Jurnal Pendidikan MIPA
e-ISSN: 2685-5488 | p-ISSN: 1411-2531
URL: https://jpmipa.fkip.unila.ac.id/index.php/jpmipa
Jounal's DOI: http://doi.org/10.23960/jpmipa

Vol. 27, No. 2, 2026
Article ID: 1137
Pages: 801-820

Confirmatory Factor Analysis of a Scientific Imagination Assessment for the Bohr Atomic Model

Fifin Dewi Ratnasari[1,3], Masturi[1,*], Suharto Linuwih[1], Ani Rusilowati[1], Endang Susilaningsih[2], & Dwi Yulianti[1]
[1]Department of Physics, Universitas Negeri Semarang, Indonesia
[2]Department of Chemistry, Universitas Negeri Semarang, Indonesia
[3]Department of Natural Science, Universitas Negeri Semarang, Indonesia

*Corresponding email: masturi@mail.unnes.ac.id

Received: 20 December 2025 *Accepted: 17 February 2026* *Published: 07 May 2026*

Abstract: The scientific imagination is not merely a source of daydreaming and creativity; it helps define, build, and apply scientific models, conduct scientific tests, develop scientific theories, and advance scientific knowledge and concepts. The focus of this study was to develop an instrument to evaluate students' scientific imagination as they explore the Bohr atomic model and to validate it through CFA. A research and development methodology was employed, with 101 students enrolled in the physics program at Universitas Negeri Semarang (UNNES) who had previously completed coursework on the Bohr model. The validity of the instrument developed for this purpose was established with confirmatory factor analysis (CFA). Descriptive analysis was conducted to profile physics students' abilities in scientific imagination when analyzing the Bohr atomic model. The findings of this study demonstrated that the scientific imagination assessment had acceptable validity and reliability based on the following parameters: comparative fit index (CFI) of 0.967, Tucker-Lewis index (TLI) of 0.957, root mean square error of approximation (RMSEA) of 0.060, standardized root mean square residual (SRMR) of 0.051, and goodness of fit index (GFI) of greater than 0.978. Scientific imagination can be divided into two dimensions: creative imagination and reproductive imagination. Creative imagination is divided into five sub-dimensions: intuition, sensitivity, productivity, exploration, and novelty. Reproductive imagination is divided into five sub-dimensions: focus, effectiveness, transformation, crystallization, and dialectics. This research successfully developed and validated a Scientific Imagination Assessment (SIA) instrument specifically designed to measure students' scientific imagination in the context of abstract physics, specifically the Bohr atomic model. This instrument employs a structured, essay-based assessment scored on a 5-point Likert scale, enabling it to capture students' cognitive and representational processes in greater depth.

Keywords: CFA, scientific imagination assessment, bohr atomic model.

Article's DOI: https://doi.org/10.23960/jpmipa.v27i2.pp801-820

INTRODUCTION

Scientific imagination refers to the capacity to envision abstract concepts in a scientifically informed manner and translate them into tangible mental representations. This is an important part of how students learn about and develop their understanding of concepts that are not physically present in their experiences or environment. The ability to generate such images will enable students to generate innovative ideas, build on them, and shape their thinking processes. The requirement to develop these skills will be very similar to what is expected of a 21st-century university grad. As with many other forms of creativity, scientific imagination is also a subset of creative activities. Scientific imagination allows people to create images in their minds. It goes beyond being a simple form of creative imagination. It supports the development of theories or models of scientific knowledge, as well as the creation of mental models of how scientists think about and interact with science. More recently, scientists have begun to view imagination as a skill that should be taught and practiced (French, 2020); since then, there has also been an increase in emphasis on the importance of imagination in the development of scientific theories, model building, experimentation, and the use of metaphor in scientific language (Murphy, 2022).

The Bohr atomic model occupies a unique epistemological position in science, as it is a theoretical construct that bridges classical physics and quantum mechanics. This model does not directly represent empirical reality but rather serves as a transitional conceptual representation, laden with assumptions, idealizations, and limitations. Therefore, understanding the Bohr atomic model requires more than simply mastering formulas or formal definitions; students need to activate their scientific imagination to construct and interpret atomic representations that are abstract, unobservable, and often at odds with everyday intuition. From a philosophy-of-science perspective, this ability serves as an epistemic mechanism that enables individuals to construct scientific meaning from theoretical models.

At the higher-education level, the Bohr atomic model is often taught as part of a continuum of the development of atomic theory. However, the cognitive processes underlying students' understanding of the model are rarely explicitly addressed. Students often treat the model as a static scientific fact, rather than as a conceptual representation that can be analyzed, transformed, and criticized. This situation highlights a gap between the objectives of modern physics learning, which emphasize understanding the model and its limitations, and evaluation practices that still focus on the final result. Without an instrument capable of measuring scientific imagination, particularly in the context of the Bohr atomic model, lecturers lack an empirical basis for identifying how students construct mental images, reconstruct concepts, and relate the model to formal scientific principles.

Therefore, developing a scientific imagination instrument contextualized to the Bohr atomic model is an urgent need, both epistemologically and pedagogically. Epistemologically, this instrument positions the process of scientific imagination as a measurable, empirical object of study in physics education. Pedagogically, the instrument serves as a diagnostic tool to holistically map students' scientific imagination profiles, enabling the design of more adaptive and targeted learning strategies (Chuang et al., 2019a; Kim, 2025; Wang et al., 2015; Zamora-Polo et al., 2019). Thus, this instrument not only supports learning evaluation but also contributes to strengthening more meaningful and reflective abstract physics learning in higher education.

The pace of technological advancement in education has resulted in numerous advantages for our society and future societies. Technology can be a tool that helps address a wide range of issues across many arenas (e.g., work, leisure, and other contexts) by enabling more effective methods of managing these challenges. At the same time, as the use of technology increases, so too must our ability to develop the skills necessary to solve problems by utilizing technology as an aid, or else we run the risk of allowing ourselves to create more problems than we solve through technology innovation example, ICT, AI, AR, etc. (Arifin et al., 2025; Hidayat et al., 2017; Kim et al., 2007; Kousloglou et al., 2023; Kunnath & Botes, 2025). One challenge for many students is conveying their abstract way of thinking and creating models to help them learn; without the ability to see the larger context of a problem, it is difficult for them to generate their own abstract ideas and apply them to real-world problems. The use of scientific imagination can help to enhance students' creative and innovative ability, thereby providing the foundation for memorizing lessons and generating new ideas that lead to solving complex problems (Hadzigeorgiou, 2016)

Science does not rely solely on observation and experimentation; it also involves conceptual creation or imaginative modeling. If teachers are to evaluate students' cognitive processing patterns as they relate to their understanding of concepts or problems, they should gain insight into students' scientific imagination (Fontes, 2024). Scientific imagination acts as a bridge connecting scientific knowledge, creativity, and critical thinking (Sanjiartha et al., 2024). Therefore, assessing the ability to imagine scientifically offers an evaluative basis for developing instruments to measure this construct. Students' scientific imagination can be systematically evaluated through well-designed assessment tools. Most existing imagination instruments, such as the Imaginative Thinking Scale (ITS) by Lin & Tsau (2013), the Four-Factor Imagination Scale (FFIS) by Zabelina & Condon (2020), or the Hunter Imagination Questionnaire (HIQ) by Jung et al. (2016) are designed to measure imagination as a general trait across domains. These instruments emphasize fluency, originality, or the frequency of imaginative activity, but do not explicitly link imagination to the process of constructing scientific knowledge.

The Imaginative Thinking Scale (ITS) is designed to measure general imaginative thinking ability across dimensions of initiation, fluency, flexibility, and originality. ITS items generally require respondents to report the frequency of idea generation or their ability to generate a variety of ideas in everyday life or common creative tasks. A major limitation of the ITS in the context of the Bohr atomic model is its lack of requirements for scientific conceptual representation. ITS items do not ask respondents to imagine unobservable entities, integrate the laws of physics, or navigate the boundaries between theoretical models and physical reality. Thus, the ITS is more sensitive to the creativity of generic ideas than to scientific imagination, which operates within a rigorous theoretical framework such as atomic structure.

The Four Factor Imagination Scale (FFIS) evaluates imagination in terms of frequency, complexity, emotional valence, and directedness. Although the directedness dimension appears relevant, FFIS items focus on internal, affective, imaginative experiences, rather than on the transformation of abstract ideas into meaningful scientific representations. In the context of the Bohr atom, students' primary challenge is not simply to imagine, but to maintain the coherence of imagination with the postulates of energy quantization, electron trajectories, and the limitations of the classical quantum model. This aspect is not accommodated in the FFIS because the instrument does not tie imagination to the epistemic structure of science.

The Hunter Imagination Questionnaire (HIQ) links imagination to creative achievement and, in some studies, to neurological correlates. HIQ items tend to assess the trait of imagination as a stable personal characteristic, such as a tendency to fantasize or engagement in creative activities. This approach is problematic in the context of abstract physics because, in the Bohr atomic model, scientific imagination is situational, contextual, and dependent on prior conceptual mastery rather than solely on personal disposition. The HIQ also does not provide a mechanism for assessing the quality of the transformation of imagination into a logically justifiable scientific representation.

In contrast, Lian et al.'s (2012) Creative Imagination and Reproductive Imagination framework explicitly positions imagination as an epistemic mechanism in science, operating in both the generative phase (idea formation, initial images) and the reconstructive phase (transformation, formalization, and systematic reasoning). This structure is philosophically aligned with the character of Bohr's atomic model as an abstract theoretical model that cannot be directly observed and must be understood through mental representation and conceptual reasoning.

Instruments such as the Verbal SIT by Wang et al. (2017), the Figural SIT by Wang (2020), or the Digital Storytelling-Based Scientific Imagination (DSSI) by Chuang et al. (2019b) emphasize early creative processes such as brainstorming, association, and figurative visualization. While relevant for idea exploration, these instruments are less sensitive to later phases of scientific understanding, such as conceptual focus, idea effectiveness, or dialectical reasoning, which are crucial in university-level physics learning. Most alternative instruments have been validated using Rasch analysis, which is powerful for unidimensional measurement and item ranking, but less than optimal for testing complex latent factor structures. In this study, the primary objective was to confirm the theoretical two-factor structure of scientific imagination and the relationships between its indicators. Therefore, the factor analysis and Confirmatory Factor Analysis (CFA) approach used in Lian et al.'s (2012) framework was more appropriate.

This strengthens the argument that Lian et al.'s (2012) framework is the most appropriate choice for examining scientific imagination in the teaching of the Bohr atomic model in higher education. Lian et al.'s (2012) model excels because it clearly distinguishes between creative imagination (intuition, productivity, novelty) and reproductive imagination (focusing, effectiveness, transformation, crystallization, dialectics). This distinction allows lecturers and researchers to more precisely identify imbalances in students' imagination profiles, such as the phenomenon found in this study: high generative ability but low application ability and conceptual effectiveness. Thus, this instrument is not only evaluative, but also diagnostic and instructional. Lian et al.'s (2012) instrument was chosen because (1) it epistemologically positions imagination as a mechanism for the formation of scientific knowledge, (2) it is pedagogically capable of mapping the gap between idea generation and conceptual application in abstract physics, and (3) it is methodologically compatible with CFA to comprehensively test the latent factor structure of scientific imagination.

Evaluation is a systematic effort to determine the extent to which learning objectives and outcomes have been achieved. All assessment tools have strengths and limitations. Multiple-choice tests are generally less effective at measuring higher-order critical thinking, whereas essay-based assessments better capture advanced reasoning and reduce guesswork. Nevertheless, essay tests also present challenges, such as longer administration and scoring time, potential subjectivity, more complex scoring procedures, and limited item coverage (Kusumadani et al., 2017).

For these reasons, the evaluation of scientific imagination in this study employed an essay-based assessment scored using a 5-point Likert-scale analytic rubric, as this format can more faithfully represent students' cognitive patterns related to imagination in scientific contexts. The developed essay test is not computer-assisted, allowing students to respond with text, graphs, or drawings that reflect their scientific imagination. This instrument was designed specifically to evaluate students' ability in scientific imagination. This study is not simply a replication of Lian et al.'s (2012) scientific imagination framework, but rather presents conceptual and methodological novelties in how the construct is operationalized, tested, and interpreted in the abstract physics domain of higher education. The main novelties of this study lie in four intersecting aspects: the scientific context, the instrument's character, the target population, and the validation approach. Thus, this study aims to develop and validate a

Scientific Imagination Assesment Instruments. The research questions of this study are: Is the Scientific Imagination Assessment a construct-valid instrument for measuring students' scientific imagination on the Bohr atomic model?

METHOD

Participants

An extensive experimental trial involved 101 physics students from Universitas Negeri Semarang. The subjects were physics students who had completed modern physics material.

Research Design and Procedures

This study employed a modified Research and Development (R&D) model based on Gall et al. (2003), with adaptations mainly concerning the number of participants at each trial stage. The modified stages included preliminary research and information gathering; planning the design of the instruments; initial product development and preliminary testing; product revision; field testing; and operational revision, as illustrated in Figure 1.

The preliminary research and information-gathering stage was conducted to analyze research needs and to obtain an initial profile of students' scientific imagination abilities using an essay-based test instrument. As this population had taken a physics course and the Bohr atom serves as an example of a highly conceptual area in science, which therefore requires a great deal of imagination, it was chosen for this phase of the study. The Bohr atomic model offers a unique combination of high abstraction, strong visual representation, epistemological rigor, and pedagogical relevance, making it a particularly suitable context for examining the structure, dynamics, and limitations of scientific imagination among higher-education physics students. It is a theoretical construct that cannot be directly verified through empirical observation but is understood through symbolic representation, analogy, and mental visualization. Unlike other abstract concepts, such as Newton's laws or the electric field, the Bohr atomic model requires students to imagine microscopic entities that explicitly contradict sensory experience, making conceptual understanding highly dependent on the ability to construct and manipulate scientific mental images.

Instruments

In planning the instrumentation design, it was necessary to develop test indicators of students' ability to use scientific imagination, as outlined by Lian et al. (2012). Scientific imagination has two aspects: creative imagination and reproductive imagination; each type of imagination has five key indicators defined for the respective dimensions of creativity and reproduction, as follows: Creative Imagination: Intuition, Sensitivity, Productive, Exploratory, Novelty; and Reproductive Imagination: Focus, Effectiveness, Transformative, Crystallized, Dialectical. The first stage of product development led to the creation of a written test based on instructional content and expectations for students at the end of the course. The Scientific Imagination Assessment was developed as part of this product, and several people (validators) reviewed it for accuracy. Validators were chosen because they had diverse backgrounds and specialties related to physics education, as well as academic credentials and/or experience teaching physics. All feedback received from validators was analyzed, and any necessary changes to meet the validity requirements were made. Once the test was deemed valid, it could be administered in classrooms.

Instrument validation was conducted using expert judgment to ensure the appropriateness of the content, the clarity of the indicators, and the representativeness of the scientific imagination construct. Validators were selected purposively based on academic and professional criteria relevant to the research objectives. The validators were experts in physics education and science learning, holding doctoral or master's degrees in physics/science education, experience teaching physics at universities, and involvement in instrument development or physics learning research. This diversity of backgrounds was intended to ensure that the instrument was reviewed from conceptual, pedagogical, and practical perspectives. The validators were asked to rate each instrument item based on its relevance, clarity of wording, and appropriateness of the indicators to the scientific imagination construct, particularly in the context of abstract physics. Quantitative assessments were conducted using a Likert scale, and Aiken's V was then used to analyze the content validity of each item (Kania et al., 2024). Aiken's V was chosen because it is specifically designed to quantify the degree of expert agreement on the relevance of items for measuring a specific construct. Items with Aiken's V scores below the established criteria were reviewed and revised before being used in the pilot testing phase.

In addition to quantitative assessments, the validators also provided qualitative feedback in the form of written comments. In general, the experts' input covered three main aspects: (1) refining the item wording to make it more explicit and unambiguous, (2) adjusting the question context to align with the character of Bohr's atomic concept and avoid non-scientific interpretations, and (3) strengthening the link between the item and the intended scientific imagination indicators. These inputs were followed up by revising the item wording, clarifying the question stimulus, and realigning the indicators with the theoretical framework of scientific imagination. After all expert input was incorporated and the content validity met the required criteria, the instrument was deemed conceptually and empirically feasible for classroom pilot testing. This validation process ensured that the instrument was not only statistically valid but also theoretically robust and pedagogically relevant. Experts rated each item using a Likert scale (Azriyanti & Syafriani, 2023).

This research project employed a small-scale experimental field test with 15 physics students enrolled at Universitas Negeri Semarang (UNNES) who had learned about the Bohr model of an atom prior to the test, and asked these students (the test subjects) to take an essay test consisting of 10 questions within a time limit of 60 minutes, not using any form of electronic devices, with pencil and paper. Students provide answers based on their imagination about Bohr's atomic model and do not seek answers from the electronic equipment used. Pilot testing of the instrument included an analysis of the internal validity through reliability testing, discrimination of response options, and an estimate of the level of difficulty of each question to determine the viability of the instrument and to meet both the validity and reliability standards necessary for use as an assessment (Kusumadani et al., 2017)

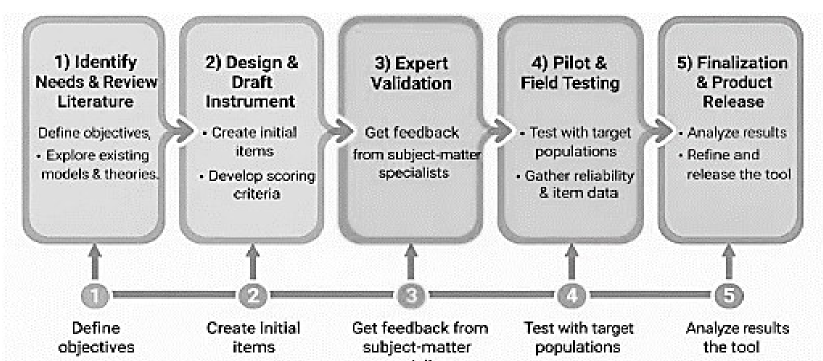

Figure 1. Scientific imagination assessment (SIA) instrument R&D process

It is important to conduct a reliability analysis to ensure that the assessment not only confirms that the tool has been developed correctly but that it produces consistent, reliable data across different applications. Instrument reliability was analyzed using an approach tailored to the data type. For essay data, inter-rater reliability was assessed using Cohen's kappa, which measures agreement beyond chance (Kolesnyk & Khairova, 2022). Meanwhile, for Likert-scale instruments, internal consistency reliability was assessed using Cronbach's Alpha, which measures inter-item integration in measuring the construct of scientific imagination (Setyaedhi, 2024). At this point, changes to the operational procedure were made based on the field test results to create a readily usable instrument for testing. Once the instrument was ready for use, it was administered to students in a physics class who had completed their studies of the Bohr atomic theory.

Data Analysis

All statistical analyses were conducted using JASP. Given that the instrument employed a 5-point Likert scale and preliminary analyses indicated minor deviations from multivariate normality, the CFA model was estimated using the robust maximum likelihood estimator (MLR). This estimation method was selected because it provides robust standard errors and fit indices that are less sensitive to violations of normality assumptions, making it appropriate for ordinal data with five or more response categories.

A two-factor measurement model was specified a priori, consisting of creative imagination and reproductive imagination, with each indicator loading on only its hypothesized latent construct. Model fit was evaluated using multiple goodness-of -fit indices, including the Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Goodness of Fit Index (GFI), Normed Fit Index (NFI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Model fit was considered acceptable when CFI, TLI, GFI, and NFI values were e 0.90, RMSEA values were d 0.08, and SRMR values were d 0.08 (Hair et al., 2019).

Construct validity was further assessed through standardized factor loadings, average variance extracted (AVE), and composite reliability (CR). Factor loadings of e 0.50, AVE values e 0.50, and CR values e 0.70 were interpreted as evidence of adequate convergent validity and internal consistency reliability (Hair et al., 2019).

The factors considered most important in deciding which measurement model to employ and how to distribute the indicators across the tests were the selection of a proper measurement model for factor analysis and the distribution of items across the indicators in the test. Confirmatory factor analysis (CFA) was used to assess the structural validity of the scientific imagination instrument used in this study (Rogers, 2024). The CFA elements included the latent variable, observed indicator variable, factor loadings, constructs, and measurement errors (Retnawati et al., 2018).

The instrument is called essay type because each item is presented in the form of an open-ended narrative prompt that asks students to explain, describe, or reflect on their scientific imagination process (e.g., how they imagined Bohr's atomic structure, how an initial idea was modified, or how an image was conceptually tested). Thus, students do not choose an answer but rather produce a written response that externalizes their cognitive and imaginative processes.

In this context, the term essay type refers to the response format (written narrative), not the scoring method. Although student responses are qualitative (essays), each response is then coded and scored using an analytical rubric based on a

5-point Likert scale. This scale represents the level of occurrence or quality of scientific imagination indicators, for example, from very low to very high. To enable essay responses to be analyzed using CFA, a rubric was developed based on indicators of scientific imagination (e.g., intuition, transformation, effectiveness). Then, trained evaluators scored each indicator on a 0–4 Likert scale. Inter-rater reliability tests (e.g., Cohen's Kappa) ensured consistency of scores. Scores are aggregated per indicator, so that each student has a numerical score for each latent construct. These scores are then used as observed variables in the CFA. Thus, the CFA is not conducted on raw essay text, but rather on indicator scores resulting from systematic coding. In the SIA, students do not assess themselves; scores are assigned based on the quality of their cognitive responses, and the assessment is conducted by an external evaluator rather than by the respondent. Therefore, referring to the instrument as a "Likert scale" without qualification is indeed inappropriate and needs to be revised to "essay-based assessment scored using a 5-point Likert analytic rubric."

RESULT AND DISCUSSION

Instruments Validity

An examination of content validity utilized Aiken's V coefficient. Validators were asked to rate all validation criteria on a scale from 0 (very inappropriate) to 4 (very appropriate), and an Aiken's V value of 0.80 or greater indicated that the items were acceptable for use without revisions (Retnawati et al., 2018). The average content validity calculations for the eight validation criteria rated by the five validators are shown in Table 1. The calculated Aiken's V value of 0.88 demonstrates that the instrument meets content validity standards.

Table 1. Validity of aiken's V content

Criterion Validity	V1	V2	V3	V4	V5	V'Aikens
Average content validity scale	3.50	3.88	3.50	3.38	3.38	0.88

Confirmatory Factor Analyses (CFA) were performed on the Scientific Imagination instrument using JASP version 0.95.4.0. Two high-level structures derived from the scientific imagination can be divided into creative and reproductive imagination (Lian et al., 2012). The creative structure contains intuition, sensibility, productivity, exploration, and novelty, while the reproductive structure contains the following elements: focusing and effectiveness, transformation, crystallization, and dialectical thinking. The instrument consists of five-point Likert items, with one item per indicator.

CFA was utilized to assess if the theoretical constructs of the Scientific Imagination instrument were consistent with empirical data from the field. The results showed that the original structure of the Scientific Imagination instrument was consistent with the data collected in the field, with excellent model fit being noted on both a Comparative Fit Index (CFI = 0.967) and a Goodness of Fit Index (GFI) being greater than 0.978 (both e" 0.90 were considered good model fit). The other fit statistics that matched the empirical data were Root Mean Square Error of Approximation (RMSEA = 0.060) and Standardized Root Mean Square Residual (SRMR = 0.051), both of which indicate excellent model fit (RMSEA d" 0.08; SRMR d" 0.08).

In conclusion, the findings imply that the empirical conditions of the scientific imagination construct are well represented by the Scientific Imagination scale; the construct structure is valid based on fit indices, such as CFI, TLI, RMSEA, SRMR, and GFI, and that the number of factors in the construct is appropriate for the scientific imagination scale. The relationships among items are consistent with the construct structure of the scientific imagination scale; thus, it is an appropriate tool for measuring scientific imagination (Rafsanjani, 2022). The feasibility and reliability of this instrument are supported by McDonald's Omega and Cronbach's alpha coefficients. (Hayes & Coutts, 2020).

Table 2. Frequentist scale reliability statistics

	Coefficient ω	Coefficient α
Creative Imagination	0.811	0.783
Reproduction Imagination	0.800	0.788
total	0.882	0.875

Table 2 presents confidence intervals for acceptable scale reliability. The ù coefficient (McDonald's omega) measures construct reliability within a factor model; ù e" 0.80 indicates a reliable instrument (Malkewitz et al., 2023). Correlation coefficients below 0.50 indicate poor reliability, 0.50–0.75 moderate, 0.75–0.90 good, and above 0.90 excellent reliability (Grgic et al., 2021). Other studies suggest coefficients above 0.60 reflect acceptable reliability, 0.61–0.80 good reliability, and above 0.81 very good reliability (Menescardi et al., 2022).

The á coefficient (Cronbach's alpha) served as a measure of internal consistency reliability based on inter-item correlations, where á e" 0.80 indicates items consistently construct the instrument (Setyaedhi, 2024). Higher á values reflect greater item consistency. Based on standardized factor loadings (SFL), all items showed SFL values greater than 0.50, indicating each item meaningfully contributed to the construct and supported the construct validity of the Scientific Imagination Test instrument.

After confirming its feasibility, validity, and reliability, the instrument was disseminated to measure students' abilities in scientific imagination in the context of Bohr's atomic theory. Mean scores for the ten indicators were calculated with reference to the scientific imagination indicators proposed by Lian et al. (2012). Student performance was evaluated using a five-point Likert scale.

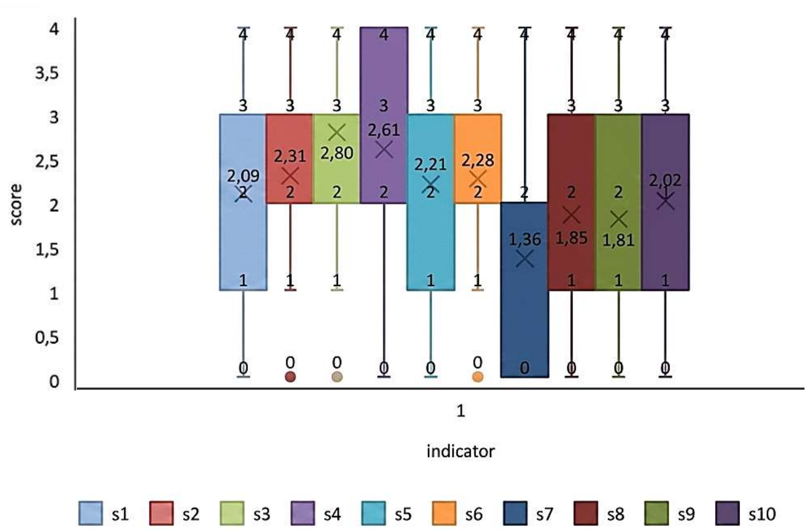

Figure 2. Scientific imagination ability profile

Figure 2 shows that the overall mean scientific imagination score was 2.13, which, based on the five-point Likert scale classification, falls within the moderate category across the ten items. Score ranges were classified as: very low (0-0.89), low (0.90–1.69), moderate (1.70–2.49), high (2.50–3.29), and very high (3.30–4.00). The average assessment of the scientific imagination indicator fell within the moderate category across 5 Likert-scale categories (0-4) based on the Scientific Imagination Assessment. The Creative imagination score of 2.40 (s1-s5) among students was higher than the reproductive imagination score of 1.80 (s6–s10), but the average of all indicators was still in the moderate category (2.31). It can be concluded that students find it easier to create the Bohr atomic model than the reproductive atomic model. Mean score s1 (2.09), s2 (2.31), s3 (2.80), s4 (2.61), s5 (2.21), s6 (2.28), s7 (1.36), s8 (1.85), s9(1.81) and s10(2.02). Based on mean scores per indicator, productivity, novelty, and intuition were categorized as high. In contrast, effectiveness was categorized as low. No indicators demonstrated scientific imagination ability in the high category.

The productivity (s3) and novelty (s4) indicators are in the high category, indicating that physics students are relatively capable of generating spontaneous, productive, and original ideas when confronted with the Bohr atomic concept. This achievement reflects the generative function of scientific imagination, namely the ability to form initial images and conceptual possibilities without the demands of strict scientific precision.

The low effectiveness score (s7) does not indicate a failure of scientific imagination but rather a gap between the imagination's generative and regulatory abilities. In the context of abstract physics, such as Bohr's atomic theory, this gap becomes even more understandable. Bohr's model requires students to abandon classical, intuitive images and replace them with quantum representations that are not directly visualizable. This condition allows students to remain creative and productive in generating ideas, but they experience difficulties when determining which ideas are most effective and scientifically sound. In other words, students' scientific imagination is more active at the exploratory stage than at the evaluative and application stages.

Confirmatory Factor Analysis

The Confirmatory Factor Analysis (CFA) model is presented in Figure 3. The model comprises two latent variables, Creative Imagination (cri), measured by five indicators, and Reproductive Imagination (rpi), also measured by five indicators. The latent variable Creative Imagination (cri) is represented by indicators s1, s2, s3, s4, and s5, while Reproductive Imagination (rpi) is represented by s6, s7, s8, s9, and s10. Each latent variable reflects the core principle of CFA as a theory-driven model rather than an exploratory approach.

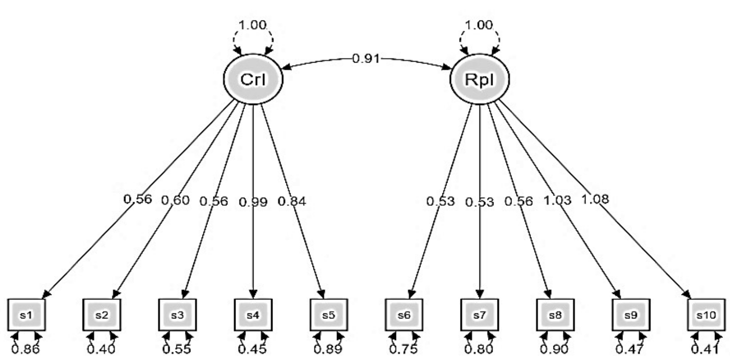

Figure 3. Model plot from the result of confirmatory factor analysis

The confirmation of the two-factor structure of scientific imagination through CFA analysis in this study has theoretical implications that go beyond mere instrument validity. The finding that scientific imagination is empirically structured into two main dimensions, creative imagination and reproductive imagination, suggests that scientific imagination is not a single, homogeneous ability, but rather a multi-layered cognitive construct with distinct yet interrelated functions. This finding provides a strong empirical foundation for the conceptual model previously formulated theoretically by Lian et al. (2012).

Table 3. show each indicator's contribution to its latent variable is reflected in standardized factor loading values. For Creative Imagination,

Table 3. Factor loading and factor covariances of indicators: scientific imagination

						95% Confidence Interval	
Factor	**Indicator**	**Std. estimate**	**Std. Error**	**z-value**	**p**	**Lower**	**Upper**
Creative Imagination	s1	0.519	0.094	5.516	< .001	0.335	0.703
	s2	0.688	0.056	12.371	< .001	0.579	0.797
	s3	0.604	0.059	10.283	< .001	0.489	0.720
	s4	0.829	0.037	22.257	< .001	0.756	0.902
	s5	0.664	0.074	8.953	< .001	0.519	0.810
Reproduction Imagination	s6	0.526	0.073	7.190	< .001	0.383	0.670
	s7	0.511	0.085	6.015	< .001	0.345	0.678
	s8	0.507	0.081	6.240	< .001	0.348	0.666
	s9	0.832	0.050	16.503	< .001	0.733	0.931
	s10	0.860	0.038	22.762	< .001	0.786	0.934

Factor Covariances

							95% Confidence Interval	
			Std. estimate	**Std. Error**	**z-value**	**p**	**Lower**	**Upper**
Creative Imagination	↔	Reproduction Imagination	0.913	0.043	21.20	< .001	0.829	0.998

factor loadings ranged from 0.519 to 0.664. Indicators s2 (0.688), s3 (0.604), s4 (0.829), and s5 (0.664) demonstrated relatively strong contributions. In contrast, indicator s1 showed a low factor loading (0.519), indicating weaker contribution and suggesting the need for further evaluation and revision. Indicator S1 represents spontaneous knowledge that emerges without formal analysis, but is rooted in scientific experience, conceptual mastery, and patterns stored in long-term memory. "When you first learned Bohr's atomic concept, what did Bohr's atom look like based on the atomic theory you had previously acquired?" This question is open-ended, reflective, and retrospective, requiring respondents to recall early cognitive experiences, access spontaneous mental images, and express images that are not necessarily stable or uniform across individuals. This characteristic differs fundamentally from other, more structured and productive indicators. Indicator S1 represents spontaneous imaginative images that emerge in the early stages of learning scientific concepts. Its low but significant loading indicates that initial intuition is personal and psychometrically unstable, so only a small portion of its variance is explained by the general SI construct.

Nevertheless, S1 remains conceptually important as an early phase in the formation of scientific imagination. S1 is not conceptually weak, but serves as an indicator of the transition from early intuitive imagination to structured creative imagination. Keeping S1 in the instrument broadens the scope of the construct, demonstrating the development of scientific imagination rather than simply its result. s1 can be positioned as a preliminary indicator or analyzed separately as intuition-based imagination. In classical physics, learners' intuition may, in principle, be shaped through classroom demonstrations of everyday phenomena. Therefore, educating intuition needs imagination (Foti et al., 2021).

For Reproductive Imagination, factor loadings ranged from 0.46 to 0.78. Indicators s6 (0.526), s7 (0.511), s8 (0.507), s9 (0.832), and s10 (0.860) showed strong representation. Indicator s8 (0.507) was at the minimum acceptable threshold, indicating a marginal but acceptable contribution. Indicator s8, representing imaginative transformation ability, showed a relatively lower factor loading, indicating that the conceptual transformation process is a complex cognitive stage and has not been evenly mastered by all respondents. This finding confirms that transformation is a key yet challenging component in the development of scientific imagination. The loading of 0.507 is within the acceptable minimum limit, indicating that imaginative transformation ability is not completely homogeneous among respondents. The variance in s8 responses is only partially explained by the construct of reproductive imagination. This process is high-level cognitive, requires mature conceptual understanding, and is strongly influenced by learning experiences and metacognitive strategies. Consequently, not all students are at the same level of transformation, resulting in greater variation in responses, reflected in lower loading values.

A factor loading value e" 0.50 indicates a strong contribution to the latent variable and meets convergent validity criteria. Values below each indicator represent residual variances, ranging from 0.261 to 0.743. Residual variance values from smallest to largest: s10 (0.261), s9 (0.308), s4 (0.313), s2 (0.526), s5 (0.559), s3 (0.635), s6 (0.723), s1 (0.731), s7 (0.739), and s8 (0.743). Residual variance d" 0.75 indicates adequate contribution and methodological acceptability. Values e" 0.75 suggest a need for revision.

Paired sample t-test results indicated that Creative Imagination scores were significantly higher than Reproductive Imagination scores at the individual student level (t H" 5.63–7.17, df = 82, p < .001), with a large effect size (Cohen's d = 0.70–0.99). This finding confirms that the dominance of Creative Imagination is not merely a descriptive difference but rather reflects a statistically significant difference in constructs. However, convergent validity analysis using Average Variance Extracted (AVE) showed that Creative Imagination had an AVE value (AVE = 0.5), compared to Reproductive Imagination (AVE = 0.5), both of which were still below the ideal threshold of 0.50.

This pattern indicates that although Creative Imagination emerged as a more prominent ability in quantitative terms, the variance explained by this latent construct across its indicators was relatively limited. In other words, the expression of students' creative imagination is heterogeneous and multidimensional, so that indicators such as intuition, productivity, and novelty do not converge strongly on a single homogeneous latent dimension. In contrast, Reproductive Imagination, despite having a lower mean score and a significantly different t-test result, shows better internal coherence, as reflected in a higher AVE.

This finding shows that the strength of performance (mean difference) and the strength of measurement (convergent validity) do not always go hand in hand. It emphasizes the complexity of scientific imagination in the context of abstract physics, such as the Bohr atomic model.

In general, it is still reasonable to measure Scientific Imagination using a cognitive-based instrument since latent variable constructs are complex and abstract. Both the CFA results support the instrument's theoretical structure and identify specific indicators that need refinement. The results of this study generally reinforce, rather than challenge, Lian et al.'s (2012) model of scientific imagination. Lian et al. proposed that scientific imagination develops through hierarchical stages, from intuition and creative exploration to transformation, crystallization, and dialectical integration. The two-factor structure confirmed in this study can be understood as an empirical condensation of these stages into two main functional clusters: the exploratory-generative stage (creative imagination) and the reconstructive-regulative stage (reproductive imagination).

However, this study also refines Lian et al.'s model by showing that these stages are not only sequential but also differ in their psychometric stability. Indicators of the early and transitional stages, such as intuition (s1) and transformation (s8), show weaker contributions than indicators of the later stages, such as crystallization and dialectical integration. This pattern does not contradict Lian et al.'s work, but rather provides empirical evidence that certain stages of scientific imagination are inherently more fragile and contextual.

Thus, these findings shift the scientific understanding of imagination from a mere developmental sequence to a dynamic structure with critical cognitive points, in which not all components develop or function equally. These findings align with Lian et al. (2012), who demonstrated creative and reproductive imagination as two interconnected dimensions of scientific imagination. Each dimension comprises five interrelated indicators that collectively contribute to the overall structure. Validation of the two-factor structure also implies that scientific imagination is best understood as a functional system, not simply a list of abilities. Creative imagination plays a role in generating initial representations, metaphors, and conceptual possibilities, while reproductive imagination functions to control, revise, and align these representations with formal scientific principles. This division provides a more operational theoretical framework than Lian et al.'s initial conceptual model, particularly for measurement and instructional design purposes.

These theoretical implications become even more significant when considered in the context of abstract physics, such as Bohr's atomic model. Abstract physics requires students to construct an understanding of entities that cannot be directly observed, thus placing a significant reliance on scientific imagination. In this context, the confirmed two-factor structure indicates that successful understanding of abstract concepts depends not only on the ability to imagine (creative imagination) but also on the ability to transform and regulate that imagination to align with formal scientific models (reproductive imagination). The relatively low contribution of the transformation indicator (s8) suggests that the transition from an intuitive image, for example, of the atom as a "miniature solar system," to an understanding of the quantum nature of Bohr's model is a major difficulty point in abstract physics. This finding extends Lian et al.'s model by demonstrating that the domain of physics, particularly modern physics, places a significant cognitive burden on the reproductive dimension of imagination. In physics studies, particularly the Bohr Atomic Model, students' cognitive understanding is a primary focus because it involves complex reasoning about electron trajectories, energy transitions, and atomic spectra. However, poor cognitive learning outcomes are often caused by learning methods that do not support the exploration and visualization of abstract concepts (Farida et al., 2025). This is in line with research by several researchers indicating that Bohr's atomic model is an abstract concept that requires further analysis.

The material on the development of atomic models by Dalton, Thomson, Rutherford, and Bohr, as well as on wave mechanics, is abstract and theoretical, requiring memorization (Sari et al., 2018). Abstract concepts are difficult to grasp in science because they require higher-order thinking, leading many students to fail to master them (Rahmawati, 2014). The material on the development of atomic models by Dalton, Thomson, Rutherford, and Bohr, as well as on wave mechanics, is abstract and theoretical, requiring memorization (Ilma & Lutfi, 2020). Abstract concepts are difficult to grasp in science because they require higher-order thinking skills, which can lead many students to fail to master them (Rachmawati, 2014). It was found that most students experienced difficulty understanding abstract concepts. Some students also experienced conceptual errors regarding atomic theory. Atomic theory is a difficult concept because it involves abstract concepts that are hard for most students to grasp (Sannah et al., 2015).

According to Anwar et al. (2015), there are several misconceptions regarding the topic of electron configuration in the Bohr atom, namely: 1. Students generally do not understand the basics of quantum theory, such as spectroscopy; it is only briefly explained. 2. Students mistakenly believe that line spectra represent atomic energy levels. However, this is not the case. 3. Students find it difficult to connect lines in emission and absorption spectra with electron transitions between energy levels. To address this, various learning methods and media have been developed and implemented. One approach implemented is the use of interactive simulation media. In their research, Ilma & Lutfi (2020) implemented PhET (Physics Education Technology) as a learning medium for atomic structure and the periodic table at Nahdlatul Ulama Vocational High School (SMK) Sugio Lamongan. The results showed that the use of PhET improved student learning outcomes and received a positive rating. Furthermore, student activity during learning was highly active, and positive student responses reached a high percentage.

CONCLUSION

The results of the Confirmatory Factor Analysis (CFA) indicate that the scientific imagination instrument has excellent model fit with the empirical data. This is demonstrated by the Comparative Fit Index (CFI = 0.967) and Tucker-Lewis Index (TLI = 0.957), which exceed the recommended fit limits, indicating that the proposed two-factor structure accurately and efficiently represents the relationships among the indicators. Furthermore, the Goodness-of-Fit Index (GFI = 0.978) indicates that the model accounts for most of the variance and covariance in the data, strongly supporting the instrument's construct validity.

From an absolute fit perspective, the Root Mean Square Error of Approximation (RMSEA = 0.060) and Standardized Root Mean Square Residual (SRMR = 0.051) values indicate a low level of model error and fall within the acceptable range. These findings indicate that the model specification is not only statistically appropriate but also stable and potentially generalizable to science learning contexts, particularly in abstract physics. Overall, the combination of model fit indices confirms that the developed scientific imagination instrument meets strong psychometric standards and is suitable for further analysis and the development of scientific imagination theory.

According to the results of these fit indices, the methods confirm that the construct structure is accurate, the number of factors is appropriate, and the item relationships are specified correctly. The instrument's appropriateness has been demonstrated. Additionally, McDonald's omega and Cronbach's alpha support the instrument's feasibility and reliability. Scientific imagination is a general structure that is consistent with Lian et al. (2012), but the distribution of indicator strengths is influenced by the epistemic characteristics of the scientific discipline. Imaginative transformation is a conceptual bottleneck in abstract physics learning, not simply an additional skill. Lian et al.'s (2012) model should be understood not simply as a linear developmental model, but as a hierarchical functional framework that can be empirically tested and refined. Thus, this study not only validates the scientific imagination instrument but also contributes to the theoretical reconstruction of the concept of scientific imagination in the context of abstract science learning. These findings pave the way for the development of a more domain-sensitive and psychometrically operationalized theory of scientific imagination.

Standardized covariance value: $r = 0.913$ ($p < 0.001$), which indicates a strong relationship between creative and reproductive imagination. However, it does not indicate extreme multicollinearity. This supports the theoretical view that the two types of imagination are related but functionally distinct. The results of this study indicate that physics students' scientific imagination skills in understanding Bohr's atomic theory did not develop evenly across all indicators. Although the average total score on the Scientific Imagination Test fell within the moderate range, a more detailed analysis revealed disparities across aspects of scientific imagination. Students tended to perform better on indicators related to the consolidation and reproduction of formal scientific models. In contrast, aspects requiring the formation of initial images, conceptual exploration, and representational transformation remained relatively weak.

These findings indicate that students' scientific imagination in the context of abstract physics functions more as a tool for knowledge reproduction than as a means of conceptual exploration and reconstruction. This disparity confirms that understanding modern physics concepts is determined not only by mastery of formal models but also by the ability to transform intuitive representations into more abstract and scientific conceptual structures. Therefore, the conclusions of this study emphasize the importance of developing learning strategies that explicitly target weaknesses in imaginative thinking, particularly in the early and transitional stages of scientific imagination, to support a deeper, more meaningful understanding of physics.

LIMITATION

This study has several limitations that need to be considered when interpreting the findings. First, the validation of the scientific imagination instrument was conducted using a single context: abstract physics material, namely, Bohr's atomic theory. Although this context is relevant for studying scientific imagination, the resulting indicator structure and profile may not fully represent the characteristics of scientific imagination in other physics topics with different epistemic characteristics, such as classical mechanics or electromagnetism. Second, the analysis used in this study was limited to a Confirmatory Factor Analysis (CFA) approach within the Classical Test Theory framework. This approach focuses on latent relationships among indicators but does not explore item characteristics in greater detail, such as difficulty level and item discrimination power across various ability levels. As a result, variations in student performance on certain indicators, especially those indicating low achievement, cannot be fully explained at the item level. Third, the research data were collected cross-sectionally, making it impossible to track the longitudinal development of students' scientific imagination. Therefore, findings showing uneven ability profiles cannot be interpreted as developmental stages, but rather as a snapshot of ability at a specific point in time. Fourth, the instrument relied on students' written and reflective responses, which may be influenced by verbal skills and answering strategies rather than solely by scientific imagination. This opens up the possibility of response bias that was not fully controlled for in this research design.

Fifth, although the measurement model demonstrated excellent overall fit, this study identified specific anomalies at the indicator level that constitute important limitations. Specifically, indicator s1 (intuition) showed a relatively high mean value and factor loading compared to other indicators, including those within the reproductive imagination dimension. This dominance of s1 potentially reflects the item salience effect, namely the tendency for students to respond highly to statements that are intuitive, spontaneous, and close to their early experiences in learning physics, regardless of the depth of the more complex construct of scientific imagination. Consequently, s1 may better represent familiar cognitive access than scientific intuition in a more rigorous theoretical sense. Furthermore, there was an imbalance in the contribution of indicators within the reproductive imagination construct, particularly indicator s7 (effectiveness), which showed the lowest descriptive value and a weaker contribution to the model. This imbalance indicates that students' ability to transform imaginative ideas into effective and purposeful solutions has not developed commensurate with their initial intuitive and creative abilities. However, the current study design does not allow for a clear distinction between these findings, whether these findings reflect the characteristics of the construct itself, biases in essay assignments, or students' limited pedagogical experience in systematically solving abstract physics problems.

Another technical limitation is CFA's inability to explain the sources of anomalies at the cognitive process level. CFA confirms the appropriateness of the factor structure. However, it does not provide information about how students reason, imagine, and externalize their scientific imagination while working on the items, especially on items like S1, which rely heavily on spontaneous responses. Without supporting qualitative data (e.g., think-aloud protocols or analyses of essay responses), interpretations of the strengths or weaknesses of particular indicators remain inferential. These limitations indicate that, while the results of this study provide important empirical and theoretical contributions, the findings need to be understood within the context of the study's scope and design.

FUTURE WORK

Based on the study's findings and limitations, several suggestions can be put forward for further research. First, future research is recommended to test the scientific imagination instrument across various physics topics at varying levels of abstraction, such as quantum mechanics, electromagnetism, and relativity. This approach is important for assessing the extent to which the two-factor structure of scientific imagination is consistent across domains or sensitive to the epistemic characteristics of each topic. Second, future research should consider using an Item Response Theory (IRT) approach as a complement to CFA. IRT analysis can provide more detailed information about the characteristics of each item, including the item's difficulty level and discriminatory power across different levels of scientific imagination ability.

Thus, indicators indicating low achievement can be analyzed more deeply to determine whether these weaknesses stem from item characteristics or from students' cognitive limitations. Third, a longitudinal study is recommended to track the development of students' scientific imagination over time. This approach will allow for a more comprehensive understanding of how aspects of scientific imagination develop, transition, and interact during the physics learning process, particularly with abstract concepts. Fourth, future research could combine test instruments with qualitative methods, such as in-depth interviews, think-aloud protocols, or analysis of students' visual representations. This mixed approach is expected to capture imaginative processes not fully revealed through written responses, thus providing a more holistic picture of scientific imagination. Fifth, conduct a Latent Profile Analysis (LPA) using 10 indicator scores to identify student subgroups with distinct imagination profiles. Sixth, suggestions for increasing the number of participants and/or adding more questionnaire items for each dimension are recommended. Questionnaire items should be sufficiently rigid, with a participant count of around 250-300, to achieve better construct validity. Overall, these suggestions are expected to expand the empirical validity of scientific imagination instruments and deepen theoretical understanding of the role of scientific imagination in abstract physics learning.

■ DECLARATION OF GENERATIVE AI USAGE IN THE WRITING PROCESS

During the writing of this manuscript, the author used Google Translate to assist with language refinement and correction and ChatGPT to create the diagrams. The author has reviewed and edited the content generated by these tools and takes full responsibility for the content of the published article.

■ REFERENCES

Anwar, B., Ernawati, Rd. R. S., Setiadi, R., & Wiji, M. (2015). *Pengembangan representasi kimia sekolah berbasis intertekstual pada subkonsep konfigurasi elektron model atom bohr yang diperluas dalam bentuk multimedia* [Development of intertextual based school chemistry representation on the sub-concept of electron configuration of the extended bohr atomic model in multimedia form]. *Jurnal Pengajaran Matematika dan Ilmu Pengetahuan Alam*, *17*(2), 278. https://doi.org/10.18269/jpmipa.v17i2.272

Arifin, Z., Saputro, S., & Nicholson, H. B. (2025). Technology enhanced inquiry based science learning with cultural integration for secondary education. *Proceeding - Int. Conf. Creat. Commun. Innov. Technol.: Empower. Transform. MATURE LEADERSHIP: Harnessing Technol. Adv. Glob. Sustain., ICCIT*. Scopus. Proceeding - 2025 4th International Conference on Creative Communication and Innovative Technology: Empowering Transformative MATURE LEADERSHIP: Harnessing Technological Advancement for Global Sustainability, ICCIT 2025. https://doi.org/10.1109/ICCIT65724.2025.11167118

Azriyanti, R. & Syafriani. (2023). Validation of the physics e-module based on problem based learning as independent teaching material to improve critical thinking skills of class xi high school students. *Jurnal Penelitian Pendidikan IPA*, *9*(11), 10223–10229. https://doi.org/10.29303/jppipa.v9i11.5809

Chuang, H.-H., Hsieh, M.-H., Cheng, Y.-Y., &

Wang, C.-C. (2019a). An instrument for assessing the development of scientific imagination via digital storytelling for elementary school students. *Creativity Research Journal*, *31*(4), 408–418. Scopus. https://doi.org/10.1080/10400419.2019.1670025

Chuang, H.-H., Hsieh, M.-H., Cheng, Y.-Y., & Wang, C.-C. (2019b). An instrument for assessing the development of scientific imagination via digital storytelling for elementary school students. *Creativity Research Journal*, *31*(4), 408–418. Scopus. https://doi.org/10.1080/10400419.2019.1670025

Farida, D., Arifin, R. A. S., Br Surbakti, S. B., & Oktavia, V. A. (2025). *Visualisasi perkembangan pengetahuan kognitif mahasiswa pada model atom bohr menggunakan microsoft excel.* [Visualization of the development of students' cognitive knowledge on the bohr atomic model using microsoft excel]. *VISA: Journal of Vision and Ideas*, *5*(1). https://doi.org/10.47467/visa.v5i1.5965

Foti, C., Anttila, D., Maniscalco, S., & Chiofalo, M. L. (2021). Quantum physics literacy aimed at k12 and the general public. *Universe*, *7*(4). Scopus. https://doi.org/10.3390/universe7040086

French, S. (2020). Imagination in scientific practice. *European Journal for Philosophy of Science*, *10*(3), 27. https://doi.org/10.1007/s13194-020-00291-z

Gall, M., Borg, W., & Gall, J. (2003). Educational research: an introduction. *British Journal of Educational Studies*, *32*. https://doi.org/10.2307/3121583

Hadzigeorgiou, Y. (2016). Imaginative science education: The central role of imagination in science education. In *Imagin. Sci. Educ.: The Cent. Role of Imagin. In Sci. Educ.* (p. 283). Springer International Publishing. Scopus. https://doi.org/10.1007/978-3-319-29526-8

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (Eighth edition). Cengage.

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than cronbach's alpha for estimating reliability. but…. *Communication Methods and Measures*, *14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Hidayat, S. R., Setyadin, A. H., Hermawan, H., Kaniawati, I., Suhendi, E., Siahaan, P., & Samsudin, A. (2017). *Pengembangan instrumen tes keterampilan pemecahan masalah pada materi getaran, gelombang, dan bunyi* [Development of problem solving skills test instruments on vibrations, waves, and sound material]. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *3*(2), 157–166. https://doi.org/10.21009/1.03206

Ilma, K., & Lutfi, A. (2020). *Penerapan phet sebagai media pembelajaran struktur atom dan sistem periodik di smk nahdlatul ulama sugio lamongan* [Development of problem solving skills test instruments on vibrations, waves, and sound material.. *UNESA Journal of Chemical Education*, *9*, 309–316. https://doi.org/10.26740/ujced.v9n3.p309-316

Jung, R. E., Flores, R. A., & Hunter, D. (2016). A new measure of imagination ability: Anatomical brain imaging correlates. *Frontiers in Psychology*, *7*(APR). Scopus. https://doi.org/10.3389/fpsyg.2016.00496

Kania, N., Kusumah, Y. S., Dahlan, J. A., Nurlaelah, E., Gürbüz, F., & Bonyah, E. (2024). Constructing and providing content validity evidence through the Aiken's V index based on the experts' judgments of the instrument to measure mathematical problem-solving skills. *REID (Research and Evaluation in Education)*, *10*(1), 64–79. https://doi.org/10.21831/reid.v10i1.71032

Kim, M. C., Hannafin, M. J., & Bryan, L. A. (2007). Technology-enhanced inquiry tools in science education: An emerging pedagogical framework for classroom practice. *Science Education*, *91*(6), 1010–1030. Scopus. https://doi.org/10.1002/sce.20219

Kim, Y. (2025). "Sacredly confidential": medical advice columns, consumerism, and literary imagination in turn-of-the-century women's magazine culture. *Configurations*, *33*(2),151–182. Scopus. https://doi.org/10.1353/con.2025.a965609

Kousloglou, M., Petridou, E., Molohidis, A., & Hatzikraniotis, E. (2023). Assessing students' awareness of 4cs skills after mobile-technology-supported inquiry-based learning. *Sustainability (Switzerland)*, *15*(8). Scopus. https://doi.org/10.3390/su15086725

Kunnath, A. J., & Botes, W. (2025). Transforming science education with artificial intelligence: Enhancing inquiry-based learning and critical thinking in South African science classrooms. *Eurasia Journal of Mathematics, Science and Technology Education*, *21*(6). Scopus. https://doi.org/10.29333/ejmste/16532

Kusumadani, A. I., Syarif, A., & Dewi, R. C. (2017). *Pengembangan tes uraian interaktif pada mata kuliah biologi lingkungan berbantuan microsoft powerpoint dan visual basic for aplication. 10* [Development of interactive essay tests in environmental biology courses using microsoft powerpoint and visual basic for applications. 10].

Lin, H.-H., & Tsau, S.-Y. (2013). The development of an imaginative thinking scale. *Imagination, Cognition and Personality*, *32*(3), 207–238. Scopus. https://doi.org/10.2190/IC.32.3.b

Menescardi, C., Villarrasa-Sapiña, I., Lander, N., & Estevan, I. (2022). Canadian agility movement skill assessment (CAMSA) in a spanish context: evidences of reliability and validity. *Measurement in Physical Education and Exercise Science*, *26*(3), 245–255. https://doi.org/10.1080/1091367X.2021.2020794

Murphy, A. (2022). Imagination in science. *Philosophy Compass*, *17*(6). Scopus. https://doi.org/10.1111/phc3.12836

Rahmawati, L. (2014). *Pengembangan dan penerapan instrumen diagnostik two tier dalam mengidentifikasi miskonsepsi siswa tentang atom dan molekul* [Development and application of two tier diagnostic instruments in identifying students' misconceptions about atoms and molecules]. *Edusentris*, *1*, 146. https://doi.org/10.17509/edusentris.v1i2.141

Retnawati, H., Djidu, H., Kartianom, K., Apino, E., & Anazifa, R. D. (2018). Teachers knowledge about higher order thinking skills and its learning strategy. *Problems of Education in the 21st Century*, *76*(2), 215–230. https://doi.org/10.33225/pec/18.76.215

Rogers, P. (2024). Best practices for your confirmatory factor analysis: A JASP and lavaan tutorial. *Behavior Research Methods*, *56*(7), 6634–6654. https://doi.org/10.3758/s13428-024-02375-7

Sannah, I. N., Kadaritna, N., & Tania, L. (2015). *Pengembangan LKS dengan model discovery learning pada materi teori atom bohr* [Development of Student worksheets with the discovery learning model on bohr's atomic theory material].

Setyaedhi, H. S. (2024). Comparative test of cronbach's alpha reliability coefficient, kr-20, kr-21, and split-half method. *Journal of Education Research and Evaluation*, *8*(1), 47–57. https://doi.org/10.23887/jere.v8i1.68164

Wang, C. (2020). Validation of the scientific imagination test-figural. In *Bulletin of Educational Psychology* (Vol. 51, Issue 3, pp. 341–367). National Taiwan Normal University. https://doi.org/10.6251/BEP.202003_51(3).0001

Wang, C., Niemi, H. M., Cheng, C., & Cheng, Y. (2017). Validation of learning progression in scientific imagination using data from Taiwanese and Finnish elementary school students. In *Thinking Skills and Creativity* (Vol. 24, pp. 73–85). Elsevier Ltd. https://doi.org/10.1016/j.tsc.2017.02.014

Wang, C.-C., Ho, H.-C., & Cheng, Y.-Y. (2015). Building a learning progression for scientific imagination: A measurement approach. *Thinking Skills and Creativity*, *17*, 1–14. Scopus. https://doi.org/10.1016/j.tsc.2015.02.001

Zabelina, D. L., & Condon, D. M. (2020). The four-factor imagination scale (FFIS): A measure for assessing frequency, complexity, emotional valence, and directedness of imagination. *Psychological Research*, *84*(8), 2287–2299. Scopus. https://doi.org/10.1007/s00426-019-01227-w

Zamora-Polo, F., Corrales-Serrano, M., Sánchez-Martín, J., & Espejo-Antúnez, L. (2019). Nonscientific university students training in general science using an active-learning merged pedagogy: gamification in a flipped classroom. *Education Sciences*, *9*(4), 297. https://doi.org/10.3390/educsci9040297