JTK: Jurnal Tadris Kimiya 11, 1 (June 2026): 132-144
Website: http://journal.uinsgd.ac.id/index.php/tadris-kimiya/index
ISSN 2527-9637 **(online) ISSN** 2527-6816 **(print)**

Computational Chemistry Software for Developing Deep Conceptual Understanding among Pre-Service Chemistry Teachers

Teguh Wibowo[1*], Ariyatun[2], Annisa Adiwena Putri[3], Ervin Tri Suryandari[3], Azlan Bin Kamari[4], and Jajang Muhariyansah[5]

[1]Department of Chemistry Education, Faculty of Science and Technology, Universitas Islam Negeri Walisongo Semarang, Semarang, 50185, Indonesia
[2]Department of Chemistry Education, Faculty of Mathematics and Natural Science, Universitas Negeri Jakarta, Jakarta Timur, 13220, Indonesia
[3]Department of Chemistry, Faculty of Science and Technology, Universitas Islam Negeri Walisongo Semarang, Semarang, 50185, Indonesia
[4]Department of Chemistry Education, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Tanjong Malim, Perak, 35900, Malaysia
[5]Graduate Institute of Science Education, College of Science, National Taiwan Normal University, Taipei, 106308, Taiwan
[*]E-mail: teguhwibowo@walisongo.ac.id

Received: March 2026; Accepted: June 2026; Published: June 2026

Abstract

Developing students' deep conceptual understanding remains a persistent challenge in chemistry education because many chemical phenomena involve abstract molecular-level representations that are difficult to visualize through conventional instruction. Although computational chemistry software (CCS) has been widely used in scientific research, empirical evidence regarding its effectiveness in promoting meaningful conceptual understanding in undergraduate chemistry education remains limited. This study investigated the effectiveness of CCS in enhancing chemistry education students' deep conceptual understanding. A quasi-experimental nonequivalent control group design was employed involving undergraduate chemistry education students assigned to experimental and control groups. The experimental group learned with computational chemistry software (Gaussian, Avogadro, ChemDraw, and Spartan), whereas the control group received conventional instruction. Students' deep conceptual understanding was assessed using a validated instrument measuring conceptual connectivity, scientific reasoning and concept application, and higher-order thinking skills. Data were analyzed using independent-samples t-tests, effect size analysis, and structural model evaluation with SmartPLS. The findings revealed that students who learned with CCS significantly outperformed those receiving conventional instruction across all dimensions of deep conceptual understanding ($p < .001$). Large effect sizes were observed for conceptual connectivity (0.89) and scientific reasoning and concept application (0.85), while a moderate-to-large effect was found for higher-order thinking skills (0.79). The measurement model also demonstrated satisfactory validity, reliability, and model fit. These findings indicate that CCS serves not only as a molecular visualization tool but also as an effective pedagogical resource for supporting scientific reasoning, conceptual integration, and higher-order thinking in chemistry learning. The study contributes empirical evidence supporting the integration of computational chemistry into undergraduate curricula to foster deeper conceptual understanding and digitally enriched chemistry instruction.

Keywords: chemistry learning, critical thinking, computational chemistry, conceptual understanding, educational software

DOI: https://doi.org/10.15575/jtk.v11i1.51648

1. Introduction

Teaching chemistry at the university level continues to face significant challenges in explaining abstract concepts such as molecular structure and geometry, bond energy, intermolecular interactions, and chemical reaction mechanisms (Corradi et al., 2014; Lee et al., 2024). Conventional teaching media, whether static images or verbal explanations, are often insufficient to help students develop a deep conceptual understanding (Arsyad et al., 2024). Consequently, learning tends to become memorization-based, offering limited opportunities for students to cultivate critical, analytical, and creative thinking skills needed to solve complex problems. Meanwhile, technological advancements have introduced Computational Chemistry Software (CCS) such as Gaussian, Avogadro, ChemDraw, and Spartan, which were initially developed for research purposes (Deringer et al., 2021; Lehtola & Karttunen, 2022). These tools allow users to model molecular structures in three dimensions, calculate physicochemical properties, and simulate reaction processes, making it possible to visualize concepts that are otherwise difficult to imagine manually (Fombona-Pascual et al., 2022; Li & Wei, 2024). Although CCS has enormous potential to bridge the gap between theoretical understanding and real-world chemical phenomena, its application in classroom learning remains limited and has yet to be fully optimized as an integral part of chemistry education.

In an educational context, the goal is not merely to transfer information but to foster deep conceptual understanding (Zhao et al., 2024). Such understanding requires that students not only recall facts or definitions but also connect interrelated concepts, explain phenomena with scientific reasoning, and apply their knowledge to new situations or complex problem-solving tasks (Vosniadou, 2019). This approach encourages the development of higher-order thinking skills such as analysis, evaluation, and creation. Deep understanding also provides a strong foundation for students to adapt to ongoing advances in science and technology while preparing them to become lifelong learners (Pramesti & Riastini, 2023; Azid et al., 2022). The integration of CCS into university-level chemistry instruction should therefore be aligned with student-centered pedagogical strategies (Byusa et al., 2022). Incorporating CCS into project-based or inquiry-based learning models can engage students in designing, simulating, and analyzing virtual experiments (Mayer, 2024). This not only enhances conceptual comprehension but also nurtures scientific curiosity, collaborative skills, and critical thinking, which are essential competencies for success in the twenty-first century.

Research in chemical education further indicates that many student misconceptions arise from weak connections between concepts (Üce & Ceyhan, 2019). Technology-enhanced learning, including the use of CCS, has been shown to effectively reduce such conceptual gaps (Erümit & Sarıalioğlu, 2025; Serhan et al., 2019). Three-dimensional molecular visualization enables students to strengthen spatial reasoning and better understand reaction mechanisms (Kuit & Osman, 2021; Serhan et al., 2019). Consequently, integrating CCS not only supports curriculum objectives but also enhances students' scientific and technological literacy, which are key competencies in the twenty-first century (Bílgín et al., 2017). In line with the principles of the Merdeka Curriculum and twenty-first-century learning demands, CCS integration supports science literacy, technological literacy, and collaboration skills. Moreover, it reinforces inquiry-based and problem-based learning approaches, where students play an active role in constructing knowledge through data exploration and simulation (Mishra, 2023).

Although numerous previous studies have highlighted the benefits of technology use in science education, particularly in chemistry, most have focused primarily on visual media or virtual laboratory applications (Nehring & Schanze, 2025). Research on the use of CCS within classroom learning contexts remains relatively limited, especially in Indonesia. In fact, CCS holds a unique role: beyond providing visual representations, it allows students to engage in data-driven computational explorations that closely resemble authentic scientific research (Du et al., 2023; Panda, 2024). This characteristic distinguishes CCS from conventional media or typical virtual laboratories, as it integrates theory, simulation, and practice into a single cohesive process. However, there remains a gap between the potential of CCS and its actual implementation in higher education. Many chemistry lecturers are still unfamiliar with CCS as a pedagogical tool, leading to its predominant use in research rather than in teaching (Pyatt, 2022).

The absence of practical guidelines and the lack of established implementation models in curricula further widen the gap between ideal pedagogical theory and classroom realities (Honke & Becker-Genschow, 2025; Ipinnaiye & Risquez, 2024; Kabudi et al., 2021). Consequently, students often gain only superficial exposure to CCS, without meaningful hands-on experiences that could strengthen their conceptual understanding. The urgency of this research also lies in responding to the demands of twenty-first-century education, which emphasizes critical thinking, problem-solving, and adaptability to digital technologies (Talanquer, 2025; Gombert et al., 2024; Rincon-Flores et al., 2024). If chemistry learning continues to rely predominantly on traditional methods, students may be inadequately prepared to face global challenges that require multidisciplinary competencies (Quan et al., 2025; Whatoni & Sutrisno, 2022). CCS can serve as a bridge for integrating scientific, technological, and computational skills into a unified learning experience. Accordingly, studies on the integration of CCS in chemistry education are not only academically relevant but also strategically significant in equipping students with more comprehensive scientific and digital literacy (Hoai et al., 2023). Thus, further investigation into the effective implementation of CCS is necessary to understand how this technology can enhance students' conceptual understanding and scientific skills.

A notable research gap also exists in evaluating the effectiveness of CCS in improving students' conceptual comprehension. Most prior studies have focused on user satisfaction or increased learning motivation (Abdo et al., 2024; Eitemüller et al., 2023), while few have examined how CCS specifically helps reduce misconceptions, strengthen conceptual interconnections, or develop higher-order thinking skills. Therefore, this study is important in addressing that gap and providing new directions for innovation in chemistry education at the tertiary level. The urgency of this research also stems from the scarcity of exploratory studies on students' needs for technology-based chemistry software that aligns with the learning context of Indonesian universities. Although several CCS tools are already available, most are commercial, foreign-language-based, require high device specifications, and are not yet adapted to the national curriculum framework or local learner characteristics (Byusa et al., 2022). Preliminary findings from observations and interviews suggest that students need software that is more intuitive, lightweight, and interactive, supporting both inquiry-based and collaborative learning processes.

2. Research Method

2.1. Research Procedure

This study employed a quasi-experimental design using the posttest-only design with a nonequivalent groups model. The procedure began with a preparatory stage, which included identifying abstract chemical concepts and developing research instruments such as the conceptual understanding test and perception questionnaire. Subsequently, lecturers and students received training on the use of CCS tools, including Gaussian, Avogadro, ChemDraw, and Spartan. The implementation phase involved dividing students into two groups: 1) an experimental group, which utilized CCS in the learning process; and 2) a control group, which followed the conventional teaching method. Data were collected through pretests and posttests to assess conceptual understanding, along with a student perception questionnaire to capture affective responses toward CCS-assisted learning.

2.2. Instruments and Data Analysis Techniques

The instruments used in this study consisted of a deep conceptual understanding test and a student perception questionnaire. The conceptual understanding test included three main indicators: 1) Interconnection among concepts; 2) Scientific reasoning and conceptual application; and 3) Higher-Order Thinking Skills. The test comprised multiple-choice items with reasoning and short-answer questions, designed to measure students' mastery of abstract chemical concepts such as molecular structure, bond energy, and reaction mechanisms. Data collection was carried out in several stages: 1) Both groups were given a pretest to measure their initial abilities; 2) During the learning process, observations were conducted to record learning activities and any technical challenges; and 3) At the end of the treatment, students took a posttest to assess cognitive improvement.

Instrument validation was conducted using SmartPLS 3.0 through several stages: 1) Convergent validity was assessed using outer loadings (factor loadings) and Average Variance Extracted (AVE); and 2) Discriminant validity was evaluated using the Fornell-Larcker criterion and cross-loading tests. The results of the convergent validity test indicated that some outer loading and AVE values did not meet the minimum standard of 0.5. Therefore, cross-loading analysis was performed to further confirm discriminant validity. In this test, the loading value of each indicator on its own construct must be higher than its loading value on other constructs. Instrument reliability was confirmed through Composite Reliability (CR) and Cronbach's alpha values, both exceeding the 0.6 threshold. Overall, the validity and reliability of the measurement model were analyzed through the outer model test in SmartPLS.

The effectiveness of CCS implementation was evaluated semi-summatively to determine whether the intervention met the predefined specifications. The analysis was carried out using the posttest-only design with nonequivalent groups and included prerequisite assumption testing. An independent sample t-test was performed to examine whether there were significant differences in students' deep conceptual understanding between the CCS and non-CCS groups. The decision criterion was based on a significance level ($p < 0.05$), where H_0 is rejected, and H_1 is accepted if the significance value is below 0.05. To further determine the magnitude of the CCS effectiveness, effect size and R-squared values were calculated.

3. Result and Discussion

3.1. Instrument Validity and Reliability

The instrument validity test was carried out through several stages using SmartPLS 3.0 software. The validity test included convergent validity, assessed through outer loadings (factor loadings) and AVE, and discriminant validity, evaluated using the Fornell–Larcker criterion and cross-loading tests. In the convergent validity test, several initial values of outer loadings and AVE did not meet the standard threshold of 0.5. The initial AVE values are presented in Table 1.

Table 1. AVE Values of Deep Conceptual Understanding

	Cronbach's Alpha	rho_A	Composite Reliability	AVE
Connectivity between concepts	0.872	0.861	0.812	0.666
Scientific reasoning and concept application	0.812	0.803	0.844	0.697
Higher-order thinking skills	0.763	0.779	0.885	0.569

Source: SmartPLS Algorithm Output (2025)

As shown in Table 1, all variables met the AVE criterion of ≥ 0.5, indicating that the convergent validity test was acceptable. The AVE values for the three dimensions of deep conceptual understanding were all above the minimum acceptable threshold (0.5), suggesting that the indicators used in this study could effectively represent their respective latent constructs. In other words, each construct consistently reflected the intended dimension of conceptual understanding. Moreover, the AVE values ranging from 0.569 to 0.697 indicate that the proportion of variance explained by the indicators exceeded the variance due to measurement error. This result confirms that the instrument demonstrated satisfactory convergent validity. Among the three dimensions, the scientific reasoning and concept application achieved the highest AVE (0.697), implying that the indicators contributed most strongly to the underlying construct.

The reliability values, reflected by Cronbach's alpha and composite reliability, also indicate good internal consistency. All dimensions had composite reliability values above 0.7, which met the reliability criteria proposed by Nunnally and Bernstein (1994). This suggests that the instrument was valid and reliable in measuring deep conceptual understanding. Therefore, based on the results of the convergent validity and reliability tests, it can be concluded that the instrument was appropriate for measuring the construct of deep conceptual understanding. The next stage involved testing discriminant validity to ensure that each construct was distinct and did not overlap in measuring different aspects of the concept.

The discriminant validity of the instrument was evaluated using the Fornell–Larcker criterion and cross-loading tests. The Fornell–Larcker test compared the square root of the AVE of each construct with its correlations with other latent constructs. The required condition was that the square root of the AVE for a construct should be greater than its correlations with any other constructs. The Fornell–Larcker criterion values are presented in Table 2.

Table 2. Fornell–Larcker Criterion Values

	Connectivity between concepts	**Scientific reasoning and concept application**	**Higher-order thinking skills**
Connectivity between concepts	0.815	-	-
Scientific reasoning and concept application	0.737	0.818	-
Higher-order thinking skills	0.760	0.822	0.794

Source: SmartPLS Algorithm Output (2025)

The results in Table 2 show that the square root of the AVE (values along the diagonal) for each construct was greater than the correlations between constructs. For instance, the construct "connectivity between concepts" had a value of 0.815, which was higher than its correlations with "scientific reasoning and concept application" (0.737) and "higher-order thinking skills" (0.760). Similarly, "scientific reasoning and concept application" had a diagonal value of 0.818, which exceeded its correlations with other constructs (0.737 and 0.822). The same pattern was observed for "higher-order thinking skills," with a diagonal value of 0.794 that was higher than its correlations with the other constructs.

These findings confirm that each construct could be clearly distinguished from the others within the research model. In other words, although the dimensions were interrelated, as shown by the relatively significant correlations, they each measured distinct aspects of deep conceptual understanding. This reinforces that the instrument met convergent validity and achieved satisfactory discriminant validity.

The relatively strong correlation between "scientific reasoning and concept application" and "higher-order thinking skills" (0.822) suggests that these two constructs were conceptually related in shaping deep conceptual understanding. However, since the diagonal values remained higher than the inter-construct correlations, the Fornell–Larcker criterion was still satisfied. Therefore, it can be concluded that the research instrument demonstrated good discriminant validity and was suitable for further analysis.

3.2. Effectiveness of Computational Chemistry Software Use on Deep Conceptual Understanding

The results of the study indicate that the use of CCS positively contributes to the enhancement of students' conceptual understanding. Data were analyzed using a posttest-only design with a nonequivalent groups model. An independent sample t-test was conducted to determine whether there was a significant difference in deep conceptual understanding between students who used CCS and those who did not. The decision rule for the independent sample t-test was that if the significance value (p) is less than 0.05, the null hypothesis (H_0) is rejected, and the alternative hypothesis (H_1) is accepted. Prior to performing the t-test, prerequisite tests were conducted to ensure that the assumptions of normality and homogeneity were met. The results of the normality test are presented in Table 3.

Table 3. Test of Normality (Shapiro–Wilk)

Indicator	Group	W	p
Connectivity between concepts	Control	0.852	0.076
	Experiment	0.987	0.146
Scientific reasoning and concept application	Control	0.887	0.083
	Experiment	0.954	0.056
Higher-order thinking skills	Control	0.897	0.083
	Experiment	0.923	0.056

Note. Significant results indicate deviation from normality

The normality test was conducted to ensure that data from the control and experimental groups were normally distributed before proceeding with parametric analysis using the independent sample t-test. As shown in Table 3, all significance (p) values obtained from the Shapiro–Wilk test were greater than 0.05. For example, in the construct "Connectivity between concepts," the control group had a p-value of 0.076 and the experimental group 0.146, both exceeding the 0.05 threshold, thus confirming normal distribution. Similarly, for "Scientific reasoning and concept application," the control group yielded p = 0.083, and the experimental group yielded p = 0.056. Although the experimental group's value approached the threshold, it remained within acceptable limits of normality. The same pattern was found for "higher-order thinking skills," with p = 0.083 (control) and p = 0.056 (experiment). Since all data met the assumption of normality, the independent sample t-test could be appropriately applied. This result reinforces the validity of the study's findings, ensuring that the differences observed in the inferential analysis are credible and not affected by non-normal data distribution. The next prerequisite test was the homogeneity of variance, verified using Levene's test, as shown in Table 4.

Table 4. Test of Equality of Variances (Levene's Test)

	F	df1	df2	p
Connectivity between concepts	0.819	1	103	0.176
Scientific reasoning and concept application	0.185	1	103	0.283
Higher-order thinking skills	0.133	1	103	0.756

The results of Levene's test indicate that all variables had significance values greater than 0.05 ($p = 0.176$, 0.283, and 0.756, respectively), confirming that the assumption of homogeneity of variances was satisfied. Therefore, the variances between the control and experimental groups were homogeneous. Meeting this assumption is essential because it ensures that any differences detected by the t-test are attributable to the treatment effect (use of CCS) rather than unequal data dispersion between groups. Given that Levene's test showed no significant difference in variances, the standard independent sample t-test assuming equal variances could be applied. The summary of the t-test results for deep conceptual understanding is presented in Table 5.

The independent sample t-test results (see Table 5) revealed significant differences between the control and experimental groups across all three indicators of deep conceptual understanding. All significance values (Sig. 2-tailed) were below 0.05, even reaching < .001, leading to the rejection of H_0 and acceptance of H_1. This result confirms that the use of CCS had a significant positive effect on students' deep conceptual understanding. More specifically, in the "Scientific reasoning and concept application" indicator, the experimental group achieved a higher mean score (M = 79.987) than the control group (M = 72.488), with a significant difference (t = -5.628, p < .001). This suggests that CCS effectively trained students to engage in scientific reasoning and apply chemical concepts more proficiently than conventional methods. For "Connectivity between concepts," the experimental group also outperformed the control group (M = 122.215 vs. M = 113.725; t = -6.023, p < .001), indicating that CCS facilitated students' ability to connect complex chemical concepts, fostering more integrative understanding. Lastly, in "Higher-order thinking skills", the experimental group recorded a higher mean (M = 78.069) than the control group (M = 76.075), with a significant difference (t = -4.519, p < .001). This result emphasizes that CCS promoted the development of higher-order thinking skills such as analysis, evaluation, and synthesis in problem-solving within chemistry learning contexts.

Table 5. Independent Sample t-Test Results for Deep Conceptual Understanding

Indicator	Group	Mean	SD	t	Sig. (2-tailed)
Scientific reasoning and concept application	Control	72.488	5.972	-5.628	< .001
	Experiment	79.987	7.993		
Connectivity between concepts	Control	113.725	15.014	-6.023	< .001
	Experiment	122.215	23.011		
Higher-order thinking skills	Control	76.075	10.830	-4.519	< .001
	Experiment	78.069	11.236		

Overall, these results demonstrate the effectiveness of CCS in enhancing students' deep conceptual understanding, both in terms of conceptual connectivity, scientific reasoning, and mastery of higher-order thinking skills. This effectiveness aligns with the literature emphasizing the importance of integrating computational technology in chemistry learning to facilitate more meaningful understanding and reduce misconceptions. Furthermore, an effect size test was conducted to determine the strength of CCS's effect on students' deep conceptual understanding, as presented in Table 6.

Table 6. Effect Size Test Results on Deep Conceptual Understanding

Indicator	Effect Size	Interpretation
Connectivity between concepts	0.89	Large
Scientific reasoning and concept application	0.85	Large
Higher-order thinking skills	0.79	Medium

The effect size test results in Table 6 show that the use of CCS had a significant effect on students' deep conceptual understanding. For the indicator "Connectivity between concepts," the effect size value of 0.89 falls within the large category. This confirmed that CCS was highly effective in helping students build connections among chemistry concepts. Through three-dimensional molecular visualizations and computational simulations, students could more easily understand how one concept was directly related to another, thus forming a more integrative network of knowledge. Next, the indicator "Scientific reasoning and concept application" also showed a large effect size value of 0.85. This finding indicates that CCS made a strong contribution to improving students' ability to provide explanations based on scientific reasoning and to apply concepts in various contexts. Activities such as computational data exploration, bond energy analysis, and reaction simulations enabled students to think more logically and systematically.
Meanwhile, the indicator "Higher-order thinking skills" obtained an effect size value of 0.79, which falls within the small-to-large category. Although slightly lower than the other two indicators, this finding still shows that CCS played a significant role in encouraging students to develop higher-order thinking skills such as analysis, evaluation, and creation. This difference can be explained by the fact that higher-order thinking skills require more time and intensive practice, so the effect is not as strong as on conceptual connectivity or scientific reasoning. Overall, the effect size test results reinforced the earlier t-test results, indicating that CCS not only produced statistically significant differences but also provided a practically strong and meaningful effect on students' deep conceptual understanding. Therefore, integrating CCS can be considered an effective and innovative approach to improving the quality of chemistry learning in higher education.

The next analysis was conducted using the Standardized Root Mean Square (SRMS) values from chemistry education students' deep conceptual understanding test, which was 0.089, less than 0.10, categorized as a model fit (Dash & Paul, 2021). The output of the model fit for the inner model of deep conceptual understanding using SmartPLS 3 is presented in Table 7.

Table 7. Model Fit Test Output

	Saturated Model	Estimated Model	Model-Fit Criteria
SRMS	0.079	0.079	SRMS<0.08
d_ULS	1.792	1.792	d_ULS > 2.000
d_G	1.402	1.402	d_G > 0,900
Chi-Square	715.722	715.722	
NFI	0.915	0.915	NFI>0.9

The model fit test results in Table 7 indicate that all indicators of deep conceptual understanding met the model's goodness-of-fit criteria based on the analysis using SmartPLS 3. First, the SRMS value of 0.079 was below the 0.08 threshold proposed by Dash & Paul (2021), indicating that the model had a good level of consistency between the actual data and the model estimation. Thus, it can be concluded that the structural model used in this study fit the analyzed data.

Second, the d_ULS value of 1.792 was lower than the 2.000 threshold, supporting that the model was free from extreme discrepancies. Meanwhile, the d_G value of 1.402, which was higher than 0.900, further strengthened the evidence that the model had a good and reliable level of fit.

Third, the NFI (Normed Fit Index) value of 0.915 exceeded the minimum threshold of 0.90, indicating that the model showed an adequate level of fit based on the comparison between the Chi-Square value of the tested model and that of the independent model. This demonstrates that the research model could explain the data efficiently. Overall, the model fit indicators obtained confirmed that the measurement instrument for students' deep conceptual understanding was valid and reliable. Moreover, these results reaffirm that the integration of CCS in chemistry learning can be properly analyzed through a well-fitted structural model, ensuring that the research findings are trustworthy for use in developing innovative learning practices in higher education.

The findings of this study revealed that the use of CCS had a significant effect on enhancing students' deep conceptual understanding, as reflected in three main indicators: connectivity between concepts, scientific reasoning and concept application, and higher-order thinking skills. The significant mean score difference between the control and experimental groups, supported by a small-to-large effect size, confirmed that the integration of CCS in chemistry learning was not only statistically effective but also practically meaningful. These results reinforce the argument that simulation-based technology can effectively bridge the gap between abstract representations and real-world phenomena, one of the major challenges in chemistry education (Zendler & Greiner, 2020).

Furthermore, these findings are consistent with recent literature emphasizing the importance of digital technology utilization in strengthening conceptual understanding through interactive visualization and dynamic simulations. As noted by Muthmainnah et al. (2022), CCS serves not merely as a medium for representation but also as an exploratory tool that enables students to test hypotheses and analyze data as in real scientific research (Bílgín et al., 2017; Nehring & Schanze, 2025; Panda, 2024). This effect was evident particularly in the construct of scientific reasoning and concept application, where students using CCS demonstrated a stronger ability to provide scientifically grounded explanations. Consequently, CCS contributes to shifting students' learning orientation from rote memorization toward analytical and applied thinking.

In terms of higher-order thinking skills, the effect size falls within the medium category, suggesting that developing such complex cognitive skills through CCS requires a more intensive process. This aligns with the findings of Du et al. (2023), who highlight that while computational simulations effectively improve analytical abilities, fostering evaluative and creative capacities requires sustained practice and the integration of inquiry-based and problem-based learning strategies. Therefore, CCS implementation should be embedded within learning designs that emphasize contextual problem exploration and reflective discussion to maximize its effect on higher-order thinking. Additionally, the model fit indices, with SRMR < 0.08 and NFI > 0.9, indicate that the research model fit the data well, thereby strengthening the validity of the findings. This is consistent with Hoai et al. (2023), who emphasized that the adequacy of structural models serves as a crucial indicator for ensuring the reliability of research inferences. Thus, this study not only demonstrates the effectiveness of CCS but also offers a robust analytical model that may serve as a reference for future studies.

From theoretical and practical perspectives, this research carries several important implications for chemistry education. Theoretically, the findings add empirical evidence that the use of CCS facilitates the development of deep conceptual understanding among students (Corradi et al., 2014; Panda, 2024). CCS functions not merely as a visual representation tool but also as a conceptual exploration medium integrating theory, simulation, and practical application. Hence, this study contributes to expanding the literature on the effectiveness of digital technology in fostering conceptual connectivity, scientific reasoning, and higher-order thinking skills (Du et al., 2023; Erümit & Sarıalioğlu, 2025). Practically, the main implication lies in the importance of integrating CCS into university-level chemistry curricula. The adoption of CCS can serve as an innovative strategy to bridge the gap between memorization-based learning and twenty-first-century competency demands, which emphasize scientific literacy, technological fluency, and analytical skills (Kabudi et al., 2021; Whatoni & Sutrisno, 2022). Chemistry lecturers can utilize CCS within inquiry-based or problem-based learning frameworks, enabling students to comprehend concepts as well as think critically, evaluate data, and solve problems collaboratively (Mayer, 2024). This study further demonstrates that integrating CCS aligns with the spirit of modern curricula that emphasize flexibility, innovation, and digital literacy enhancement in teaching and learning. Therefore, higher education institutions should provide infrastructural support, lecturer training, and pedagogical guidelines to optimize CCS utilization in classrooms. Such support ensures that technology does not remain a supplementary tool but becomes an integral component of the learning process.

4. Conclusion

This study demonstrates that the use of Computational Chemistry Software (CCS) had a significant effect on enhancing chemistry education students' deep conceptual understanding. The results of the validity and reliability tests confirmed that the indicators, covering connectivity between concepts, scientific reasoning and concept application, and higher-order thinking skills, were valid and reliable measures for assessing deep conceptual understanding. The independent sample t-test results revealed a statistically significant difference in mean scores between the experimental group using CCS and the control group that did not. Furthermore, the effect size, which falls within the medium to large category, reinforces that CCS was not only statistically effective but also practically meaningful in improving students' ability to connect concepts, engage in scientific reasoning, and develop higher-order thinking skills. In addition, the model fit results obtained through SmartPLS analysis indicate that the research model demonstrates good validity, confirming that the findings are robust and reliable for guiding future instructional innovations. The implications of this study highlight the importance of integrating CCS into chemistry curricula at the tertiary level as an innovative pedagogical strategy aligned with the Merdeka Curriculum and twenty-first-century competency demands. Thus, CCS serves not only as a visualization tool but also as an exploratory learning medium that effectively strengthens students' deep conceptual understanding in chemistry. Therefore, based on the results of this study, further research into the development of CCS can be conducted, with the hope that the developed CCS will be useful and easier to apply to chemistry learning.

References

Abdo, S. N., Hsu, J. L., Kapetanakis, C., Newman, D. L., Wright, L. K., & Bailey, J. (2024). An exploration of spatial visualization skills: Investigating students' use of 3D models in science problems during think-aloud interviews. *Journal of Chemical Education, 101*(9), 3624–3634. https://doi.org/10.1021/acs.jchemed.3c01355

Arsyad, M., & Syakhrani, A. W. (2024). The efficiency of using visual learning media in improving the understanding of science concepts in elementary school students. *Indonesian Journal of Education (INJOE)*, *4*(1), 775-787. Retrieved from https://felifa.net/index.php/INJOE/article/view/234

Azid, N., Ali, R. M., Khuluqo, I. Purwanto, S. E. & Susanti, E. N. (2022). Higher order thinking skills, school-based assessment and students' mathematics achievement: Understanding teachers' thoughts. *International Journal of Evaluation and Research in Education (IJERE)*, *11*(1), 290-302. https://doi.org/10.11591/ijere.v11i1.22030

Bílgín, A. K., Yürükel, F. N. D., & Yígít, N. (2017). The effect of a developed REACT strategy on the conceptual understanding of students: "Particulate nature of matter." *Journal of Turkish*

Science Education, *14*(2), 65–81. Retrieved from https://www.tused.org/index.php/tused/article/view/155

Byusa, E., Kampire, E., & Mwesigye, A. R. (2022). Game-based learning approach on students' motivation and understanding of chemistry concepts: A systematic review of literature. *Heliyon*, *8*(5), e09541. https://doi.org/10.1016/j.heliyon.2022.e09541

Corradi, D. M. J., Elen, J., Schraepen, B., & Clarebout, G. (2014). Understanding possibilities and limitations of abstract chemical representations for achieving conceptual understanding. *International Journal of Science Education*, *36*(5), 715–734. https://doi.org/10.1080/09500693.2013.824630

Dash, G. & Paul, J. (2021). CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change,* 173, 121092. https://doi.org/10.1016/j.techfore.2021.121092

Deringer, V. L., Bartók, A. P., Bernstein, N., Wilkins, D. M., Ceriotti, M., & Csányi, G. (2021). Gaussian process regression for materials and molecules. *Chemical reviews*, *121*(16), 10073-10141. https://doi.org/10.1021/acs.chemrev.1c00022.

Du, D., Baird, T. J., Bonella, S., & Pizzi, G. (2023). OSSCAR, an open platform for collaborative development of computational tools for education in science. *Computer Physics Communications*, 282. https://doi.org/10.1016/j.cpc.2022.108546

Eitemüller, C., Trauten, F., Striewe, M., & Walpuski, M. (2023). Digitalization of multistep chemistry exercises with automated formative feedback. *Journal of Science Education and Technology, 32*(3), 453–467. https://doi.org/10.1007/s10956-023-10043-2

Erümit, A. K., & Sarıalioğlu, R. Ö. (2025). Artificial intelligence in science and chemistry education: a systematic review. *Discover Education*, *4*(1). https://doi.org/10.1007/s44217-025-00622-3

Fombona-Pascual, A., Fombona, J., & Vicente, R. (2022). Augmented reality, a review of a way to represent and manipulate 3D chemical structures. *Journal of Chemical Information and Modeling, 62*(8), 1863–1872. https://doi.org/10.1021/acs.jcim.1c01255

Gombert, S., Fink, A., Giorgashvili, T., Jivet, I., Di Mitri, D., Yau, J., Frey, A., & Drachsler, H. (2024). From the automated assessment of student essay content to highly informative feedback: a case study. *International Journal of Artificial Intelligence in Education, 34*(4), 1378–1416. https://doi.org/10.1007/s40593-023-00387-6

Hoai, V. T. T., Son, P. N., Em, V. V. D., & Duc, N. M. (2023). Using 3D molecular structure simulation to develop chemistry competence for Vietnamese students. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(7). https://doi.org/10.29333/ejmste/13345

Honke, N., & Becker-Genschow, S. (2025). Adaptive learning in bionics: transforming science education. *Frontiers in Education*, 10(February), 1–21. https://doi.org/10.3389/feduc.2025.1427083

Ipinnaiye, O., & Risquez, A. (2024). Exploring adaptive learning, learner-content interaction and student performance in undergraduate economics classes.

Computers and Education, 215(March), 105047. https://doi.org/10.1016/j.compedu.2024.105047

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, 2(December 2020), 100017. https://doi.org/10.1016/j.caeai.2021.100017

Kuit, V. K., & Osman, K. (2021). Eficacia del módulo electrónico CHEMBOND3D para mejorar el conocimiento de los estudiantes sobre el concepto de enlace químico y las habilidades visoespaciales. *European Journal of Science and Mathematics Education*, *9*(4), 252–264. https://doi.org/10.30935/scimath/11263

Lee, O. S., Gather, M. C. & Zysman-Colman, E. (2024). Digichem: computational chemistry for everyone. *Digital Discovery*, 3, 1695-1713. https://doi.org/10.1039/D4DD00147H

Lehtola, S., & Karttunen, A. J. (2022). Free and open-source software for computational chemistry education. *Wiley Interdisciplinary Reviews: Computational Molecular Science, 12*(5), 1–33. https://doi.org/10.1002/wcms.1610

Li, H., & Wei, X. (2024). A concise review of biomolecule visualization. *Current Issues in Molecular Biology*, *46*(2), 1318-1334. https://doi.org/10.3390/cimb46020084

Mayer, R. E. (2024). The Past, present, and future of the cognitive theory of multimedia learning. *Educational Psychology Review*, *36*(1), 1–25. https://doi.org/10.1007/s10648-023-09842-1

Mishra, J. K. (2023). Inquiry based learning: issues and challenges. *Echetana: International Journal of Education, 8*(1), 157-163. Retrieved from https://echetana.com/wp-content/uploads/2023/04/738.-R-E-Hari-K.-Sharma.pdf

Muthmainnah, Ibna Seraj, P. M., & Oteir, I. (2022). Playing with AI to investigate human-computer interaction technology and improving critical thinking skills to pursue 21st century age. *Education Research International*, 2022. https://doi.org/10.1155/2022/6468995

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill.

Nehring, A. & Schanze, S. (2025). Turning the Plurality of Chemistry into a Resource for Learning: A Core Competency of Chemistry Teachers. *Science & Education, 34*, 2051–2078. https://doi.org/10.1007/s11191-025-00624-5

Panda, D. R. (2024). Human computer interaction strategies for effective digital learning experiences: From classroom to screen. *International Journal of Scientific Research in Engineering and Management*, 08(008), 1–5. https://doi.org/10.55041/ijsrem37370

Pramesti, N. K. A. & Riastini, P. N. (2023). Professional status: teachers' difficulties in compiling questions based on higher order thinking skills. *International Journal of Elementary Education*, *7*(3), 516-523. https://doi.org/10.23887/ijee.v7i3.61625

Pyatt, K. (2022). Use of chemistry software to teach and assess model-based reaction and equation knowledge. *Journal of Technology and Science Education*, *4*(4), 215–227. http://www.jotse.org/index.php/jotse/article/view/110/142

Quan, W., Yue, D. W. & Bei Er, L. (2025). The application and challenges of virtual reality (VR) painting technology in art education in china: a case study of universities in mianyang, sichuan province. *International Journal of Academic Research in Progressive Education & Development, 14*(4), 335-359. http://dx.doi.org/10.6007/IJARPED/v14-i4/26534

Rincon-Flores, E. G., Castano, L., Guerrero Solis, S. L., Olmos Lopez, O., Rodríguez Hernández, C. F., Castillo Lara, L. A., & Aldape Valdés, L. P. (2024). Improving the learning-teaching process through adaptive learning strategy. *Smart Learning Environments, 11*(1). https://doi.org/10.1186/s40561-024-00314-9

Serhan, M., Sprowls, M., Jackemeyer, D., Long, M., Perez, I. D., Maret, W., Tao, N., & Forzani, E. (2019). Total iron measurement in human serum with a smartphone. *AIChE Annual Meeting, Conference Proceedings*, 2019-Novem. https://doi.org/10.1109/JTEHM.2020.3005308

Talanquer, V. (2025). Exploring the Plurality of Chemical Modeling: Implications for Chemistry Teaching. *J. Chem. Educ., 102* (8), 3090-3095. https://doi.org/10.1021/acs.jchemed.5c00244

Üce, M., & Ceyhan, İ. (2019). Misconception in chemistry education and practices to eliminate them: Literature analysis. *Journal of Education and Training Studies, 7*(3), 202. https://doi.org/10.11114/jets.v7i3.3990

Vosniadou, S. (2019). The development of students' understanding of science. *Frontiers in Education*, 4(April), 1–6. https://doi.org/10.3389/feduc.2019.00032

Whatoni, A. S., & Sutrisno, H. (2022). Development of a learning module supported by augmented reality on chemical bonding material to improve interest and motivation of students learning for senior high school. *Jurnal Penelitian Pendidikan IPA, 8*(4), 2210–2218. https://doi.org/10.29303/jppipa.v8i4.2057

Zendler, A., & Greiner, H. (2020). The effect of two instructional methods on learning outcome in chemistry education: The experiment method and computer simulation. *Education for Chemical Engineers*, *30*, 9-19. https://doi.org/10.1016/j.ece.2019.09.001

Zhao, L., Wu, G., Shao, W. & Ma, X. (2024). Conceptual understanding and cognitive patterns construction for physical education teaching based on deep learning algorithms. *Scientific Reports*, 14(31409). https://doi.org/10.1038/s41598-024-83028-9