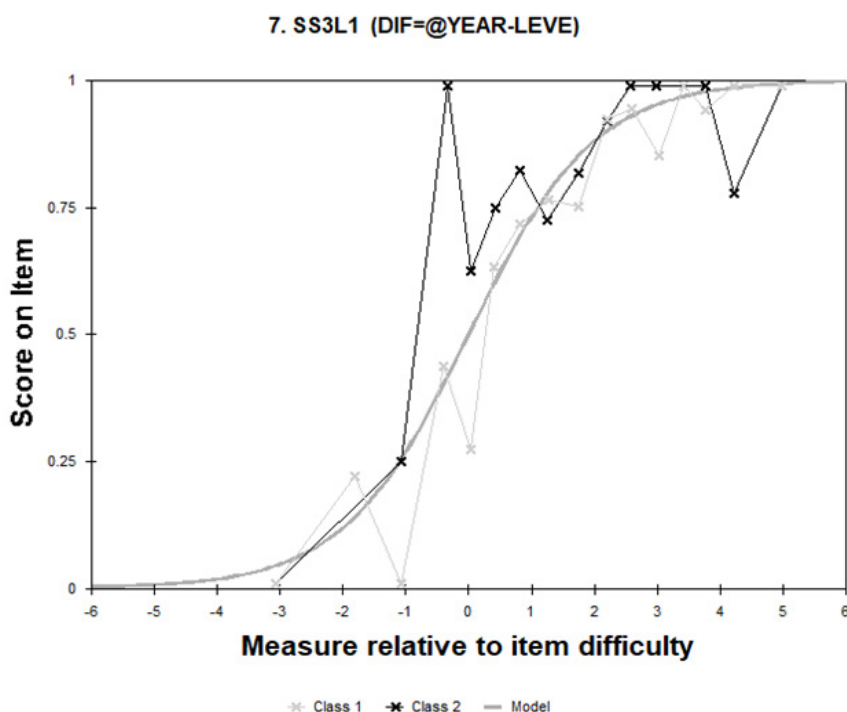
Table 3 summarizes the differential item functioning (DIF) analysis. DIF was considered present when the contrast exceeded 0.50 logits and the Welch test was statistically significant ($p < .05$) (Linacre, 2012; Moradi et al., 2022). Only Item 7 (SS3L1) satisfied both criteria, indicating differential functioning across the groups examined.

Table 3*DIF of Items in VRTC*

Item	Code	Person	DIF Measure	DIF SE	Person	DIF Measure	DIF SE	DIF Contrast	Joint SE	Rasch-Welch t	df	p
1	PM1L1	C1	-0.22	0.10	C2	-0.03	0.14	-0.19	0.17	-1.11	470	.269
2	PM2L2	C1	0.69	0.10	C2	0.79	0.15	-0.10	0.18	-0.55	467	.582
3	BSM1L1	C1	-0.39	0.10	C2	-0.57	0.15	0.19	0.17	1.07	466	.285
4	PM3L2	C1	0.28	0.10	C2	0.24	0.15	0.04	0.17	0.24	472	.810
5	SS1L2	C1	0.04	0.10	C2	0.07	0.14	-0.03	0.17	-0.20	470	.845
6	SS2L1	C1	-0.51	0.10	C2	-0.51	0.14	0.00	0.17	0.00	470	1.000
7	SS3L1	C1	-1.43	0.11	C2	-2.01	0.18	0.59	0.21	2.76	417	.006
8	SS4L1	C1	0.33	0.10	C2	0.47	0.15	-0.15	0.18	-0.83	466	.405
9	SS5L2	C1	-0.10	0.10	C2	-0.38	0.14	0.29	0.17	1.66	469	.097
10	SS6L2	C1	0.62	0.10	C2	0.52	0.15	0.10	0.18	0.56	475	.579
11	SS7L1	C1	-0.24	0.10	C2	-0.22	0.14	-0.02	0.17	-0.13	470	.897
12	SS8L1	C1	-0.14	0.10	C2	-0.12	0.14	-0.02	0.17	-0.14	470	.887
13	SS9L1	C1	-0.99	0.10	C2	-0.99	0.15	0.00	0.18	0.00	467	1.000
14	SS10L1	C1	-0.88	0.10	C2	-0.77	0.15	-0.11	0.18	-0.62	473	.536
15	SS11L3	C1	0.22	0.10	C2	0.45	0.15	-0.23	0.18	-1.29	464	.197
16	SS12L3	C1	0.74	0.10	C2	0.81	0.15	-0.07	0.18	-0.40	468	.688
17	BSM2L1	C1	-0.67	0.10	C2	-0.49	0.14	-0.18	0.17	-1.01	474	.311
18	BSM3L2	C1	0.08	0.10	C2	0.05	0.14	0.03	0.17	0.18	471	.856
19	BSM4L1	C1	-0.42	0.10	C2	-0.12	0.14	-0.30	0.17	-1.74	472	.083
20	BSM5L1	C1	0.05	0.10	C2	0.30	0.15	-0.25	0.17	-1.45	466	.148
21	BSM6L3	C1	0.63	0.10	C2	0.63	0.15	0.00	0.18	0.00	471	1.000
22	EPM1L2	C1	0.30	0.10	C2	0.26	0.15	0.04	0.17	0.23	472	.817
23	EPM2L2	C1	0.17	0.10	C2	-0.14	0.14	0.30	0.17	1.76	474	.079
24	EPM3L1	C1	-0.54	0.10	C2	-0.51	0.14	-0.03	0.17	-0.17	471	.863
25	3DP1L3	C1	1.05	0.11	C2	1.15	0.16	-0.10	0.19	-0.50	467	.616
26	3DP2L3	C1	1.39	0.11	C2	1.00	0.16	0.39	0.19	2.04	504	.042

Note: C1 (Class 1) = first year cohort; C2 (Class 2) = second year cohort

As shown in Figure 4, both groups had a low probability of answering the item correctly at lower ability levels. However, between approximately -2.0 and -1.0 logits, students who had completed pre-medical chemistry courses showed a higher probability of a correct response despite relatively low overall ability. At higher ability levels, the probability of success increased for both groups. However, students who had not yet completed pre-medical chemistry courses demonstrated a slightly higher probability of answering the item correctly.

Figure 4*Characteristic Curve of Item 7 (SS3L1)*

Discussion

Visual representations in chemistry provide epistemic access to chemical phenomena (Evagorou et al., 2015; Silva & Sasseron, 2025). Previous studies emphasized the need to examine how students interpret the implicit and explicit features of visual representations and develop structure–property reasoning in chemistry (Chiu et al., 2018; Kohn et al., 2018; Talanquer, 2018; Talanquer, 2022). Although chemical expertise is important, chemistry education in pre-medical and health science programs often emphasizes applying chemical concepts in professional courses and in understanding disease processes, diagnostics, nutrition, and pharmacology (Armstrong & Pöe, 2020; Barcelo & Ferido, 2023; Matlin et al., 2017). Thus, visual representations are necessary for transferring chemical knowledge to health-related applications. Assessing pre-medical students' knowledge of visual representations can therefore inform teaching interventions that improve conceptual understanding and support the application of chemistry to health concepts in prerequisite courses.

Overall, the results suggest that the VRTC exhibits adequate reliability, item fit, and unidimensionality. The item hierarchy in the person-item map also supports the construct validity of the hypothesized progression of visual representation knowledge. Items requiring interpretation of explicit features, such as counting carbon atoms or determining molecular formulas from structures, showed low difficulty. Items involving implicit features were moderately difficult, whereas items requiring representational equivalence and inference of physical or chemical properties were more difficult. This hierarchy suggests that difficulty in visual representations is anchored in Cognitive Load Theory and Cognitive Theory of Multimedia Learning. As Talanquer (2022) noted, making molecular inferences is difficult when implicit features are not recognized. Cognitive load may also increase when students lack familiar schemata for interpreting a representation (Schuessler et al., 2024).

However, some items showed different levels of difficulty than initially hypothesized. For example, Item 8 (SS4L1) and Item 20 (BSM5L1) were initially classified as easy but were found to be moderately difficult. Item 8 required the identification of lone pairs, while Item 20 required determining a molecular formula from a ball-and-stick model. Although these features were explicitly shown, students may not have been familiar with the related concepts needed to interpret them. Consistent with the hypothesized difficulty, three-dimensional projection items, such as Item 25 (3DP1L3) and Item 26 (3DP2L3), involved Level 3 skills and required mental manipulation of visual representations.



Item 2 (PM2L2) was also more difficult than some Level 3 items. This item involved iodine sublimation after heating, where some students appeared to assume that covalent bonds in diatomic iodine molecules were disrupted. Item 16 (SS12L3), classified as difficult, involved two competencies: knowledge of wedge-dash diagrams and bond angles based on VSEPR theory. Its difficulty may stem from students' unfamiliarity with wedge-dash diagrams, which have been reported as challenging in previous research (Taskin et al., 2015). Similarly, Item 10 (SS6L2) involved orbital hybridization, a concept known to be abstract, theoretical, and complex (Elliott & Hughes, 2025). Although hybridization is important for understanding chemical bonding, it has also been identified as difficult for students (Brass et al., 2024; Marifa et al., 2023; Wegwerth et al., 2021).

A good assessment tool should include items with varying levels of difficulty to measure a construct effectively (Nedungadi et al., 2019). However, the person-item map indicated that several items had similar difficulty levels, including BSM1L1, EPM3L1, and SS2L1. Although removing items with similar difficulty has been recommended (Boone, 2016), these items showed acceptable item fit and point-measure correlation values. Thus, they remain useful for measuring knowledge of visual representations and should not be removed from the VRTC immediately. However, items with similar logit values may be deleted if they have standardized residual correlations greater than 0.30, indicating that the items do not exhibit local independence, thereby inflating the test's reliability (Linacre, 2023). Since the largest standardized residual correlation in this study was less than 0.30, items with similar logit values may be modified further, but not automatically deleted. The person-item map also revealed gaps in student abilities. These gaps may suggest limitations in item precision, but they may also reflect sampling limitations. Since the test was designed to monitor pre-medical students' progression before and after prerequisite chemistry courses, more respondents with heterogeneous levels of knowledge and understanding of visual representations may be needed in future revisions.

Item 7 (SS3L1) exhibited differential item functioning across the groups examined. This item required students to determine the number of carbon atoms represented in a skeletal structure. Although this task was classified as a selection-level item (Level 1), its interpretation depends on familiarity with the line-angle convention that carbon atoms are implied at vertices. The item was easier for students who had completed additional coursework in general, organic, and analytical chemistry. The difference may be attributed to a greater exposure to skeletal structures and repeated opportunities to interpret implicit carbon atoms during prior chemistry instruction. Thus, the finding may reflect curricular sequencing, instructional differences, or other unmeasured cohort characteristics. While this item may exhibit differential item functioning, the item-fit statistics and a positive point-measure correlation suggest that the item is helpful in measuring the construct and may not necessarily require deletion. However, this item needs further review and re-evaluation using more heterogeneous samples in succeeding iterative analyses.

The results indicate the need for further refinement of the test or additional data collection involving pre-medical students with more heterogeneous knowledge of visual representations in chemistry. Rasch analysis provided evidence of item fit and unidimensionality, suggesting that the items are appropriate for measuring the intended construct. However, the findings may also indicate that the participants did not represent the full range of abilities needed to evaluate the first version of the test. Overall, the first version of the Visual Representations Test in Chemistry appears to be reliable and valid and may be used to measure the construct among pre-medical students particularly in the Philippine setting.

Conclusions and Implications

Rasch analysis provided evidence for the reliability, unidimensionality, item fit, and construct validity of the first version of VRTC. The observed item hierarchy was consistent with the proposed progression of the construct, allowing ability estimates to be interpreted in relation to this developmental sequence. These findings support the potential use of the VRTC as a diagnostic assessment in pre-medical programs to identify students' representational strengths and learning needs. The instrument may also guide chemistry educators in designing learning activities that strengthen students' use of visual representations and support the transfer of chemistry knowledge to health-related contexts. For students needing support at Level 1, teachers may use guided feature-identification tasks, labelling activities, and explicit discussion of what visual elements represent. For Level 2, teachers may use paired representations and prompts that require students to infer information not directly displayed, such as polarity, electron density, or molecular geometry. For Level 3, teachers may use translation tasks across skeletal, ball-and-stick, wedge-dash, Fischer, and Newman representations, followed by structure–property reasoning in chemistry and health-related contexts.



Limitations

This study has several limitations. First, the proposed progression of the construct was derived from Filipino pre-medical students with varying levels of exposure to college chemistry, ranging from those who had completed general chemistry to those who had taken organic or analytical chemistry. Its applicability to other student populations and curricular contexts, therefore, requires further investigation. Second, the VRTC focused on selected chemical representations, including particle models, skeletal structures, ball-and-stick models, electrostatic potential maps, Newman projections, and Fischer projections; it does not capture the full range of representations used in chemistry. Further research should examine differential item functioning across grade levels and genders, particularly because students' exposure to chemistry increases across the curriculum, and pre-medical cohorts may be unevenly distributed by gender. Increasing the item pool would also improve coverage and precision across the ability continuum. In addition, the present multiple-choice format used three distractors per item. Future studies may explore a two-tier format to capture students' reasoning and distinguish correct responses based on conceptual understanding from those obtained through guessing. Finally, only four institutions participated during data gathering. The data may not reflect variations in curricula, instructional practices, educational resources, and student characteristics across Philippine higher education. Replication with a larger, more diverse sample of institutions across geographic regions and institutional types is needed to strengthen the generalizability of the findings.

Acknowledgements

The authors would like to acknowledge the Department of Science and Technology - Capacity Building Program in Science and Mathematics Education for providing support during the conduct of the study.

Ethics Statement

The protocol for the study was submitted for ethics review to the Saint Louis University Research Ethics Committee on March 6, 2019. The protocol was approved on April 10, 2019, and was assigned ethics approval certificate number SLU-REC 2019-044.

Declaration of Interest

The authors have no conflicts of interest to declare.

References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>
- Al-Balushi, S. M., & Al-Hajri, S. H. (2014). Associating animations with concrete models to enhance students' comprehension of different visual representations in organic chemistry. *Chemistry Education Research and Practice*, 15(1), 47–58. <https://doi.org/10.1039/C3RP00074E>
- Armstrong, D., & Poë, J. C. (2020). The science of human health—a context-based chemistry course for non-science majors incorporating systems thinking. *Journal of Chemical Education*, 97(11), 3957–3965. <https://doi.org/10.1021/acs.jchemed.0c00887>
- Arnold, J. C., Boone, W. J., Kremer, K., & Mayer, J. (2018). Assessment of competencies in scientific inquiry through the application of Rasch measurement techniques. *Education Sciences*, 8(4), Article 184. <https://doi.org/10.3390/educsci8040184>
- Aryadoust, V., Ng, L. Y., & Sayama, H. (2020). A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research. *Language Testing*, 38(1), 6–40. <https://doi.org/10.1177/0265532220927487>
- Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79–86. <https://doi.org/10.1177/0748175613513808>
- Barcelo, J. M., & Ferido, N. M. B. (2023). Validating the progression of chemistry-based health literacy: An application of Rasch analysis. In X. Liu & W. J. Boone (Eds.), *Advances in applications of Rasch measurement in science education* (pp. 213–239). Springer. https://doi.org/10.1007/978-3-031-28776-3_9
- Barcelo, J. M., & Ferido, N. M. (2025). Evaluation of false, reductionist part-whole relationships in biochemistry and its effect on health science students' chemistry-based health literacy. *Science & Education*, 34(6), 4683–4716. <https://doi.org/10.1007/s11191-025-00642-3>
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Lawrence Erlbaum Associates.



- Boone, W. J. (2016). Rasch analysis for instrument development: Why, when, and how? *CBE—Life Sciences Education*, 15(4), rm4. <https://doi.org/10.1187/cbe.16-04-0148>
- Boone, W. J., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education*, 4(1), Article 1416898. <https://doi.org/10.1080/2331186X.2017.1416898>
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer. <https://doi.org/10.1007/978-94-007-6857-4>
- Brass, R., Braithwaite, E., Edwards, H. E., Kaur, J., Kleanthous, A., Madhlangobe, T. T., Mistry, A. D., Suma, A., Hin, S. L., & Williams, D. P. (2024). The need for SPD: A hybridization-based card game for students on introductory chemistry courses. *Journal of Chemical Education*, 101(9), 4030–4036. <https://doi.org/10.1021/acs.jchemed.3c01204>
- Carvalho, I., Borges, Á. D. L., & Bernardes, L. S. C. (2005). Medicinal chemistry and molecular modeling: An integration to teach drug structure-activity relationship and the molecular basis of drug action. *Journal of Chemical Education*, 82(4), 588. <https://doi.org/10.1021/ed082p588>
- Chi, S., Wang, Z., Luo, M., Yang, Y., & Huang, M. (2018). Student progression on chemical symbol representation abilities at different grade levels (Grades 10–12) across gender. *Chemistry Education Research and Practice*, 19(4), 1055–1064. <https://doi.org/10.1039/C8RP00010G>
- Chiu, M.-H., Chou, C.-C., Chen, Y.-H., Hung, T., Tang, W.-T., Hsu, J.-W., Liaw, H. L., & Tsai, M.-K. (2018). Model-based learning about structures and properties of chemical elements and compounds via the use of augmented realities. *Chemistry Teacher International*, 1(1), Article 20180002. <https://doi.org/10.1515/cti-2018-0002>
- Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, 90(6), 1073–1091. <https://doi.org/10.1002/sce.20164>
- Cook, M., Wiebe, E. N., & Carter, G. (2008). The influence of prior knowledge on viewing and interpreting graphics with macroscopic and molecular representations. *Science Education*, 92(5), 848–867. <https://doi.org/10.1002/sce.20262>
- Cooper, M. M., Corley, L. M., & Underwood, S. M. (2013). An investigation of college chemistry students' understanding of structure-property relationships. *Journal of Research in Science Teaching*, 50(6), 699–721. <https://doi.org/10.1002/tea.21093>
- Davis, D. R., & Boone, W. J. (2021). Using Rasch analysis to evaluate the psychometric functioning of the other-directed, light-hearted, intellectual, and whimsical (OLIW) adult playfulness scale. *International Journal of Educational Research Open*, 2, Article 100054. <https://doi.org/10.1016/j.ijedro.2021.100054>
- Elliott, M. C., & Hughes, C. E. (2025). The interface of hybridization, hyperconjugation, and conformation in organic chemistry teaching. *Journal of Chemical Education*, 102(8), 3459–3467. <https://doi.org/10.1021/acs.jchemed.5c00445>
- Evagorou, M., Erduran, S., & Mäntylä, T. (2015). The role of visual representations in scientific practices: From conceptual understanding and knowledge generation to “seeing” how science works. *International Journal of STEM Education*, 2, Article 11. <https://doi.org/10.1186/s40594-015-0024-x>
- Fang, X., Cao, S., & Liu, R. (2011). Interaction of bisphenol A with bovine hemoglobin using spectroscopic and molecular modeling methods. *Applied Spectroscopy*, 65(11), 1250–1253. <https://doi.org/10.1366/11-06357>
- Fulmer, G. W. (2014). Validating proposed learning progressions on force and motion using the Force Concept Inventory: Findings from Singapore secondary schools. *International Journal of Science and Mathematics Education*, 13(6), 1235–1254. <https://doi.org/10.1007/s10763-014-9553-x>
- Hagquist, C., & Andrich, D. (2017). Recent advances in analysis of differential item functioning in health research using the Rasch model. *Health and Quality of Life Outcomes*, 15, Article 181. <https://doi.org/10.1186/s12955-017-0755-0>
- Höst, G. E., Schönborn, K. J., & Palmerius, K. E. (2012). Students' use of three different visual representations to interpret whether molecules are polar or nonpolar. *Journal of Chemical Education*, 89(12), 1499–1505. <https://doi.org/10.1021/ed2001895>
- Kohn, K. P., Underwood, S. M., & Cooper, M. M. (2018). Connecting structure–property and structure–function relationships across the disciplines of chemistry and biology: Exploring student perceptions. *CBE—Life Sciences Education*, 17(2), ar21. <https://doi.org/10.1187/cbe.18-01-0004>
- Kozma, R., Chin, E., Russell, J., & Marx, N. (2000). The roles of representations and tools in the chemistry laboratory and their implications for chemistry learning. *Journal of the Learning Sciences*, 9(2), 105–143. https://doi.org/10.1207/S15327809JLS0902_1
- Li, C.-Y., Romero, S., Bonilha, H. S., Simpson, K. N., Simpson, A. N., Hong, I., & Vellozo, C. A. (2016). Linking existing instruments to develop an activity of daily living item bank. *Evaluation & the Health Professions*, 41(1), 25–43. <https://doi.org/10.1177/0163278716676873>
- Linacre, J. M. (2012). *A user's guide to WINSTEPS MINISTEP Rasch-model computer programs*. Winsteps.com.
- Linacre, J. M. (2023). *A user's guide to WINSTEPS MINISTEP Rasch-model computer programs*. Winsteps.com. <https://www.winsteps.com/a/Winsteps-Manual.pdf>
- Linacre, J. M., & Tennant, A. (2009). More about critical eigenvalue sizes (variances) in standardized-residual principal components analysis (PCA). *Rasch Measurement Transactions*, 23(3), 1228.
- Liu, X., & Boone, W. J. (2023). Introduction to advances in applications of Rasch measurement in science education. In X. Liu & W. J. Boone (Eds.), *Advances in applications of Rasch measurement in science education* (pp. 1–11). Springer. https://doi.org/10.1007/978-3-031-28776-3_1
- Luisi, P. L. (2002). Emergence in chemistry: Chemistry as the embodiment of emergence. *Foundations of Chemistry*, 4(3), 183–200. <https://doi.org/10.1023/A:1020672005348>
- Luxford, C. J., & Bretz, S. L. (2014). Development of the bonding representations inventory to identify student misconceptions about covalent and ionic bonding representations. *Journal of Chemical Education*, 91(3), 312–320. <https://doi.org/10.1021/ed400700q>
- Marifa, H. A., Abukari, M. A., Samari, J. A., Dorsah, P., & Abudu, F. (2023). Chemistry teachers' pedagogical content knowledge in teaching hybridization. *Pedagogical Research*, 8(3), Article em0132. <https://doi.org/10.29333/pr/13168>
- Mason, B., Rau, M. A., & Nowak, R. (2019). Cognitive task analysis for implicit knowledge about visual representations with similarity learning methods. *Cognitive Science*, 43(9), Article e12744. <https://doi.org/10.1111/cogs.12744>



- Matlin, S. A., Mehta, G., Krief, A., & Hopf, H. (2017). The chemical sciences and health: Strengthening synergies at a vital interface. *ACS Omega*, 2(10), 6819–6821. <https://doi.org/10.1021/acsomega.7b01463>
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- Mohamad, M. M., Sulaiman, N. L., Sern, L. C., & Salleh, K. M. (2015). Measuring the validity and reliability of research instruments. *Procedia—Social and Behavioral Sciences*, 204, 164–171. <https://doi.org/10.1016/j.sbspro.2015.08.129>
- Moradi, E., Ghabanchi, Z., & Pishghadam, R. (2022). Reading comprehension test fairness across gender and mode of learning: Insights from IRT-based differential item functioning analysis. *Language Testing in Asia*, 12, Article 20. <https://doi.org/10.1186/s40468-022-00192-3>
- Nedungadi, S., Paek, S. H., & Brown, C. E. (2019). Utilizing Rasch analysis to establish the psychometric properties of a concept inventory on concepts important for developing proficiency in organic reaction mechanisms. *Chemistry Teacher International*, 2(2), Article 20190004. <https://doi.org/10.1515/cti-2019-0004>
- Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, 50(2), 162–188. <https://doi.org/10.1002/tea.21061>
- Nickel, S., Brockmüller, S., Boone, W. J., & Habig, S. (2025). Measuring representational competence—Analyses of dimensionality and the relationship to general content knowledge. *International Journal of Science Education*, 1–25. Advance online publication. <https://doi.org/10.1080/09500693.2025.2515621>
- Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, 15(2), Article 020111. <https://doi.org/10.1103/PhysRevPhysEduRes.15.020111>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Prigozhin, M. B., Scott, G. E., & Denos, S. (2014). Mechanical modeling and computer simulation of protein folding. *Journal of Chemical Education*, 91(11), 1939–1942. <https://doi.org/10.1021/ed400719c>
- Rau, M. A. (2015). Enhancing undergraduate chemistry learning by helping students make connections among multiple graphical representations. *Chemistry Education Research and Practice*, 16(3), 654–669. <https://doi.org/10.1039/C5RP00065C>
- Rau, M. A. (2016). Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review*, 29(4), 717–761. <https://doi.org/10.1007/s10648-016-9365-3>
- Rau, M. A. (2017a). Making connections among multiple visual representations: How do sense-making skills and perceptual fluency relate to learning of chemistry knowledge? *Instructional Science*, 46(2), 209–243. <https://doi.org/10.1007/s11251-017-9431-3>
- Rau, M. A. (2017b). Do knowledge-component models need to incorporate representational competencies? *International Journal of Artificial Intelligence in Education*, 27(2), 298–319. <https://doi.org/10.1007/s40593-016-0134-8>
- Rau, M. A., Michaelis, J. E., & Fay, N. (2015). Connection making between multiple graphical representations: A multi-methods approach for domain-specific grounding of an intelligent tutoring system for chemistry. *Computers & Education*, 82, 460–485. <https://doi.org/10.1016/j.compedu.2014.12.009>
- Rodemer, M., Eckhard, J., Graulich, N., & Bernholt, S. (2021). Connecting explanations to representations: Benefits of highlighting techniques in tutorial videos on students' learning in organic chemistry. *International Journal of Science Education*, 43(17), 2707–2728. <https://doi.org/10.1080/09500693.2021.1985743>
- Schnotz, W. (2014). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). Cambridge University Press.
- Schönborn, K. J., & Anderson, T. R. (2010). Bridging the educational research-teaching practice gap. *Biochemistry and Molecular Biology Education*, 38(5), 347–354. <https://doi.org/10.1002/bmb.20436>
- Schuessler, K., Rodemer, M., Giese, M., & Walpuski, M. (2024). Organic Chemistry and the challenge of representations: Student difficulties with different representation forms when switching from paper–pencil to digital format. *Journal of Chemical Education*, 101(11), 4566–4579. <https://doi.org/10.1021/acs.jchemed.4c00303>
- Sevian, H., & Talanquer, V. (2014). Rethinking chemistry: A learning progression on chemical thinking. *Chemistry Education Research and Practice*, 15(1), 10–23. <https://doi.org/10.1039/C3RP00111C>
- Shulman, J. I. (2013). Chemistry in the premedical curriculum: Considering the options. *Journal of Chemical Education*, 90(7), 813–815. <https://doi.org/10.1021/ed3008632>
- Silva, F. C., & Sasseron, L. H. (2025). The positioning of visual representations as epistemic objects for the teaching of organic chemistry. *Journal of Chemical Education*, 102(2), 615–620. <https://doi.org/10.1021/acs.jchemed.4c01018>
- Sim, J. H., & Daniel, E. G. S. (2014). Representational competence in chemistry: A comparison between students with different levels of understanding of basic chemical concepts and chemical representations. *Cogent Education*, 1(1), Article 991180. <https://doi.org/10.1080/2331186X.2014.991180>
- Souza, M. A., Coster, W. J., Mancini, M. C., Dutra, F. C., Kramer, J., & Sampaio, R. F. (2017). Rasch analysis of the participation scale (P-scale): Usefulness of the P-scale to a rehabilitation services network. *BMC Public Health*, 17, Article 934. <https://doi.org/10.1186/s12889-017-4945-9>
- Stieff, M. (2011). When is a molecule three dimensional? A task-specific role for imagistic reasoning in advanced chemistry. *Science Education*, 95(2), 310–336. <https://doi.org/10.1002/sce.20427>
- Stull, A. T., Gainer, M., Padalkar, S., & Hegarty, M. (2016). Promoting representational competence with molecular models in organic chemistry. *Journal of Chemical Education*, 93(6), 994–1001. <https://doi.org/10.1021/acs.jchemed.6b00194>
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive load theory in perspective. In *Cognitive load theory* (pp. 237–242). Springer. https://doi.org/10.1007/978-1-4419-8126-4_18



- Talanquer, V. (2007). Students' predictions about the sensory properties of chemical compounds: Additive versus emergent frameworks. *Science Education*, 92(1), 96–114. <https://doi.org/10.1002/sce.20235>
- Talanquer, V. (2018). Progressions in reasoning about structure–property relationships. *Chemistry Education Research and Practice*, 19(4), 998–1009. <https://doi.org/10.1039/C7RP00187H>
- Talanquer, V. (2022). The complexity of reasoning about and with chemical representations. *JACS Au*, 2(12), 2658–2669. <https://doi.org/10.1021/jacsau.2c00498>
- Taskin, V., Bernholt, S., & Parchmann, I. (2015). An inventory for measuring student teachers' knowledge of chemical representations: Design, validation, and psychometric analysis. *Chemistry Education Research and Practice*, 16(3), 460–477. <https://doi.org/10.1039/C4RP00214H>
- Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2023a). Interpreting results from Rasch analysis 1: The “most likely” measures coming from the model. *Disability and Rehabilitation*, 46(3), 591–603. <https://doi.org/10.1080/09638288.2023.2169771>
- Tesio, L., Caronni, A., Simone, A., Kumbhare, D., & Scarano, S. (2023b). Interpreting results from Rasch analysis 2: Advanced model applications and the data-model fit assessment. *Disability and Rehabilitation*, 46(3), 604–617. <https://doi.org/10.1080/09638288.2023.2169772>
- Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education*, 91(9), 1426–1431. <https://doi.org/10.1021/ed500076x>
- Tsarpalis, G., Pantazi, G., Pappa, E. T., & Byers, B. (2021). Using electrostatic potential maps as visual representations to promote better understanding of chemical bonding. *Chemistry Teacher International*, 3(4), 391–411. <https://doi.org/10.1515/cti-2021-0012>
- Wang, L. (2012). Using molecular modeling in teaching group theory analysis of the infrared spectra of organometallic compounds. *Journal of Chemical Education*, 89(3), 360–364. <https://doi.org/10.1021/ed200538c>
- Wegwerth, S. E., Overby, J. S., Douglas, C. J., Winter, J. E., Manchester, G. J., & Engalan, J. (2021). From abstract to manipulatable: The hybridization explorer, a digital interactive for studying orbitals. *Journal of Chemical Education*, 98(2), 655–661. <https://doi.org/10.1021/acs.jchemed.0c00847>
- Weinrich, M. L., & Talanquer, V. (2016). Mapping students' modes of reasoning when thinking about chemical reactions used to make a desired product. *Chemistry Education Research and Practice*, 17(2), 394–406. <https://doi.org/10.1039/C5RP00208G>
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- Zarkadis, N., Papageorgiou, G., & Stamovlasis, D. (2017). Studying the consistency between and within student mental models for atomic structure. *Chemistry Education Research and Practice*, 18(4), 893–902. <https://doi.org/10.1039/C7RP00135E>
- Zoechling, S., Hopf, M., Woithe, J., & Schmeling, S. (2022). Students' interest in particle physics: Conceptualisation, instrument development, and evaluation using Rasch theory and analysis. *International Journal of Science Education*, 44(15), 2353–2380. <https://doi.org/10.1080/09500693.2022.2122897>

Received: May 03, 2026

Revised: June 26, 2026

Accepted: August 22, 2026

Cite as: Barcelo, J. M., & Araneta, P. L. G. N. (2026). Development and validation of the visual representations test in chemistry for pre-medical students. *Journal of Baltic Science Education*, 25(4), 600–616. <https://doi.org/10.33225/jbse/26.25.600>



Jonathan M. Barcelo
(Corresponding author)

PhD, RN, Professor, Saint Louis University, Baguio City, Benguet, Philippines.
E-mail: JonMBarc3y@gmail.com
ORCID: <https://orcid.org/0000-0002-0022-7470>

Precious Lady Gine N. Araneta

MPH, RMT, Assistant Professor, Saint Louis University, Baguio City, Benguet, Philippines.
E-mail: plgnaraneta@slu.edu.ph
ORCID: <https://orcid.org/0009-0007-6135-600X>

