Jurnal Pembelajaran IPA Terpadu: PELITA 6(2), July - December 2026

Development and Feasibility Testing of a Project-Based Learning (PjBL)-Based Electronic Portfolio Assessment to Measure Creative Thinking Skills in Chemical Bonding

Afidatul Mutmainah [1], Nahadi [2*], Triannisa Rahmawati [3], Wiwi Siswaningsih [4]

[1, 2, 3, 4] Universitas Pendidikan Indonesia, Indonesia

* nahadi@upi.edu

Abstract

The urgency of this study lies in the need for an innovative and feasible Project-Based Learning (PjBL)-based electronic portfolio assessment that can authentically measure students' creative thinking skills in understanding the abstract and complex concepts of chemical bonding. Creative thinking is a core dimension of Higher Order Thinking Skills (HOTS) that remains difficult to cultivate in chemistry instruction, particularly on the abstract topic of chemical bonding, where conventional written assessments have yet to accommodate authentic, process-based evaluation of students' creative thinking. This study employed a Research and Development (R&D) approach based on the 4D model, limited to the define, design, and develop stages, to develop and test the feasibility (validity and reliability) of a Project-Based Learning (PjBL)-based electronic portfolio assessment instrument for evaluating students' creative thinking skills in chemical bonding. Content validity was examined through expert judgment involving five validators, comprising three chemistry education lecturers and two chemistry teachers, using the Content Validity Ratio (CVR), while inter-rater reliability was examined through the Intraclass Correlation Coefficient (ICC) Two-Way Mixed-Effects Model of Consistency involving five raters who scored the work of 19 students selected through total sampling. The results show that the PjBL-based electronic portfolio assessment instrument possesses highly adequate quality, with a perfect CVR value of 1.00 (Valid) for all task items. Furthermore, the inter-rater reliability test showed consistent stability, with ICC values ranging from 0.949-0.981 (excellent) for Task 1 (Open-Ended Essay), 0.866-0.942 (good to excellent) for Task 2 (Concept Map), and 0.850-0.936 (good to excellent) for Task 3 (Digital Poster). The resulting instrument offers teachers a systematic and objective alternative for assessing creative thinking through project-based tasks, encourages students' active engagement and self-reflection through digital portfolios, and provides a validated reference for future development of technology-based authentic assessment in science education.

Keywords: *Creative Thinking, Chemical Bonding, Electronic Portfolio, Validity and Reliability*

Introduction

Chemistry instruction at the senior high school level plays a crucial role in laying the foundation of students' scientific knowledge, particularly for abstract topics. Through this subject, students are guided to understand the phenomena of matter, microscopic concepts, and logical, systematic patterns of reasoning. Chemistry is not merely regarded as the mastery of formulas or rote cognitive theory, but rather as a strategic medium for training higher order thinking skills (HOTS), problem solving, and scientific literacy (Huda et al., 2025). The characteristics of the modern curriculum demand that science learning take place actively, with students directly involved in constructing understanding through contextual and meaningful learning experiences. Within this learning ecosystem, assessment becomes an integral

https://doi.org/10.54065/pelita.6.2.2026.1700

component for measuring the attainment of established instructional objectives. Assessment is a systematic process encompassing the collection, analysis, and interpretation of information to inform educational decision-making, serving both an evaluative and a diagnostic function as a means of providing feedback.

In line with these developments, the Merdeka Curriculum directs learning toward a student-centered orientation with a stronger emphasis on creative thinking skills. Grounded in constructivist theory, students' knowledge is formed through direct experience and active interaction with their cognitive environment, allowing them to process information through inquiry and independent reconstruction of ideas. Creative thinking, as part of the HOTS dimension, plays an important role in helping students understand concepts in depth while formulating innovative solutions to problems. Creative thinking is the ability to generate various alternative solutions to a problem, centered on four key indicators: fluency, flexibility, originality, and elaboration. This cognitive capacity is rooted in Williams' theory, which asserts that creativity is not a static innate potential but a cognitive capacity that can be nurtured and developed through structured learning scenarios.

Although fostering creativity has become a mandatory outcome under curriculum regulations, its implementation in the field still faces obstacles due to the abstract nature of chemistry content, chemical bonding being a prime example. This topic requires students to integrate basic concepts, differentiate the characteristics of substances, and connect the three levels of chemical representation (macroscopic, microscopic, and symbolic) with everyday phenomena. Empirical evidence in the field, however, shows that students frequently experience significant difficulty in distinguishing the characteristics of ionic and covalent bonds and relating them to the physical properties of compounds the previous studied reported the occurrence of misconceptions that leave students' understanding of chemical bonding shallow (Hendrawani, 2023). This phenomenon is reinforced by the findings of previous studied regarding misconceptions in the concept of covalent bonding (Istiqomah et al., 2021). And further supported by the another studies which confirm that difficulties in understanding chemical bonding remain widespread among students (Warsito et al., 2020; Zulkhairi, 2022).

This weak conceptual understanding and limited creative thinking skill are further compounded by a school evaluation system still dominated by conventional methods. The assessments applied by teachers generally rely on written tests that measure only cognitive end results (assessment of learning), without being balanced by authentic assessment instruments capable of comprehensively capturing students' creative thinking processes and portfolios of work. Chemistry instruction consequently becomes trapped in expository methods or theoretical drilling that emphasizes rote memorization. Yet stress the importance of shifting the function of evaluation toward assessment for learning, serving as a formative feedback instrument throughout the learning process, as well as assessment as learning, which trains students' capacity for independent self-reflection (Rutherford et al., 2025). Further add that school assessment should not merely pursue a summative-administrative function to produce final grades, but should instead be directed toward formative assessment that supports the ongoing development of students' individual competencies (Ramasari et al., 2024).

This problem of conventional evaluation was confirmed by a preliminary study consisting of interviews with a chemistry teacher at a public senior high school in Bandung. The teacher confirmed that portfolio assessment methods had never been implemented in chemistry instruction. The use of digital technology thus far has been limited to facilitating cognitive tests via the Google Forms platform to obtain instant scores, as well as using Google Drive to collect independent assignments. This use of digital platforms was judged not yet to touch the essence of systematically organized authentic assessment management. The assessment patterns consistently used remain limited to Daily Tests and Semester Examinations to measure final achievement, so that evaluation of the learning process is neglected, ultimately contributing to students' low creative thinking ability.

As an effort to bridge the gap between the curriculum's ideal demands and classroom reality, the Project-Based Learning (PjBL) model can serve as a relevant pedagogical foundation, as it actively engages students in collaboratively designing and completing meaningful projects (Chang et al., 2024). The PjBL syntax proceeds sequentially through determining a driving question, designing the project steps, creating a schedule, implementation and monitoring, testing the product outcome, and finally evaluating the experience (Pratiwi & Ikhsan, 2024). Through this model, students can visualize the abstract concept of chemical bonding into a tangible product that fully connects the three levels of representation of matter (Syahidah & Rusmini, 2025; Situmorang et al., 2022). Project design activities require students to generate new ideas and original solutions that foster the growth of creative thinking skills (Ernawati et al., 2023; Pratiwi & Ikhsan, 2024). Furthermore, the PjBL model aligns closely with the principle of authentic assessment, as evaluation takes place continuously throughout the project process rather than focusing solely on the final outcome (Lestari et al., 2024; Chang et al., 2024).

However, the implementation of authentic assessments such as portfolios in schools today remains limited and largely manual, paper-based practice, making it impractical and failing to take advantage of digital technology (Tur et al., 2019). These limitations of physical formats and manual feedback constraints can be resolved in the digital era through the use of electronic portfolios, or e-portfolios. An e-portfolio is a collection of digital documents and artifacts that continuously record students' learning process, reflection, achievement, and competency development (Yang et al., 2016). As a modern assessment method, the e-portfolio functions as a dynamic medium for providing ongoing feedback rather than merely a static compilation of products (Zhang & Tur, 2023; Ariningsih et al., 2021; Aini et al., 2024). The application of e-portfolios has proven efficient in data storage capacity, facilitates independent reflection through digital logs, enhances online feedback interaction, and presents evidence-based assessment holistically (Mudau, 2022).

The successful implementation of e-portfolios will be more optimal when integrated with a Learning Management System (LMS) already familiar within the school ecosystem, such as Google Classroom. This use of IT-based e-portfolieo media aligns with (Garcia, 2025), Likewise, Project-based digital learning tools can effectively support the improvement of students' creative thinking skills in chemistry learning (Pane & Simamora, 2025). In addition, research by previous studied demonstrates that combining the PjBL model with portfolio assessment offers an adaptive alternative for simultaneously measuring learning process and outcomes (Masruroh et al., 2024; Haryati et al., 2024; Sari et al., 2023). Global literature further supports this view highlights the importance of authentic assessment in measuring the ability to apply knowledge to real-world contexts, while previous studied show that the use of electronic portfolio assessment in chemistry learning can enhance students' creative thinking skills as well as learning outcomes (Edwards et al., 2025).

In order to objectively measure the creative thinking indicators proposed within PjBL-based chemical bonding instruction, this e-portfolio instrument was designed around three interrelated task types. The first task consists of open-ended essay questions based on contextual problems intended to spark students' critical reasoning in solving real-world issues. The second task involves constructing a concept map based on the principles of previous studied as a hierarchical visual representation showing the relationships among sub-concepts through propositions, thereby contributing to the development of flexibility and elaboration in creative thinking (Novak & Cañas, 2006; Rahmawati et al., 2019). The third task is a Digital Poster serving as the final product, integrating text and visual elements to effectively communicate the concept of chemical bonding to readers through an engaging and easily understood scientific representation (Pedwell et al., 2017).

The poster assessment rubric in this study was developed based on widely accepted principles of effective scientific poster presentation, including the clarity of the title, organization of content, visual design, and ease of communication. Furthermore, the assessment indicators were formulated based on the principles of analytic rubrics so that each aspect can be evaluated objectively, systematically, and consistently across raters (Gosselin & Golick, 2020). The novelty of this study lies in the development of a Google Classroom-integrated electronic portfolio assessment instrument for the PjBL model, whose feasibility quality is tested comprehensively. Testing was carried out through content validation by experts using the Content Validity Ratio (CVR) coefficient, as well as an inter-rater reliability stability test using Intraclass Correlation Coefficient (ICC) analysis on the results of the three types of student creativity tasks (Ngui et al., 2022). This step offers education practitioners an alternative for capturing students' authentic creative-thinking profiles digitally, objectively, and systematically. Based on this background, the present study aims to develop and test the feasibility of a Project-Based Learning (PjBL)-based electronic portfolio assessment instrument to evaluate students' creative thinking skills in the topic of chemical bonding.

Method

The development of the evaluation instrument in this study applied the Research and Development (R&D) method ((Sumarni, 2019). The product developed was an electronic portfolio (e-portfolio) assessment based on Project-Based Learning (PjBL) via the Google Classroom platform, intended to measure students' creative thinking skills in the topic of chemical bonding. The instrument modification process referred to the 4D model, which was limited to the third stage, namely develop. This study was limited to the develop stage and did not proceed to the disseminate stage for three principal reasons. First, the primary objective of this study was to establish the initial feasibility of the instrument namely its content validity and inter-rater reliability which must be firmly demonstrated through expert validation and limited trial before the instrument can be responsibly disseminated on a wider scale (Thiagarajan, 1974). Disseminating an instrument whose measurement quality has not yet been empirically confirmed risks generating misleading assessment data when applied across broader, more heterogeneous settings. Second, dissemination constitutes a substantively different type of study, requiring trials across multiple schools and regions with more diverse student and teacher characteristics; this research was intentionally scoped to focus on establishing instrument feasibility, consistent with common practice in instrument-development research, where validation and large-scale implementation are treated as distinct, sequential lines of inquiry. Third, the develop stage already yields the core contribution targeted by this research, namely a validated and reliable e-portfolio assessment instrument ready for use at the classroom level. The dissemination stage is therefore recommended as a direction for future research, particularly to examine the instrument's generalizability across a wider range of schools, chemistry topics, and student populations. The procedure in this study began with the define stage, followed by the design stage, and concluded with the develop stage, each of which is elaborated descriptively below.

The define stage was carried out through needs analysis, an analysis of the senior high school chemistry curriculum, and an analysis of student characteristics to map the problems in assessing creative thinking within chemical bonding instruction. The results of this stage served as the reference for determining the scope of the material and the assessment objectives. In the design stage, an initial e-portfolio draft was prepared on Google Classroom, along with a test blueprint, the design of three task types, and the development of an analytic assessment rubric. This rubric was developed based on the creative-thinking indicators. which include fluency, flexibility, originality, and elaboration. The tasks designed consisted of a contextual open-ended essay (Task 1), a concept map (Task 2), and a digital poster (Task 3). The develop stage was carried out to test the feasibility of the instrument through validation by an expert panel and a limited trial with students to obtain empirical data.

Logical validation of the instrument draft involved five expert validators serving as expert judges. This expert panel consisted of three chemistry education lecturers (each specializing in digital learning media, chemistry assessment, and learning innovation models, respectively) and two chemistry teachers serving as school practitioners. Input from these experts and practitioners was used to refine both the theoretical aspects and the practical feasibility of the instrument prior to trial testing. The limited trial was conducted at a private senior high school in Bandung, in class XI MIPA 1, which had already covered the topic of chemical bonding. Sampling was carried out using a saturated (total) sampling technique, so that all 19 students in the class were fully included. The digital task products completed by these 19 students were then independently scored by a team of raters to test the reliability of the instrument.

Data collection techniques in this study included interviews, analysis of learning documents, expert validation sheets, and scoring of task products. An initial interview was conducted with a chemistry teacher to identify obstacles to assessment in schools. Document analysis was carried out on curriculum instruments, including Learning Outcomes (CP), Learning Objectives (TP), Learning Objective Flow (ATP), and Learning Objective Achievement Indicators (IKTP). The expert validation sheet used a dichotomous "Yes" (appropriate/valid) or "No" (inappropriate/invalid) response format to assess three aspects: the alignment between indicators and tasks, the alignment between tasks and the rubric, and the alignment between tasks and the PjBL model. Meanwhile, the product scoring sheet was used by raters to score students' work on the three tasks.

Data from the interviews and learning documents (CP, TP, ATP, IKTP) were analyzed using descriptive qualitative methods as the basis for instrument development. Data from the expert validation sheets were analyzed quantitatively using the Content Validity Ratio (CVR) coefficient of Lawshe (1975), calculated using the formula:

$$CVR = \frac{ne - N/2}{N/2}$$

In this formula, the first term denotes the number of validators giving a "Yes" rating, and the second denotes the total number of validators. Based on Lawshe's standard, the minimum CVR coefficient value for a panel of five validators is 0.99 for an instrument item to be declared valid. Subsequently, the reliability of the assessment rubric was calculated based on inter-rater consistency using the Intraclass Correlation Coefficient (ICC) of McGraw and Wong (1996, as cited in (Koo & Li, 2016). Testing employed the Two-Way Mixed-Effects Model of Consistency using SPSS software. The resulting ICC coefficients were then classified according to the criteria presented in Table 1.

***Table 1**. Classification Criteria for the ICC Inter-rater Reliability Coefficient*

ICC Value	Reliability Criteria
< 0.50	Low
0.5–0.74	Moderate
0.75–0.89	Good
≥ 0.90	Excellent

The criteria in Table 1 served as the basis for determining the level of consistency of the creative-thinking performance assessment rubric. A high ICC coefficient value indicates that the analytic rubric for all three electronic portfolio tasks is consistent, objective, and suitable for use, as it can minimize bias in inter-rater scoring.

Results

Definition Stage (Define)

The development of this assessment instrument began with the define stage, carried out through curriculum analysis, mapping of actual conditions in schools, and identification of teacher and student needs. The partner school has currently implemented the Merdeka Curriculum, which requires an in-depth understanding of Learning Objectives (TP). Nevertheless, field data gathered through interviews with a chemistry teacher at a senior high school in Bandung revealed a conceptual gap. Learning evaluation in the classroom to date has still relied on conventional summative multiple-choice tests, while alternative assessment models based on digital portfolios had never been attempted. The teacher also confirmed that chemical bonding is considered a complex topic because it simultaneously involves three levels of representation (macroscopic, microscopic, and symbolic), making it prone to triggering misconceptions if not monitored through continuous evaluation. Based on these conditions, a structured electronic portfolio assessment tool integrated with digital media was designed to document students' creative-thinking profiles. The resulting task analysis produced three forms of digital tasks aligned with the Project-Based Learning (PjBL) model: Task 1, an open-ended essay based on a contextual problem concerning differences in compound solubility; Task 2, the hierarchical reconstruction of a chemical-bonding concept map based on Novak's principles (without cross-links); and Task 3, the creation of a group digital poster as the project's final product. All of these task drafts were subsequently managed online through the Google Classroom platform.

***Table 2.** Synthesis of Interview Results with the Chemistry Teacher*

Aspect	Synthesis Result
Chemical bonding instruction	Chemical bonding is abstract, so instruction remains dominated by lectures, discussion, practice problems, and simple teaching aids; practicum has not been optimal due to facility limitations.
Learning models used	Teachers use various models, from teacher-centered to PBL and PjBL, but their application to chemical bonding remains suboptimal.
Forms of assessment	Assessment remains dominated by written tests such as daily tests, quizzes, practice problems, and semester exams.
Process assessment	Assessment focuses more on final outcomes, while the learning process has not been assessed systematically.
Creative thinking skills	Students' creative thinking skills are considered to remain at a low to moderate level.

Aspect	Synthesis Result
Digital technology utilization	Technology use remains limited to assignment collection via Google Drive and testing via Google Forms.
Electronic portfolio assessment	Teachers have never implemented electronic portfolio assessment, only collecting assignments as files.
Development needs	Teachers stated that a PjBL-based electronic portfolio assessment is needed to help assess creative thinking skills in chemical bonding.

Table 2 presents a thematic synthesis of the interview results conducted with a chemistry teacher prior to instrument development. The findings indicate that chemical bonding instruction remains predominantly conventional and lecture-based, assessment practices are dominated by written tests emphasizing final outcomes, and digital technology utilization for assessment purposes remains minimal. These findings confirm the existence of a conceptual and practical gap that justifies the development of a PjBL-based electronic portfolio assessment instrument capable of documenting students' creative-thinking processes in an authentic and systematic manner, as well as the teacher's explicitly expressed need for such an instrument.

Design Stage

In the design stage, the main focus was on developing the blueprint and initial components of the electronic portfolio instrument (Draft I). The first step began with formulating a task blueprint specifically designed to map the four aspects of creative thinking: fluency, flexibility, originality, and elaboration. In addition, a project-based Student Worksheet (LKM) draft was prepared as a technical guide for students, alongside the development of a draft assessment rubric equipped with explicit operational descriptors to reduce rater subjectivity. The concept-map rubric focused on assessing propositions and hierarchical structuring, while the poster rubric assessed the elements of title, author identity, content substance, and visual aesthetics. Alongside the design of the physical instrument, the researcher also structured the architecture of the e-portfolio environment on the Google Classroom platform. All task instructions, submission deadlines, and rubric drafts were uploaded to the Classwork menu. This digital-space management was prepared so that the complete trail of students' digital artifacts could be neatly documented. This design stage successfully produced a complete draft of the instrument, ready for feasibility testing in the next phase.

Table 3. *Expert Validator Comments and Corresponding Instrument Revisions*

Aspect Revised	Validator Input	Improvement Made
Initial PjBL stage in the LKPD	The LKPD had not yet facilitated students in posing questions.	Added a contextual phenomenon and scaffolding questions so students could formulate a driving question.
Initial analysis in Task 1	The initial analysis did not yet depict a phenomenon close to daily life.	Added the phenomenon of water, oil, and soap along with probing questions to activate students' prior knowledge.
PjBL syntax	The tasks did not yet show full alignment with PjBL syntax.	Added a driving question, project-planning activities, and connecting narratives so that the Task 1 → Task 2 → Task 3 flow aligned with PjBL syntax.
Task 2 (Concept Map)	The concept map did not yet reflect project planning and its link to the project was unclear.	Revised the function of the concept map as the basis for planning the poster project and clarified its connection to the project.
Assessment rubric	The creative-thinking skill indicators were not yet clearly reflected in the instrument and rubric.	Aligned the fluency, flexibility, originality, and elaboration indicators with each task's characteristics and the assessment rubric.

Aspect Revised	Validator Input	Improvement Made
Presentation of the LKPD and instrument	The presentation of questions and instructions was not yet sufficiently informative.	Added titles, sources, introductory sentences, a sample concept map, poster requirements, and completed the LKPD documentation.

Table 3 summarizes the qualitative feedback provided by the five expert validators during the instrument validation process, together with the corresponding revisions made to the electronic portfolio instrument and its accompanying LKPD. The feedback predominantly concerned strengthening the alignment between each task and the syntax of PjBL particularly the explicit presentation of a driving question in Task 1 clarifying the role of the concept map as a problem-solving and project-planning aid in Task 2, adjusting the creative-thinking indicators to more accurately reflect each task's characteristics, and enriching technical specifications and contextual information in Task 3. These revisions were incorporated prior to the empirical trial, thereby strengthening the content validity and instructional coherence of the instrument.

Development Stage (Develop)

The development stage produced the final product: an electronic portfolio assessment instrument together with an assessment rubric validated by an expert panel (expert judgment). This theoretical validation was crucial for testing the initial feasibility of the product before it was deployed in schools. The validation process involved five validators, comprising three chemistry education lecturers from Universitas Pendidikan Indonesia (UPI) and two chemistry teachers serving as school practitioners. Data from the validation sheets were then calculated using Lawshe's Content Validity Ratio (CVR) formula. The results showed that all items in Task 1, Task 2, and Task 3, together with their respective validated assessment rubrics, obtained a perfect CVR coefficient of 1.00. This figure exceeds the minimum standardized threshold (0.99) for a five-member panel, so the instrument was declared highly valid. Beyond the quantitative data, the expert team also provided qualitative input to refine the draft, such as reconstructing the contextual stimulus in LKM 1 by replacing a generic animated video with a video depicting oil-water interaction and the action of soap, as well as inserting scaffolding questions into the task instruction sheet. After the draft was revised based on this input, the instrument was declared ready for empirical trial testing.

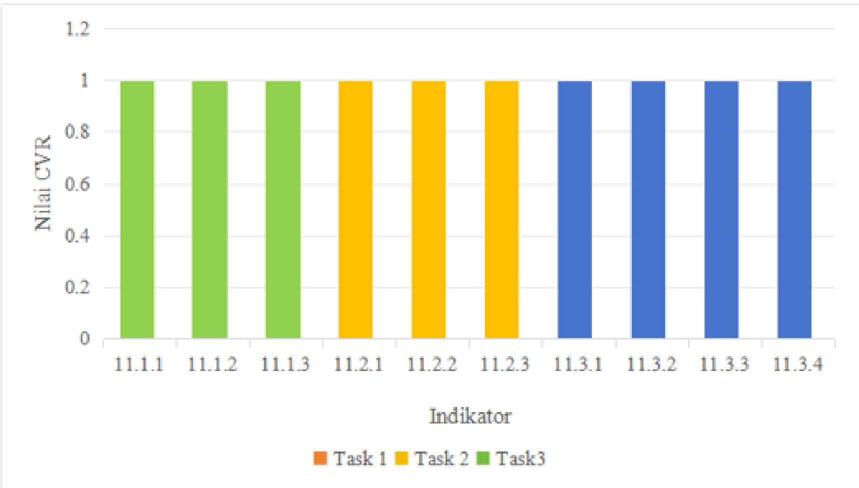

Figure 1. *CVR Values for Each Indicator on Task 1, Task 2, and Task 3*

Figure 1 presents the Content Validity Ratio (CVR) values obtained for each assessment indicator across the three tasks. All ten indicators—three under Task 1 (11.1.1–11.1.3), three under Task 2 (11.2.1–11.2.3), and four under Task 3 (11.3.1–11.3.4)—achieved a uniform CVR value of 1.00. This uniformity indicates unanimous agreement among all five expert validators that every indicator is essential and relevant to the construct being measured, confirming that the instrument possesses excellent content validity across all task types before proceeding to empirical reliability testing.

Empirical testing to measure the instrument's consistency was carried out through a limited trial involving the digital task products of 19 grade XI students at a senior high school in Bandung. The digital portfolio products submitted by these students were independently scored by five raters. The level of inter-rater agreement consistency was analyzed using the Intraclass Correlation Coefficient (ICC) formula with the Two-Way Mixed-Effects Consistency model, the results of which are summarized in Table 2.

Table 4. *ICC Inter-rater Reliability Coefficient Results for the Assessment Instrument*

Task	ICC Coefficient Range	Reliability Criteria
Task 1 (Open-Ended Essay)	0.949-0.981	Excellent
Task 2 (Concept Map)	0.866-0.942	Good to Excellent
Task 3 (Digital Poster)	0.850-0.936	Good to Excellent

Referring to Table 4, the ICC coefficient values for all three digital portfolio task formats exceeded the required reliability threshold. ICC scores ranging from 0.850 to 0.981 indicate that the instrument's level of consistency falls into the Good to Excellent category, with the highest consistency found in Task 1. This empirical data confirms that all three portfolio tasks, together with their validated assessment rubrics, possess high objectivity and remain stable when used by raters to map students' creative-thinking profiles.

Detailed inter-rater reliability output from SPSS (Two-Way Mixed-Effects Model, ICC Consistency) for each indicator within Task 1, Task 2, and Task 3 is presented in Tables 5 through 7.

Task 1 (Open-Ended Essay)

Table 5. *Detailed Inter-rater Reliability (ICC) Output for Task 1 Indicators (11.1.1–11.1.3)*

Indicator	Aspect	ICC (Average Measures)	Reliability Category
11.1.1	A.1	0.949	Excellent
11.1.2	A.2	0.981	Excellent
11.1.3	A.3	0.964	Excellent

Table 5 presents the SPSS ICC (Average Measures) results for the three Task 1 indicators. Indicator 11.1.1 (Aspect A.1) obtained an ICC value of 0.949, indicator 11.1.2 (Aspect A.2) obtained 0.981, and indicator 11.1.3 (Aspect A.3) obtained 0.964, all of which fall within the Excellent reliability category. These consistently high values indicate that the five raters exhibited very strong agreement in scoring students' responses to the open-ended essay task, confirming the stability of the rubric used for this task.

Task 2 (Concept Map)

Table 6. *Detailed Inter-rater Reliability (ICC) Output for Task 2 Indicators (11.2.1–11.2.3)*

Indicator	Aspect	ICC (Average Measures)	Reliability Category
11.2.1	B.1	0.940	Excellent
11.2.2	B.2	0.866	Good
11.2.3	B.3	0.892	Good

Table 6 presents the SPSS ICC (Average Measures) results for the three Task 2 indicators. Indicator 11.2.1 (Aspect B.1) obtained an ICC value of 0.940, falling within the Excellent category, while indicator 11.2.2 (Aspect B.2) obtained 0.866 and indicator 11.2.3 (Aspect B.3) obtained

0.892, both falling within the Good category. Although slightly lower than Task 1, all three indicators remain well above the minimum reliability threshold, indicating consistent rater agreement in assessing students' concept maps.

Task 3 (Digital Poster)

Table 7. Detailed Inter-rater Reliability (ICC) Output for Task 3 Indicators (11.3.1–11.3.4)

Indicator	Aspect	ICC (Average Measures)	Reliability Category
11.3.1	C.1	0.936	Excellent
11.3.2	C.2	0.850	Good
11.3.3	C.3	0.924	Excellent
11.3.4	C.4	0.880	Good

Table 7 presents the SPSS ICC (Average Measures) results for the four Task 3 indicators. Indicators 11.3.1 (Aspect C.1) and 11.3.3 (Aspect C.3) obtained ICC values of 0.936 and 0.924 respectively, falling within the Excellent category, while indicators 11.3.2 (Aspect C.2) and 11.3.4 (Aspect C.4) obtained 0.850 and 0.880, both falling within the Good category. These results confirm that the analytic rubric for the digital poster task produced consistent and objective scores across all five raters.

Discussion

The prevailing pattern of conventional, largely summative-administrative evaluation in schools often neglects the dimension of higher order thinking processes (HOTS), particularly in capturing students' creative capacity. This is supported by (Ramasari et al., 2024), who state that school assessment orientation should not merely pursue a summative function to fulfill final grade requirements, but should instead be directed toward formative assessment that supports the development of individual competencies. On the other hand, chemistry topics such as chemical bonding are laden with abstract concepts that require the full alignment of three levels of representation. The PjBL model enables students to visualize such abstract concepts into tangible products that fully connect the three levels of representation of matter (Situmorang et al., 2022). When classroom evaluation relies solely on theoretical multiple-choice tests, the space for students to explore scientific reasoning becomes constrained. This problem aligns with the reports of previous studied regarding the widespread misconceptions that leave students' understanding of chemical bonding shallow (Hendrawani, 2023; Istiqomah et al., 2021).

Restructuring the function of assessment through the paradigms of assessment for learning and assessment as learning is therefore a pedagogical necessity, consistent with the view ofprevious studied, who position assessment as both a formative feedback instrument and a means of training students' capacity for self-reflection (Rutherford et al., 2025). The use of the Google Classroom platform as a digital artifact storage space in this study offers the practicality of a systematic digital log. This finding is supported by previous studied, who confirm that the implementation of e-portfolios is highly efficient in terms of data storage capacity, facilitates reflection, and overcomes the limitations of the impractical physical format of conventional paper portfolios (Tur et al., 2019; Mudau, 2022). Through organized digital-architecture management, this e-portfolio tool can serve as a dynamic instrument for documenting the developmental trail of students' creative-thinking capacity. Beyond functioning as a medium for documenting learning outcomes, the e-portfolio also supports the implementation of authentic assessment through the continuous collection of learning evidence, allowing students' competency development to be monitored more comprehensively (Trejo González, 2024). This is in line with previous studied, who argue that electronic portfolios are highly effective as a dynamic medium for continuously recording students' learning processes and competency development (Garcia, 2025; Zhang & Tur, 2023).

The theoretical feasibility test of the developed assessment instrument showed a very high level of feasibility quality. The perfect Content Validity Ratio (CVR) coefficient of 1.00 obtained from the expert panel confirms that all task items, together with the blueprint and validated assessment rubrics, possess strong content relevance to the applicable curriculum, which previous studied indicates a highly valid instrument standardization. This alignment occurred because the instrument blueprint was systematically designed by adopting the sequential syntax of Project-Based Learning (PjBL) (Ngui et al., 2022). This is supported by previous studied, who state that project design activities within the PjBL model align closely with the principle of authentic assessment, as they actively evaluate students throughout the project process (Chang et al., 2024; Pratiwi & Ikhsan, 2024).

This task design was specifically constructed to elicit the four key indicators of creativity: fluency, flexibility, originality, and elaboration. These four indicators are also still recognized as the primary dimensions in creative-thinking assessment based on the expert consensus reported by previous studied, and therefore remain relevant as the basis for developing the instrument in this study (López Martínez et al., 2024). This conditioning of indicators is consistent with Williams' theory, which asserts that creativity can be nurtured and developed through structured learning scenarios. Combining the PjBL model with portfolio assessment offers an adaptive alternative for simultaneously measuring learning process and outcomes (Haryati et al., 2024; Sari et al., 2023). From a methodological standpoint, this attainment of perfect content validity forms the initial foundation confirming that the resulting electronic portfolio instrument is valid for measuring students' target creative competencies without departing from the essential boundaries of the chemical bonding topic.

The quality of this instrument was subsequently reinforced empirically through an inter-rater reliability test, with ICC coefficient analysis ranging from 0.850 to 0.981. According to the standardization criteria of previous studied, achievement within the Good to Excellent category indicates a robust level of consistency in the instrument's operationalization. Consistency of perception among the five raters was driven by the availability of clear operational descriptors in the validated assessment rubric drafts, which minimized rater subjectivity bias (Koo & Li, 2016). In Task 2 (Concept Map), the decision to exclude the cross-links component previous studied original principle proved effective in maintaining rater scoring stability at the secondary-school level from (Novak & Cañas, 2006). This is supported by (Rahmawati et al., 2019), who note that focusing assessment on the parameters of scientific propositions and hierarchical concept structuring already makes a substantial contribution to the flexibility and elaboration aspects of students' ideas in a manner realistic for the senior high school level. This finding is also consistent with the meta-analysis results of the previous studied which show that the use of concept maps consistently has a positive impact on students' conceptual understanding and academic achievement, making it relevant as one of the task formats in project-based assessment (Izci & Akkoc, 2024).

A similar pattern of consistency is also reflected in Task 3 (Digital Poster). A similar pattern of consistency is also reflected in Task 3 (Digital Poster). The assessment rubric was developed based on established principles of effective scientific poster presentation, emphasizing the clarity of the title, organization of information, visual design, and effective communication of scientific content. The use of a structured analytic rubric with well-defined criteria is consistent with previous studied, who explain that analytic rubrics can improve objectivity and consistency in inter-rater scoring (Gosselin & Golick, 2020). In addition, using a poster as the final product allows students to effectively communicate scientific concepts through the integration of text, visual illustrations, and a layout that supports readability and reader comprehension (Pedwell et al., 2017).

Among the three task instruments developed, Task 1 (Open-Ended Essay) recorded the highest level of inter-rater agreement reliability (ICC = 0.949-0.981). In contrast to conventional rote memorization tests, which previous studied foster shallow thinking on the topic of chemical bonding, the authentic problem evaluation in Task 1 instead provides students with cognitive freedom (Warsito et al., 2020; Zulkhairi, 2022). This is consistent with the another studied view that authentic assessment through open-ended problems can independently measure students' ability to apply knowledge to real-world contexts (McArthur, 2023). When students are given the freedom to explore the underlying phenomenon of differences in compound solubility using their own wording, the creative-thinking indicators of fluency and flexibility are optimally stimulated.

The incorporation of qualitative input from the expert validator team such as providing scaffolding questions and contextualizing the water-soap interaction phenomenon directly contributed to increasing the precision of this instrument as a reliable measurement tool. Overall, the results of the CVR validation and ICC reliability indices demonstrate that the developed electronic portfolio assessment tool has met mature feasibility criteria. This conclusion is supported by the previous studied, who affirm that the use of a valid and reliable electronic portfolio assessment in chemistry instruction is essential for informing educational decision-making while objectively capturing students' creative-thinking skill profiles (Edwards et al., 2025).

Conclusion

Based on the research findings, the development of the Project-Based Learning (PjBL)-based electronic portfolio assessment instrument for chemical bonding, using the 4D model, was limited to the development stage only, without carrying out the disseminate stage. This study successfully achieved its stated objective of producing a valid and reliable evaluation instrument for measuring students' creative thinking skills. The theoretical validation results from the expert panel showed a very high level of feasibility, with a perfect Content Validity Ratio (CVR) coefficient of 1.00 obtained for all task component items along with their validated assessment rubric drafts. In terms of empirical consistency, the inter-rater reliability test results, through Intraclass Correlation Coefficient (ICC) analysis, showed a robust level of consistency, with value ranges of 0.949–0.981 for Task 1 (Excellent), 0.866–0.942 for Task 2 (Good to Excellent), and 0.850–0.936 for Task 3 (Good to Excellent). These findings confirm that the use of a Google Classroom-integrated electronic portfolio assessment tool has a positive implication as an objective and consistent measurement tool, making it a viable and innovative alternative evaluation instrument in chemistry instruction, particularly for abstract topics requiring the integration of three levels of representation of matter. This tool also facilitates the systematic documentation of digital artifact trails to support the paradigms of assessment for learning and assessment as learning.

However, this study has limitations in that the model stages were restricted to the development phase only, the trial subjects were limited in scope to a single senior high school in Bandung, and the material coverage was limited to the concept of chemical bonding. Future research is therefore recommended to continue development through to the disseminate stage, to test the feasibility of this e-portfolio assessment tool on other chemistry topics, to involve a broader range of education levels, or to combine it with other digital platforms in order to obtain a more comprehensive generalization of the instrument.

Acknowledgment

-

References

Aini, N. Q., Nahadi, N., & Fatimah, S. S. (2024). Development of electronic portfolios to assess concept mastery and creative thinking on project-based acid-based concepts. *Jurnal Pendidikan Kimia*, 16(3). https://doi.org/10.24114/jpkim.v16i3.65188

Ariningsih, N. K. D., Artini, L. P., & Marsakawati, N. P. E. (2021). The effect of e-portfolio in project-based learning toward learner autonomy and writing competency. *Journal of Education Research and Evaluation*, 5(1), 154–162. https://doi.org/10.23887/jere.v5i1.29982

Chang, Y., Choi, J., & Şen-Akbulut, M. (2024). Undergraduate Students' Engagement in Project-Based Learning with an Authentic Context. *Education Sciences*, 14(2), 168. https://doi.org/10.3390/educsci14020168

Edwards, R. A., White, B. A. A., & Findyartini, A. (2025). Editorial: Innovations in teaching and learning for Health Professions Educators. *Frontiers in Medicine*, 12. https://doi.org/10.3389/fmed.2025.1611578

Ernawati, M. D. W., Minarni, Dewi, F., & Yusnidar. (2023). Project-Based Learning Innovations to Improve Students' Creative Thinking Ability in Chemistry Learning Process Development Courses. *Indonesian Journal of Educational Research and Review*, 6(2), 31–321. https://doi.org/10.23887/ijerr.v6i2.66089

Garcia, M. B. (2025). Self-Coded Digital Portfolios as an Authentic Project-Based Learning Assessment in Computing Education: Evidence from a Web Design and Development Course. *Education Sciences*, 15(9), 1150. https://doi.org/10.3390/educsci15091150

Gosselin, D. C., & Golick, D. (2020). Posters as an effective assessment tool for a capstone course. *Journal of Environmental Studies and Sciences*, 10(4), 426–437. https://doi.org/10.1007/s13412-020-00612-x

Haryati, T., Rusdi, M., Asyhar, R., Hadisaputra, S., & Hasibuan, M. H. E. (2024). The Effect of Scaffolding and Creative Thinking Skills in an Acid and Base Learning Project on Students' Science Process Skills. *Jurnal Penelitian Pendidikan IPA*, 10(3), 1083–1092. https://doi.org/10.29303/jppipa.v10i3.4674

Hendrawani, H. (2023). Katalog Miskonsepsi dalam Pembelajaran Ikatan Kimia. *Empiricism Journal*, 4(2), 648–656. https://doi.org/10.36312/ej.v4i2.1729

Huda, P. C., Widiyanto , B., & Hayati , M. N. (2025). Kemampuan Berpikir Tingkat Tinggi melalui Tes Three-Tier Berbasis Indikator HOTS Brookhart pada Pembelajaran IPA SMP Materi Ekologi. *Jurnal Pelita: Jurnal Pembelajaran IPA Terpadu*, 5(2), 419–431. https://doi.org/10.54065/pelita.5.2.2025.889

Istiqomah, W., Rahayu, S., & Muchson, M. (2021). Identifikasi miskonsepsi materi ikatan kovalen pada mahasiswa kimia tahun pertama universitas negeri malang menggunakan tes diagnostik two-tier. *Jurnal MIPA Dan Pembelajarannya*, 1(6). https://doi.org/10.17977/um067v1i6p482-492

Izci, E., & Akkoc, E. A. (2024). The impact of concept maps on academic achievement: A meta-analysis. *Heliyon*, 10(1), e23290. https://doi.org/10.1016/j.heliyon.2023.e23290

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lestari, N. A., Liliasari, Irawan, N. Z. P., & Musthapa, I. (2024). Research Trends of Project-Based Learning Model in Chemistry Learning Through Bibliometric Analysis. *Indonesian Journal of Educational Research and Review*, 7(2), 404–415. https://doi.org/10.23887/ijerr.v7i2.79200

López Martínez, O., Lorca Garrido, A. J., & de Vicente-Yagüe Jara, M. I. (2024). Indicators of verbal creative thinking: results of a Delphi panel. *Frontiers in Psychology*, 15. https://doi.org/10.3389/fpsyg.2024.1397861

Masruroh, M., Purwanto, B. E., & Kusrina, T. (2024). Implementasi Pembelajaran Kimia Berbasis Project untuk Meningkatkan Kreativitas Berpikir dan Literasi Sains. *Journal of Education Research*, 5(3), 2864–2867. https://doi.org/10.37985/jer.v5i3.1361

McArthur, J. (2023). Rethinking authentic assessment: work, well-being, and society. *Higher Education*, 85(1), 85–101. https://doi.org/10.1007/s10734-022-00822-y

Mudau, P. K. (2022). Lecturers' Views on the Functionality of e-Portfolio as Alternative Assessment in an Open Distance e-Learning. *International Journal of Educational Methodology*, 8(1), 81–90. https://doi.org/10.12973/ijem.8.1.81

Ngui, W., Pang, V., & Hiew, W. (2022). Implementing E-Portfolio as an Assessment Tool for Academic Writing in Higher Education. *International Journal of Academic Research in Progressive Education and Development*, 11(3). https://doi.org/10.6007/IJARPED/v11-i3/15402

Novak, J. D., & Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. *Florida Institute for Human and Machine Cognition*, 1(1), 1–31. https://cmap.ihmc.us/docs/theory-of-concept-maps

Pane, E. N. A., & Simamora , P. (2025). Penerapan Model PjBL Upaya Meningkatkan Critical Thinking, Creative Thinking, Collaboration dan Communication (4C) Siswa pada Pembelajaran Fisika. *Jurnal Pelita: Jurnal Pembelajaran IPA Terpadu*, 5(2), 508–522. https://doi.org/10.54065/pelita.5.2.2025.890

Pedwell, R. K., Hardy, J. A., & Rowland, S. L. (2017). Effective visual design and communication practices for research posters: Exemplars based on the theory and practice of multimedia learning and rhetoric. *Biochemistry and Molecular Biology Education*, 45(3), 249–261. https://doi.org/10.1002/bmb.21034

Pratiwi, E. A., & Ikhsan, J. (2024). Project Based Learning (PjBL) in Chemistry Learning: Systematic Literature and Bibliometric Review 2015 - 2022. *Jurnal Penelitian Pendidikan IPA*, 10(6), 343–354. https://doi.org/10.29303/jppipa.v10i6.7017

Rahmawati, Y., Ridwan, A., Hadinugrahaningsih, T., & Soeprijanto. (2019). Developing critical and creative thinking skills through STEAM integration in chemistry learning. *Journal of Physics: Conference Series*, 1156, 012033. https://doi.org/10.1088/1742-6596/1156/1/012033

Ramasari, A. D., Jannah, M., Agustina, S. P., & Zora, F. (2024). Analisis Peran Evaluasi dalam Pembelajaran di Sekolah. *Jurnal Ilmiah Pendidikan Scholastic*, 8(3), 9–16. https://doi.org/10.36057/jips.v8i3.694

Rutherford, S., Pritchard, C., & Francis, N. (2025). Assessment <scp>IS</scp> learning: developing a student-centred approach for assessment in Higher Education. *FEBS Open Bio*, 15(1), 21–34. https://doi.org/10.1002/2211-5463.13921

Sari, A. M., Suryana, D., Bentri, A., & Ridwan, R. (2023). Efektifitas Model Project Based Learning (PjBL) dalam Implementasi Kurikulum Merdeka di Taman Kanak-Kanak. *Jurnal Basicedu*, 7(1), 432–440. https://doi.org/10.31004/basicedu.v7i1.4390

Situmorang, M., Sinaga, M., Sitorus, M., & Sudrajat, A. (2022). Implementation of Project-based Learning Innovation to Develop Students' Critical Thinking Skills as a Strategy to Achieve Analytical Chemistry Competencies. *Indian Journal of Pharmaceutical Education and Research*, 56(1s), s41–s51. https://doi.org/10.5530/ijper.56.1s.41

Sumarni, S. (2019). *Model penelitian dan pengembangan (R&D) lima tahap (MANTAP)*. http://digilib.uin-suka.ac.id/id/eprint/39153

Syahidah, N. A., & Rusmini, R. (2025). Development of Project-Based Learning-Based Learner Activity Sheet to Improve Creative Thinking Skills on Chemical Bonding. *Jurnal Paedagogy*, 12(3), 786. https://doi.org/10.33394/jp.v12i3.16013

Thiagarajan, S. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*.

Trejo González, H. (2024). Evaluación auténtica en contexto universitario a través de portafolios electrónicos de aprendizaje. *Revista Portuguesa de Educação*, 37(2), e24028. https://doi.org/10.21814/rpe.29644

Tur, G., Urbina, S., & Forteza, M. D. (2019). Rubric-Based Formative Assessment in Process Eportfolio: Towards Self-Regulated Learning. *Digital Education Review*, 18–35. https://doi.org/10.1344/der.2019.35.18-35

Warsito, J., Subandi, S., & Parlan, P. (2020). Identifikasi Miskonsepsi Siswa pada Topik Ikatan Kimia Serta Perbaikannya dengan Pembelajaran Model ECIRR (Elicit, Confront, Identify, Resolve, Reinforce). *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 5(11), 1563–1572. https://doi.org/10.17977/jptpp.v5i11.14158

Yang, M., Tai, M., & Lim, C. P. (2016). The role of e-portfolios in supporting productive learning. *British Journal of Educational Technology*, 47(6), 1276–1286. https://doi.org/10.1111/bjet.12316

Zhang, P., & Tur, G. (2023). Collaborative e-Portfolios Use in Higher Education During the COVID-19 Pandemic: A Co-Design Strategy. *International Journal of Educational Methodology*, 9(3), 585–601. https://doi.org/10.12973/ijem.9.3.585

Zulkhairi, Z. (2022). Studi literatur: Analisis miskonsepsi siswa pada topik ikatan kimia. *Jurnal Kajian Pendidikan IPA*, 2(1), 133–140. https://doi.org/10.52434/jkpi11625