ienci

Investigações em Ensino de Ciências (ISSN: 1518-8795)
2026, vol. 31, n. 3, pp. 01-24
https://doi.org/10.22600/1518-8795.ienci/2026v31n3p01

PESQUISA EMPÍRICA

Web-Based Interactive Multimedia as A Catalyst for Science Process Skills and Scientific Literacy in Inquiry-Based Learning

Multimídia Interativa Baseada na Web como Catalisadora das Habilidades de Processos Científicos e do Letramento Científico na Aprendizagem Baseada em Investigação

Syarifah Rita Zahara [a], Sirry Alvina [b], Riska Imanda [b], Mellyzar Mellyzar [b]

[a] Department of Physics Education, Universitas Malikussaleh, Lhokseumawe, Indonesia; [b] Department of Chemistry Education, Universitas Malikussaleh, Lhokseumawe, Indonesia.

Abstract. This study examined the role of Web-Based Interactive Multimedia Learning (WIML) as a catalyst for enhancing students' Science Process Skills (SPS) and Scientific Literacy (SL) within inquiry-based chemistry learning on Green Chemistry topics. Grounded in inquiry learning theory and multimedia learning principles, a quasi-experimental non-equivalent control group design was employed with 157 senior high school students in North Aceh, Indonesia. Both groups engaged in equivalent inquiry cycles, while the experimental group was supported by WIML features, including interactive visualizations, simulations, inquiry prompts, and automated feedback. SPS were assessed through structured classroom observation using a validated seven-indicator rubric, and SL improvement was measured using normalized N-gain scores. The results showed that the experimental group demonstrated significantly higher SPS engagement (Mann–Whitney U, $p < 0.001$, $r = 0.71$) and greater SL gains (mean N-gain = 0.52) compared to the control group (0.40). These findings indicate that WIML enhances inquiry enactment, strengthens evidence-based reasoning, and supports more inclusive learning trajectories without altering inquiry fidelity. The study highlights the pedagogical value of multimedia-supported inquiry for promoting sustainable chemistry education.

Keywords: Web-based interactive multimedia, Science process skills, Scientific literacy, Inquiry-based learning, Green chemistry education.

Submetido em 10/01/2026
Aceito em 14/07/2026
Publicado em 16/09/2026

Resumo. Este estudo investigou o papel da Aprendizagem por Multimídia Interativa Baseada na Web (Web-Based Interactive Multimedia Learning – WIML) no desenvolvimento das Habilidades de Processos Científicos (Science Process Skills – SPS) e do Letramento Científico em aulas de Química baseadas em investigação, com foco em Química Verde. Fundamentado na teoria da aprendizagem por investigação e nos princípios da aprendizagem multimídia, adotou-se um delineamento quase experimental com grupo de controle não equivalente, envolvendo 157 estudantes do ensino médio em Aceh Norte, Indonésia. Ambos os grupos participaram de ciclos investigativos equivalentes, sendo que o grupo experimental contou com o suporte do WIML, incluindo visualizações interativas, simulações, prompts investigativos e feedback automatizado. As SPS foram avaliadas por observações estruturadas com rubrica validada de sete indicadores, e o letramento científico foi mensurado por meio do ganho normalizado (N-gain). Os resultados indicaram maior engajamento em SPS no grupo experimental (Mann–Whitney U, $p < 0,001$, $r = 0,71$) e ganhos superiores em letramento científico (N-gain médio = 0,52) em comparação ao grupo de controle (0,40). Os achados evidenciam que o WIML fortalece a aprendizagem investigativa, o raciocínio baseado em evidências e trajetórias de aprendizagem mais inclusivas, destacando seu valor pedagógico na educação em Química sustentável.

Palavras-chave: Multimídia interativa baseada na Web, Habilidades de processos científicos, Letramento científico, Aprendizagem baseada em investigação, Educação em Química Verde.

Contatos: syarifah.rita@unimal.ac.id; sirry.alvina@unimal.ac.id; riska.imanda@unimal.ac.id; mellyzar@unimal.ac.id

Introduction

Science Process Skills (SPS) and Scientific Literacy (SL) are widely recognized as essential competencies in chemistry education, supporting students' capacity to think critically, analyze empirical evidence, conduct systematic investigations, and apply scientific knowledge to real-world problems responsibly (Hikmah et al., 2018; Maison et al., 2019; Orosz et al., 2023; Tseng et al., 2022). SPS encompass both basic skills—such as observing, classifying, measuring, predicting, inferring, and communicating—and integrated inquiry competencies, including formulating hypotheses, designing experiments, interpreting multilevel chemical representations, evaluating evidence patterns, and constructing justified conclusions (Alvina et al., 2025; Daniel et al., 2025; Dewi & Kim, 2025; Hacıeminoğlu et al., 2022). SL, on the other hand, refers to students' ability to interpret science-related issues, evaluate competing evidence claims, and make informed decisions grounded in scientific warrants (OECD, 2023; Reyes et al., 2023).

Strengthening SPS and SL requires innovative pedagogical approaches, particularly inquiry-based learning (IBL), which organizes instruction through epistemic scientific cycles—posing questions, planning investigations, manipulating variables, interpreting evidence, constructing explanations, and reflecting on claims (Pedaste et al., 2015; Potier, 2023). IBL has been validated as a primary framework for developing SPS and SL in chemistry classrooms because it positions students as active agents who generate, test, and refine knowledge through evidence-driven reasoning rather than passive reception of information (Arantika et al., 2019; Georgiou & Kyza, 2023; Nedungadi et al., 2015).

Green Chemistry (GC) provides a particularly productive context for inquiry-based SPS and SL development at the senior high school level. GC is defined as the design of chemical products and processes that reduce or eliminate the generation and use of hazardous substances (Anastas & Warner, 1998), operationalized through twelve guiding principles including atom economy maximization, safer synthesis design, waste prevention, and biodegradable product engineering (Idul et al., 2025; Mashami et al., 2025). GC topics enable students to evaluate chemical processes using sustainability criteria in authentic environmental contexts, thereby connecting abstract chemical concepts with consequential real-world decision-making (Linkwitz & Eilks, 2022; Listyarini et al., 2019; López-Fernández et al., 2025).

Despite these affordances, persistent challenges remain. Teachers often lack preparedness for sustainability-integrated chemical inquiry, laboratory resources are frequently constrained, and GC content risks being positioned as declarative add-on knowledge rather than inquiry-investigable reasoning criteria (Linkwitz & Eilks, 2022; Quiroz-Martinez & Rushton, 2025; Widyantoro et al., 2025). These challenges highlight the need for supportive learning technologies that can bridge gaps between pedagogical intent and classroom enactment.

Web-Based Interactive Multimedia Learning (WIML) has emerged as a promising instructional medium for addressing these challenges. WIML refers to web-accessible digital learning environments that integrate multiple representational modalities—including animations, simulations, interactive visualizations, embedded videos, and linked symbolic-submicroscopic-macroscopic representations—to present dynamic chemical content and support real-time learning interactions (Hammer & Avram, 2024; Handoyo et al., 2024). Theoretically, WIML's pedagogical value is grounded in Mayer's (2014) Cognitive Theory of Multimedia Learning (CTML), which posits that learners process verbal and visual information through dual channels with limited capacity, and that well-designed multimedia can reduce extraneous cognitive load—the processing demands imposed by poor instructional design—while increasing germane cognitive load, which refers to the mental effort devoted to constructing and automating cognitive schemas (Sweller, 1988; Sweller et al., 2019). Extraneous load is reduced through principles such as spatial contiguity (placing related text and images near each other), signaling (using cues to highlight essential information), and segmenting (breaking complex content into manageable units).

In inquiry-based chemistry classrooms, WIML plays a strategic role as a catalyst for investigation rather than a replacement for the inquiry model itself. By providing accessible digital content, supporting observation and data interpretation, enabling hypothesis-driven reasoning through interactive prompts, and delivering automated real-time feedback, WIML strengthens the quality of students' inquiry enactment while students remain the primary agents of investigative activity (Al-Duhani et al., 2024; Batamuliza et al., 2024; Hinostroza et al., 2024). This catalytic function is consistent with the scaffolding literature, which demonstrates that inquiry is most effective when supported by well-timed scaffolds that reduce novice cognitive overload while preserving student-driven reasoning and explanation building (Moon et al., 2017; Szalay et al., 2023; Stiller & Wilde, 2025).

Although previous research has established the individual benefits of multimedia environments for chemistry learning (Hu et al., 2024; Peperkorn et al., 2024) and the effectiveness of inquiry pedagogy for SPS and SL development (Chengere et al., 2025; Georgiou & Kyza, 2023), limited empirical evidence exists regarding the combined impact of WIML-supported inquiry instruction on both SPS and SL within GC contexts at the secondary level. Furthermore, most existing studies have assessed SPS using paper-based procedural tests rather than structured behavioral observation during authentic inquiry activities, limiting ecological validity. This study addresses these gaps by investigating WIML's catalytic role in inquiry-based chemistry learning using performance-based SPS assessment and standardized SL measurement.

Research objective

This study aims to investigate the effectiveness of WIML as a catalytic support in inquiry-
This study aims to investigate the effectiveness of WIML as a catalytic support in inquiry-
based chemistry learning, with particular attention to students' SPS and SL in GC contexts. Specifically, this study seeks to: (a) examine the effect of WIML as a catalyst in inquiry-based chemistry learning on students' observable SPS engagement; (b) examine the impact of multimedia-supported inquiry learning on students' SL gains in GC topics; and (c) analyze the relationship between students' SPS engagement and their SL improvement in multimedia-supported versus conventional inquiry learning environments.

2026, vol. 31, n. 3, pp. 01-24

Research hypotheses

H1: The integration of WIML as a catalyst within inquiry-based chemistry learning leads to significantly stronger observable student engagement in SPS compared to inquiry learning without multimedia support.

H2: Students participating in multimedia-supported inquiry learning demonstrate higher normalized gains in SL than students in inquiry learning without multimedia support.

H3: Students' observable engagement in SPS is positively and significantly associated with their Scientific Literacy N-gain improvement in both instructional conditions, with stronger learning coherence emerging in the multimedia-supported inquiry environment.

Theoretical Background

Science Process Skills (SPS) in chemistry inquiry

SPS represent the cognitive and psychomotor competencies that scientists employ when investigating natural phenomena (Padilla, 1990). In chemistry education, SPS enable students to convert abstract chemical concepts into observable investigative actions, including generating researchable questions, designing evidence-based tests, controlling variables, interpreting data trends, drawing warranted conclusions, and communicating claims supported by empirical evidence (Sailaubay et al., 2024; Szalay et al., 2023). Contemporary frameworks conceptualize SPS along a developmental continuum from basic skills (observing, classifying, measuring, predicting, inferring, and communicating) to integrated inquiry skills (formulating hypotheses, designing experiments, interpreting multilevel chemical representations—that is, macroscopic observations, submicroscopic particle models, and symbolic equations—evaluating evidence patterns, and constructing justified conclusions) (Alvina et al., 2025; Daniel et al., 2025; Dewi & Kim, 2025; Senisum et al., 2022).

Recent intervention research has demonstrated that SPS development improves significantly when inquiry learning is supported through explicit scaffolding, structured investigation phases, and defensible rubric-based classroom observations that ensure inter-rater reliability and valid performative evidence (Hsu et al., 2015; Lin et al., 2023; Stiller & Wilde, 2025). Performance-based assessment through structured behavioral observation during authentic inquiry activities offers greater ecological validity than paper-based procedural tests, as it captures real-time inquiry enactment in naturalistic classroom settings (López-Fernández et al., 2025; Vo & Simmie, 2025).

Scientific Literacy (SL) in sustainable chemistry education

SL in the context of sustainable chemistry education refers to students' capacity to interpret sustainability-related chemistry problems, evaluate competing evidence claims, justify solution trade-offs, and make environmentally consequential decisions using chemical warrants such as toxicity, resource consumption, atom economy, waste generation, and life-cycle considerations (Araripe & Zuin Zeidler, 2024; Aydin Gunbatar et al., 2025; OECD, 2023; Reyes et al., 2023). The OECD's PISA framework defines SL across three core competencies: explaining phenomena scientifically, evaluating and designing scientific enquiry, and interpreting data and evidence scientifically (OECD, 2023). In GC-framed curriculum materials, these competencies are operationalized through tasks that require students to reason toward evidence-justified decisions about chemical processes and sustainability trade-offs rather than memorize sustainability slogans (Idul et al., 2025; Karpudewan & Kulandaisamy, 2018; Lembens et al., 2022; López-Fernández et al., 2025).

Empirical evidence supports the notion that socio-scientific inquiry-based learning approaches in chemistry education effectively enhance students' SL, critical chemical literacy, ethics-aligned reasoning, and decision justification competencies (Georgiou & Kyza, 2023; Mamlok-Naaman & Mandler, 2020). SL becomes more visible and measurable when students must defend decisions under real constraints through socio-scientific inquiry problems, positioning sustainable chemistry as a productive context for strengthening critical chemical literacy (Khuadthong et al., 2025) and ethics-aligned reasoning (Zidny et al., 2021).

Inquiry-based learning (IBL) as a core framework

IBL has been validated as a primary pedagogical framework for strengthening both SPS and SL in chemistry classrooms because it organizes learning through epistemic scientific cycles (Arantika et al., 2019; Georgiou & Kyza, 2023; Nedungadi et al., 2015). The inquiry cycle, as synthesized by Pedaste et al. (2015), comprises five general phases: orientation, conceptualization (questioning and hypothesis gereration), investigation (experimentation and data interpretation), conclusion, and discussion. Within this framework, SPS are treated as observable inquiry performances embedded in investigation cycles rather than as add-on competencies, which explains the growing reliance on performance-based tasks and structured classroom observations to capture real-time inquiry enactment.

The literature has clarified that inquiry is most effective when supported with well-timed scaffolds that reduce novice cognitive overload while preserving student-driven reasoning and explanation skills. Guided inquiry, in which teachers provide structured support while students retain investigative agency, improves students' capacity to reason with evidence, evaluate trade-offs, interpret representations, and justify scientific claims in sustainable contexts (Moon et al., 2017; Szalay et al., 2023).

Cognitive load theory and multimedia learning principles

The theoretical rationale for WIML's effectiveness as na inquiry catalyst draws primarily on two complementary cognitive frameworks. Sweller's (1988; 2019) Cognitive Load Theory (CLT) posits that human working memory has limited capacity and that instructional design should minimize extraneous cognitive load (processing demands arising from poor design), manage intrinsic load (complexity inherent to the content), and maximize germane load (effort directed toward schema construction and automation). In chemistry learning, intrinsic load is often high due to the need to coordinate multiple representational levels – macroscopic phenomena observed in laboratories, submicroscopic particle-level explanations, and symbolic chemical equations (Johnstone, 1991; Hu et al., 2024).

Mayer's (2014) Cognitive Theory of Multimedia Learning (CTML) complements CLT by specifying that meaningful learning occurs when learners engage in selecting, organizing, and integrating verbal and pictorial information through dual processing channels. CTML identifies several evidence-based design principles that reduce extraneous load and promote deeper learning, including the multimedia principle (combining words with pictures), the contiguity principle (placing related text and visuals together in space and time), the signaling principle (using cues to direct attention), and the segmenting principle (presenting information in learner-paced chunks rather than continuous streams). WIML environments that adhere to these principles can serve as procedural reasoning scaffolds—that is, supports that guide students through the logical steps of scientific investigation—by presenting dynamically linked multiple representations, reducing extraneous processing demands, and enabling iterative variable manipulation with real-time feedback (Al-Duhani et al., 2024; Batamuliza et al., 2024; Hinostroza et al., 2024; Peperkorn et al., 2024).

Pedagogical affordances of WIML

WIML has been increasingly validated as an inquiry-process amplifier that strengthens investigation quality through several affordances: web accessibility enabling anytime-anywhere learning, multimodal chemical visualization (animations, simulations, videos, and linked representations presenting the same phenomenon at macroscopic, submicroscopic, and symbolic levels simultaneously), iterative variable manipulation allowing students to test hypotheses virtually, digital data organization tools, and real-time actionable feedback that enhances learners' self-regulation during inquiry enactment (Al-Duhani et al., 2024; Batamuliza et al., 2024; Hinostroza et al., 2024). The term 'multimodal features' here refers to the integration of multiple sensory channels (visual, auditory, textual) and representational formats (static images, dynamic animations, interactive simulations, and symbolic notations) within a single learning environment.

Simulation-rich WIML environments improve students' capacity to extract evidence patterns, test hypotheses, interpret multilevel chemical representations, align data to claims, and build scientific explanations when supported through inquiry-aligned prompts (Hu et al., 2024; Peperkorn et al., 2024). Furthermore, inquiry-aligned multimedia feedback loops—defined as cyclical processes in which student actions within the multimedia generate immediate, contextually relevant information that students then use to adjust their reasoning—significantly strengthen procedural reasoning, evidence interpretation, and self-regulated investigation performance, particularly when feedback is timely, adaptive, and embedded within tasks (Hu et al., 2024; Peperkorn et al., 2024).

GC learning in senior high schools: relevance, challenges, and learning opportunities

GC learning at the senior high school level has gained increasing relevance as it provides a sustainability-oriented decision framework grounded in principles such as atom economy maximization, safer synthesis design, waste prevention, degradability, and biodegradable product engineering, enabling students to operationalize inquiry skills, interpret chemical evidence, evaluate alternative synthesis pathways, justify trade-offs, and defend sustainability-related claims with data-supported warrants (Idul et al., 2025; Mashami et al., 2025).

Despite strong pedagogical relevance, recent secondary-education studies have repeatedly reported persistent challenges such as limited teacher preparedness for sustainability-integrated chemical inquiry, constrained laboratory resources, and risks of GC being positioned as declarative add-on knowledge rather than inquiry-investigable reasoning criteria (Linkwitz & Eilks, 2022; Quiroz-Martinez & Rushton, 2025; Widyantoro et al., 2025).

The literature has consistently shown that contextualized inquiry-framed GC tasks—such as comparing alternative chemical pathways by hazards, waste, resource use, and biodegradation potential—effectively elicit SPS performances and strengthen measurable SL outcomes. Recent work has also proposed that localized GC modules and teacher professional development improve implementation defensibility, rater-guided scoring consistency, and student evidence-based decision justification quality in chemistry classrooms (López-Fernández et al., 2025).

Methods

Research design

This study employed a quasi-experimental non-equivalent control group pretest-posttest design (Campbell & Stanley, 1963) to evaluate the effect of WIML embedded within inquiry-based chemistry instruction on students' SPS and SL. In this design, pre-existing intact classes are assigned to experimental and control conditions without random assignment; both groups complete a pretest prior to the intervention and a posttest following it, permitting assessment of treatment effects through between-group comparisons of observed change while controlling statistically for baseline differences. The non-equivalent designation acknowledges that, in the absence of randomization, the two groups may differ on unmeasured background characteristics; group equivalence is therefore established empirically through pretest comparisons rather than assumed through random allocation. This design is appropriate for the present study because it preserves the natural composition of intact school classrooms – thereby maintaining ecological and instructional validity – while still enabling rigorous causal inference about the WIML treatment through pretest-controlled, between-group comparisons. The design is represented as follows:

$$\text{Experimental group:} \quad O_1 \quad X \quad O_2$$

$$\text{Control group:} \quad O_3 \quad – \quad O_4$$

where O_1 and O_3 represent pretest administrations, X represents the WIML intervention, and O_2 and O_4 represent posttest administrations. Both groups received equivalent inquiry-based instruction on GC topics; the experimental group was additionally supported by WIML, while the control group received the same inquiry procedures using conventional instructional materials (printed worksheets, textbook resources, and teacher-led demonstrations) without multimedia integration.

Research participants

The participants consisted of 157 first-semester students (aged 15-16 years) from three senior high schools in North Aceh, Indonesia. Six intact classrooms were selected using a non-randomized cluster sampling approach, preserving the natural class composition to maintain ecological validity. Each school contributed one experimental and one control class (3 schools × 2 classes = 6 classrooms). The final distribution was: Experimental-1 (n = 29), Experimental-2 (n = 25), Experimental-3 (n = 25), Control-1 (n = 27), Control-2 (n = 27), and Control-3 (n = 24), yielding a total of 79 experimental and 78 control participants. This intact-class configuration preserved instructional authenticity and minimized disruption to existing school learning schedules. Group equivalence was assessed through comparison of pretest SL scores using the Mann-Whitney U test, which confirmed no statistically significant difference between groups prior to intervention ($p > 0.05$), supporting the validity of subsequent between-group comparisons.

Development and validation of WIML

The WIML environment (shown in Figure 1) was developed prior to classroom implementation using the 4-D instructional design model (Define, Design, Develop, Disseminate) proposed by Thiagarajan et al. (1974). The development process comprised:

(a) needs analysis, including curriculum alignment with GC topics and identification of student learning difficulties;

(b) interface and content design following Mayer's (2014) multimedia learning principles (contiguity, signaling, segmenting, and coherence);

(c) expert validation by six content and pedagogical experts;

(d) small-scale pilot testing with 30 students from a non-participating school to evaluate usability and identify technical issues, and;

(e) final dissemination through an online learning platform.

The WIML platform integrated four key features:

(a) dynamic chemical visualizations presenting GC principles at macroscopic, submicroscopic, and symbolic levels;

(b) interactive simulations allowing students to manipulate variables in virtual experiments (e.g., comparing synthesis pathways by atom economy and waste generation);

(c) hypothesis-driven inquiry prompts embedded at critical investigation stages to scaffold reasoning transitions, and ;

(d) automated real-time feedback providing immediate responses to student inputs during inquiry activities.

The multimedia functioned as a cognitive and procedural catalyst—supporting students' reasoning, investigation, and scientific engagement while preserving the core structure of teacher-facilitated inquiry-based learning. The platform is accessible at: https://sites.google.com/view/weblearningipas.

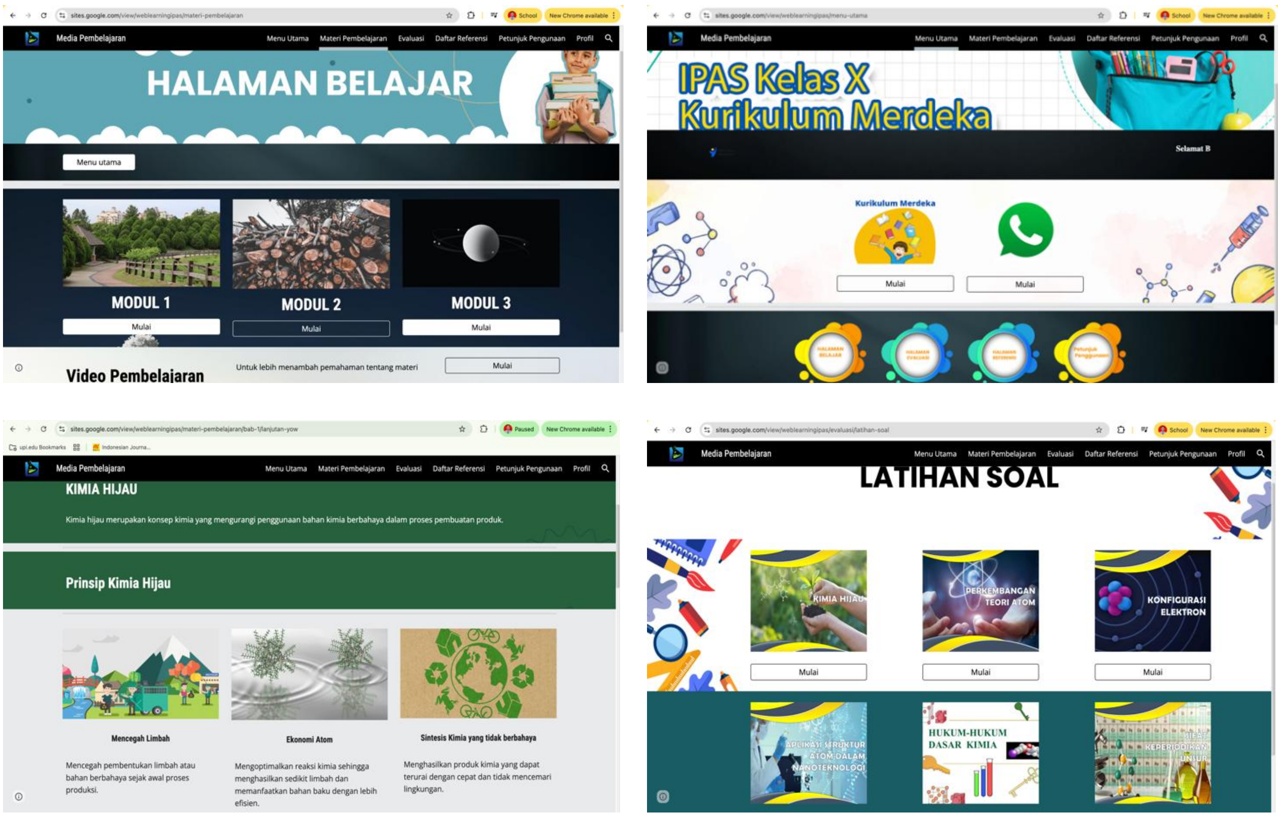

Figure 1. Overview of the WIML platform for GC instruction.

Research instruments

Two primary instruments were used for data collection: Science Concept Test (SL Measure) and SPS Observation Sheet.

Science Concept Test (SL Measure)

A paper-based SL test consisting of 10 contextual items was developed to measure students' ability to interpret evidence, apply inquiry reasoning, and make evidence-based decisions within GC inquiry contexts. The test assessed three SL competencies aligned with the OECD (2023) framework: explaining phenomena scientifically, evaluating and designing scientific enquiry, and interpreting data and evidence scientifically. Test content focused on sustainable chemical principles including atom economy, safer synthesis, biodegradable materials, waste reduction, and non-toxic formulations. All items were dichotomously scored (2 = correct, 0 = incorrect), yielding a maximum possible score of 20. The same test was administered as both pretest and posttest to enable N-gain calculation. The complete set of

10 SL test items, with corresponding OECD competency classifications and Green Chemistry content alignments, is provided in Supplementary Material S1.

A sample item illustrating the evidence-interpretation competency read: "A factory produces cleaning products using two alternative synthesis pathways. Pathway A generates 15% waste by mass while using a non-toxic solvent, whereas Pathway B generates 5% waste but requires a chlorinated solvent classified as hazardous. Based on Green Chemistry principles, which pathway is more sustainable, and what evidence supports your decision?" Students were required to evaluate trade-offs between waste minimization and hazard reduction, demonstrating integrated SL reasoning.

SPS Observation Sheet

Students' SPS were evaluated using a structured behavioral observation rubric covering seven inquiry indicators: observing, predicting, making hypotheses, conducting experiments, interpreting data, concluding, and communicating. Each indicator was rated using a four-point observational scale with clearly defined behavioral anchors: 1.00–<2.00 = low (minimal or no engagement in the target skill), 2.00–<3.00 = moderate (partial engagement with limited independence), and 3.00–4.00 = high (consistent, independent, and elaborated skill demonstration). Two trained observers independently recorded students' inquiry behaviors throughout the investigation sessions using a structured observation protocol to capture authentic SPS enactment during GC inquiry activities.

Instrument validation

Both instruments were validated by six experts comprising three chemistry education lecturers with doctoral qualifications and three senior high school chemistry teachers with more than 10 years of professional teaching experience. Content relevance at the item level was quantified using the Item-Level Content Validity Index (I-CVI) following established protocols (Lynn, 1986; Polit & Beck, 2006). Expert judgments confirmed high instrument validity, yielding mean I-CVI values of 0.93 for the science concept test and 0.96 for the observation sheet, both exceeding the recommended minimum threshold of 0.78 (Polit & Beck, 2006). For the SPS observation rubric, inter-rater reliability was assessed through parallel coding of a subsample by two trained observers, yielding Cohen's kappa (κ) values ranging from 0.78 to 0.91 across indicators, indicating substantial to almost perfect agreement (Landis & Koch, 1977). For the science concept test, internal consistency was evaluated using the Kuder-Richardson 20 (KR-20) formula, yielding a coefficient of 0.81, indicating good reliability.

Intervention procedure

The intervention was conducted over four weeks (eight 90-minute sessions) with the following standardized procedure applied across both conditions:

- Week 1 (sessions 1–2) - Pretest administration and orientation. All students completed the 10-item SL pretest under standardized conditions. Teachers introduced the GC inquiry framework, including the twelve principles of Green Chemistry, and demonstrated the inquiry cycle structure (orientation, conceptualization, investigation, conclusion, discussion). Experimental group students additionally received a guided introduction to the WIML platform, including navigation, simulation tools, and inquiry prompt features.
- Weeks 2–3 (sessions 3–6) - Inquiry-based GC investigation. Both groups engaged in a guided inquiry project requiring students to produce liquid soap using biodegradable, non-toxic ingredients while emphasizing low-waste formulations and avoidance of toxic additives. The inquiry cycles proceeded through: (a) observing properties of raw materials and formulations; (b) predicting outcomes based on GC principles; (c) formulating hypotheses about optimal sustainable formulations; (d) conducting hands-on experiments to test hypotheses; (e) interpreting data from experimental results; (f) drawing evidence-based conclusions; and (g) communicating findings through written and oral presentations. In the experimental group, each inquiry phase was supported by corresponding WIML features: interactive visualizations during observation, virtual simulations during hypothesis testing, embedded prompts during data interpretation, and automated feedback during conclusion formulation. The control group completed identical inquiry activities using conventional materials (printed GC worksheets, textbook resources, and teacher-led demonstrations). Throughout all sessions, trained observers recorded SPS behaviors using the structured observation rubric.
- Week 4 (sessions 7–8) - Final inquiry presentations and posttest. Students presented their GC inquiry findings. All students completed the identical 10-item SL posttest under the same standardized conditions as the pretest.

Data analysis

SL improvement was analyzed using Hake's (1998) normalized N-gain formula:

$$\textit{N-gain} = (\textit{posttest score} - \textit{pretest score}) / (\textit{maximum score} - \textit{pretest score})$$

N-gain values were categorized as low (<0.30), moderate (0.30–0.70), or high (>0.70) following Hake's (1998) classification. Preliminary normality testing using the Kolmogorov-Smirnov test with Lilliefors correction indicated non-normal distributions for both SPS scores and N-gain values ($p < 0.05$ in all cases). Accordingly, non-parametric statistical procedures were employed throughout. Between-group SPS differences were analyzed using the Mann-Whitney U test. The magnitude of instructional impact was quantified using the rank-biserial correlation coefficient ($r = Z/\sqrt{N}$), interpreted following Cohen's (1988) conventions: small ($r = 0.10$–0.30), medium ($r = 0.30$–0.50), and large ($r > 0.50$). To examine the relationship between students' observable SPS engagement and their SL gains, Spearman's rank correlation (ρ) was applied. A two-tailed significance threshold of $\alpha = 0.05$ was used for all inferential decisions. All statistical analyses were performed using SPSS Version 26.

Research ethics and permissions

This study was conducted in accordance with institutional research ethics standards. Ethical clearance was approved by the Research Ethics Committee of Universitas Malikussaleh. Official permissions for intervention and data collection were granted by all three participating schools (SMAN 1 Dewantara, SMAN 1 Muara Batu, and SMAN 1 Iskandar Muda).

Written informed consent was obtained from all participating students and their parents or guardians prior to data collection. Participation was voluntary, and students were informed of their right to withdraw at any time without academic penalty. Participant confidentiality and anonymity were maintained throughout data collection, analysis, and reporting.

Results

Distribution of students' SPS achievement

The distribution of students across SPS achievement levels differed markedly between groups. In the experimental group (n = 79), most students were classified in the moderate category (73.42%), followed by the high category (25.32%), with only 1.27% in the low category. In the control group (n = 78), the moderate category dominated (84.62%), while only 5.13% reached the high category and 10.26% remained in the low category.

The experimental group thus exhibited a more favorable achievement profile, characterized by a substantially higher proportion of students achieving high SPS levels and a notably smaller proportion demonstrating low SPS engagement.

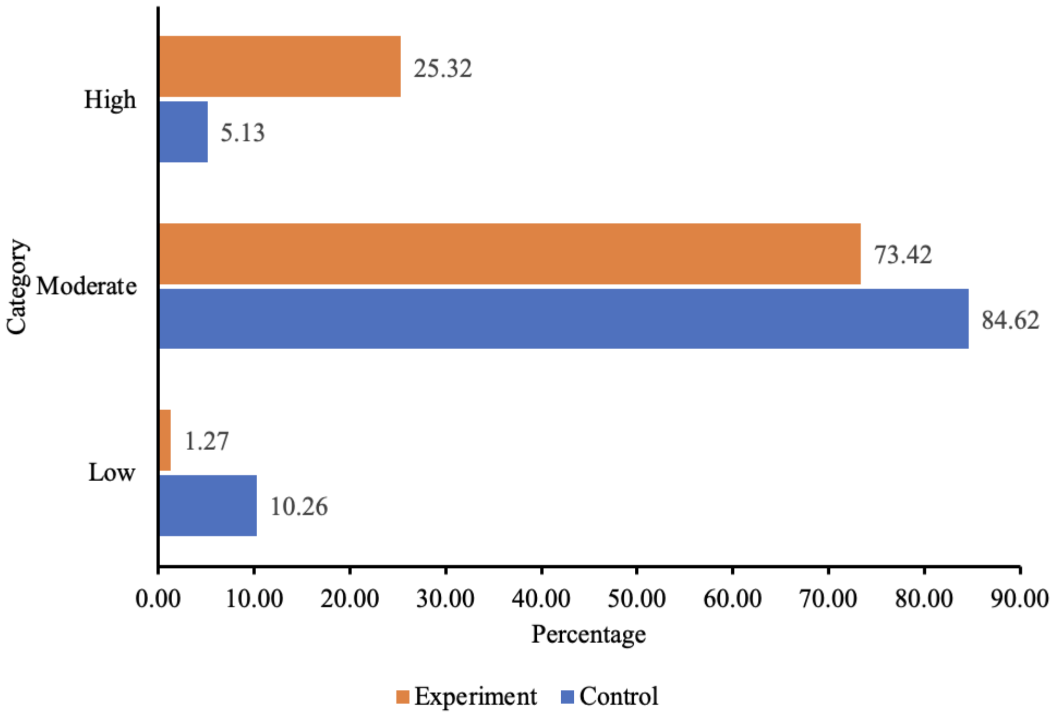

Figure 2. Distribution of students' SPS achievement levels.

Performance across SPS indicators

A comparison of mean scores across the seven SPS indicators revealed that the experimental group consistently outperformed the control group on all indicators.

The highest mean scores in the experimental group were recorded for observing (M = 3.28, SD = 0.48) and concluding (M = 3.28, SD = 0.52), followed by communicating (M = 3.20, SD = 0.55), reflecting strong performance in core inquiry and inference processes.

In the control group, SPS performance remained at moderate levels, with the lowest mean scores recorded for making hypotheses (M = 2.27, SD = 0.63) and interpreting data (M = 2.23, SD = 0.58). Standard deviation values (0.43–0.63) indicated relatively consistent SPS performance within each group.

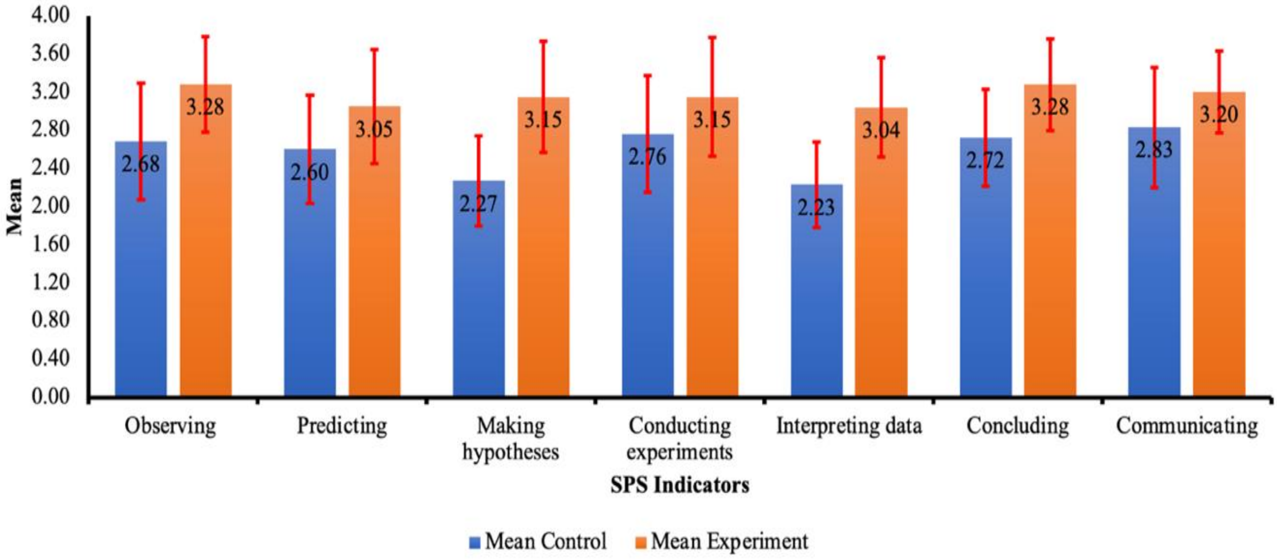

Figure 3. Mean scores of SPS indicators

Assumption testing and group comparison

Normality testing using the Kolmogorov–Smirnov test with Lilliefors correction revealed that SPS scores in both groups were not normally distributed (p = 0.000), as summarized in Table 1. A non-parametric comparison using the Mann–Whitney U test (Table 2) indicated a statistically significant difference in SPS performance between groups (U = 563.500, Z = −8.925, $p < 0.05$).

The calculated effect size (r = 0.71) demonstrated a large practical impact, suggesting that the multimedia-supported inquiry treatment meaningfully enhanced students' SPS.

Table 1. Kolmogorov–Smirnov normality test.

Variable	Class	Statistic	df	Sig. (p-value)
SPS Score	Control	0.242	78	0.000
	Experiment	0.165	79	0.000

a. Lilliefors Significance Correction

Table 2. Mann–Whitney test: test statistics.

Statistics[a]	SPS Score
Mann–Whitney U	563.500
Wilcoxon W	3644.500
Z	-8.925
Asymp. Sig. (2-tailed)	0.000

a. Grouping Variable: Class

SPS indicator interrelationships

Spearman correlation analysis within the experimental group (Figure 4) revealed multiple significant positive associations among SPS indicators, particularly at the investigation and inference stages. The strongest correlations were observed between conducting experiments and concluding ($\rho = 0.430$, $p < 0.001$), concluding and communicating ($\rho = 0.440$, $p < 0.001$), and observing and conducting experiments ($\rho = 0.397$, $p < 0.001$). However, several early-stage reasoning transitions showed non-significant or negligible relationships, including predicting and making hypotheses ($\rho = -0.107$, $p = 0.352$) and observing and interpreting data ($\rho = -0.200$, $p = 0.079$), indicating a segmental rather than fully cyclical SPS progression pattern. In the control group (Figure 5), SPS indicator correlations were fewer and largely non-sequential. Significant relationships appeared between predicting and communicating ($\rho = 0.530$, $p < 0.001$), conducting experiments and concluding ($\rho = 0.472$, $p < 0.001$), and making hypotheses and concluding ($\rho = 0.441$, $p < 0.001$). Most foundational inquiry links were not significant, such as observing and communicating ($\rho = 0.058$, $p = 0.615$) and interpreting data and concluding ($\rho = 0.039$, $p = 0.731$). This pattern suggests that SPS competencies in the control group developed independently rather than as connected inquiry phases.

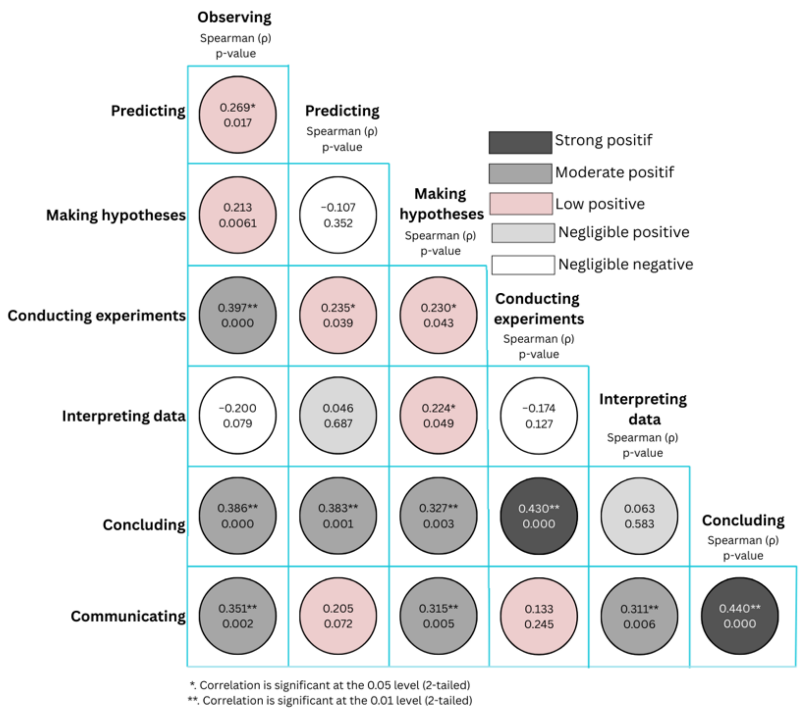

Figure 4. Spearman correlation matrix of SPS indicators (experimental class).

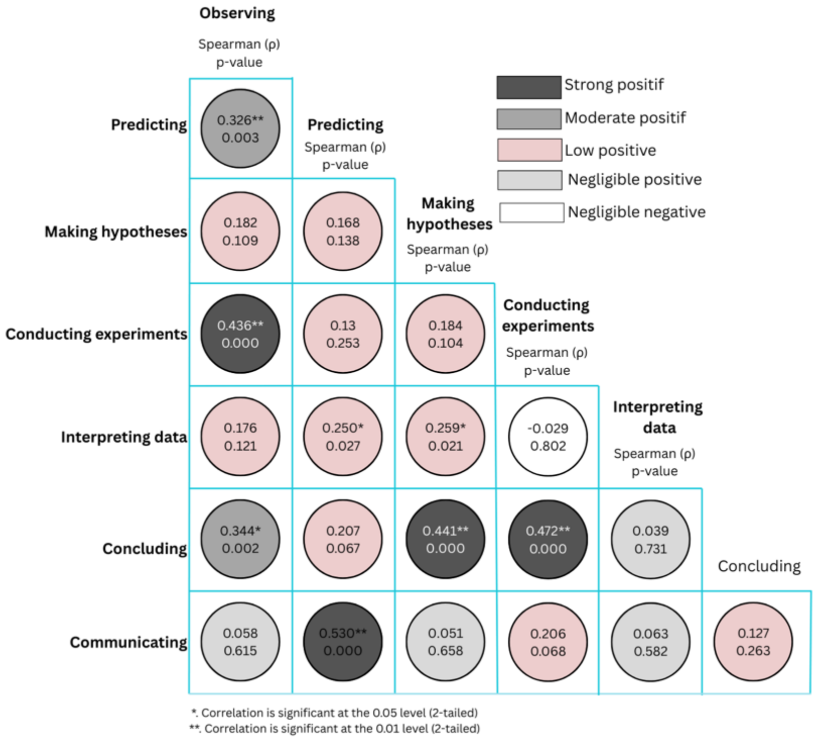

Figure 5. Spearman correlation matrix of SPS indicators (control class).

SL Gains

Both groups demonstrated moderate improvement in SL, with higher gains observed in the experimental group (Table 3; Figure 6). The experimental group achieved a mean N-gain of 0.52 (SD = 0.18) compared to the control group's mean N-gain of 0.40 (SD = 0.12). Although both fell within the moderate gain category per Hake's (1998) classification, the experimental group showed a substantially broader distribution toward high-gain achievement: 13.92% of experimental students reached the high N-gain category compared to only 1.28% of control students. Conversely, a larger proportion of control students fell in the low N-gain category (23.08%) compared to the experimental group (11.39%).

The Mann-Whitney U test confirmed a statistically significant difference in SL N-gain between groups ($U = 1807.00$, $Z = -4.492$, $p < 0.001$), with a moderate effect size ($r = 0.36$), indicating that the instructional treatment had an educationally meaningful influence on SL development.

Table 3. Descriptive statistics of SL N-gain

Group	N	Mean N-gain	SD	Category
Control	78	0.40	0.12	Moderate
Experiment	79	0.52	0.18	Moderate

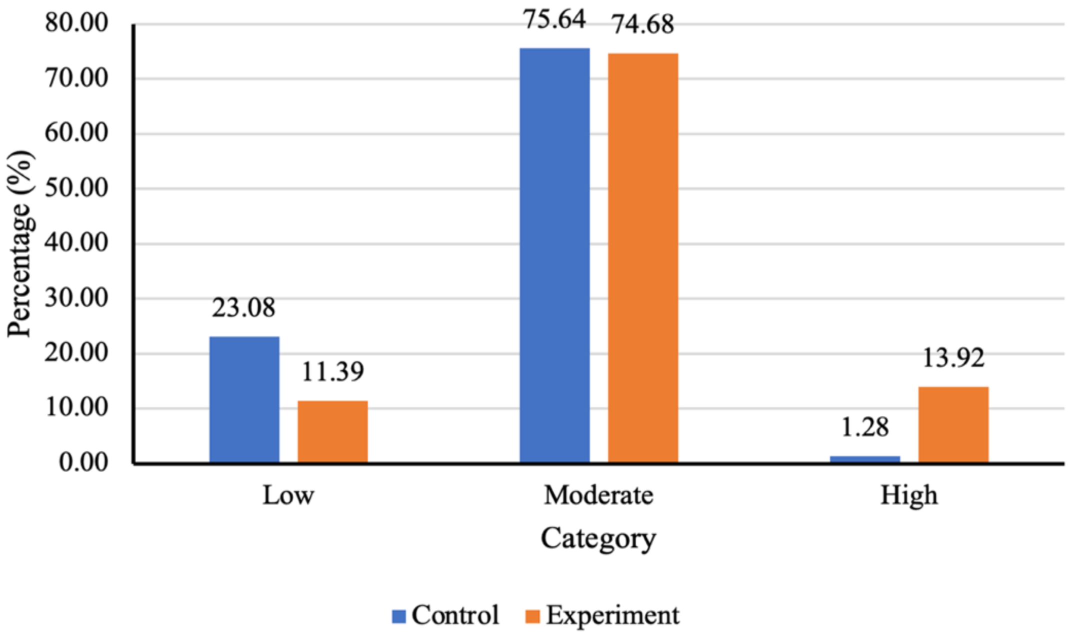

Figure 6. Distribution of N-gain categories.

Relationship between SPS and SL N-gain

Normality of gain scores was evaluated using the Kolmogorov–Smirnov test with Lilliefors correction, which indicated significant deviations from normality in both the control ($p = 0.004$) and experimental group ($p = 0.011$), confirming non-normal SPS distributions (Table 4). Accordingly, group differences were examined using a non-parametric Mann–Whitney U test, revealing a statistically significant difference in SPS performance between the experimental and control groups ($U = 1807.00$, $Z = -4.492$, $p < 0.001$; Table 5). The effect size ($r = 0.36$) suggested a moderate practical impact, indicating that the instructional treatment had an educationally meaningful influence on SPS development.

Table 4. Kolmogorov–Smirnov normality test.

Variable	Class	Statistic	df	Sig. (p-value)
N-Gain	Control	0.124	78	0.004
	Experiment	0.115	79	0.011

a. Lilliefors Significance Correction

Table 5. Mann–Whitney test: test statistics.

Statistics[a]	SPS Score
Mann–Whitney U	1807.000
Wilcoxon W	4888.000
Z	-4.492
Asymp. Sig. (2-tailed)	0.000

b. Grouping Variable: Class

Subsequently, Spearman's rank correlation analysis demonstrated significant positive associations between SPS performance and SL gains (N-gain) in both groups (Table 6). The experimental class showed a moderate positive correlation ($\rho = 0.354$, $p = 0.001$), indicating that students with stronger engagement in SPS tended to achieve greater literacy gains. The control class also exhibited a stronger moderate positive correlation ($\rho = 0.486$, $p < 0.001$), suggesting that although SPS was linked to N-gain in both conditions, the strength of the association differed between groups. Overall, these results confirmed that SPS performance was significantly associated with improvements in SL, regardless of instructional condition, albeit with varying magnitudes.

Table 6. Spearman's correlation between SPS and SL N-gain.

Class	N	Spearman ρ	p-value	Interpretation
Experiment	79	0.354**	0.001	Significant
Control	78	0.486**	0.000	Significant

**. Correlation is significant at the 0.01 level (2-tailed).

Discussion

WIML's Impact on SPS Engagement (H1)

The results provide strong support for H1, demonstrating that students in the multimedia-supported inquiry condition exhibited significantly higher SPS engagement than their counterparts in conventional inquiry instruction. The large effect size ($r = 0.71$) indicates a substantively meaningful difference that extends beyond statistical significance. Since SPS were assessed through structured behavioral observation during authentic inquiry activities rather than retrospective paper-based tests, the findings reflect students' actual enactment of inquiry behaviors, offering stronger ecological validity than procedural knowledge tests (López-Fernández et al., 2025; Vo & Simmie, 2025).

These findings are consistent with CLT predictions that well-designed multimedia scaffolds reduce extraneous cognitive load by presenting information through complementary channels, freeing working memory resources for germane processing (Sweller et al., 2019). The WIML features—interactive visualizations, embedded task prompts, guided inquiry sequencing, and immediate feedback—likely strengthened students' performance in observing patterns, synthesizing evidence, forming conclusions, and communicating explanations. These indicators align with inquiry phases previously identified as most responsive to structured scaffolds and cognitively rich multimedia interaction (Ajayan et al., 2025; Hu et al., 2024). The high mean scores for observing and concluding ($M = 3.28$) in the experimental group suggest that the WIML environment particularly enhanced students' capacity at inquiry entry and exit points, where evidence processing and reasoning externalization are most cognitively demanding.

Despite the overall improvement, the segmental SPS progression pattern—with weak or negative correlations between early-stage reasoning transitions (predicting–hypothesis formulation; observing–data interpretation)—deserves careful interpretation. This pattern suggests that while WIML effectively scaffolded investigation and inference phases, it did not automatically bridge the cognitive transitions between predictive reasoning and formal hypothesis generation. This finding aligns with prior research indicating that these transitions require explicit teacher-led epistemic prompts and sustained reasoning scaffolds rather than technology-mediated support alone (Szalay et al., 2023; Tseng et al., 2022). The control group's even more fragmented SPS inter-indicator relationships further confirm that inquiry learning without multimedia support maintains moderate procedural engagement but with limited inter-indicator connectivity and fewer opportunities for iterative reasoning feedback (Chen & Chu, 2024; Orosz et al., 2023).

WIML's Impact on SL Gains (H2)

H2 was supported: the experimental group achieved significantly higher SL N-gains than the control group, with a moderate effect size (r = 0.36). The pattern of results—particularly the substantially larger proportion of experimental students reaching high N-gain levels (13.92% vs. 1.28%) and the smaller proportion remaining at low gain levels (11.39% vs. 23.08%) – indicates that WIML scaffolds supported more inclusive learning trajectories that benefited students across the achievement spectrum rather than only high-achieving students.

From a CTML perspective (Mayer, 2014), the multimedia environment's dynamically linked multiple representations likely helped students coordinate macroscopic observations, submicroscopic particle models, and symbolic chemical equations, thereby reducing the split-attention effect and enabling deeper conceptual integration. This interpretation aligns with prior research documenting that multimedia-supported inquiry environments help learners iteratively explore abstract science concepts, visualize inquiry phenomena, and replay reasoning processes when needed, enabling differentiated cognitive trajectories (Ajayan et al., 2025; Al-Duhani et al., 2024; Lai et al., 2019). The moderate overall effect size (r = 0.36) for SL gains, though smaller than the effect on SPS (r = 0.71), is educationally meaningful and suggests that SL development is a gradual, cumulative process influenced by multiple factors including learning duration, cognitive feedback richness, and reflection quality (Chengere et al., 2025; Mutlu, 2020).

SPS-SL Relationship (H3)

H3 was partially supported. SPS engagement was positively and significantly correlated with SL N-gain in both instructional conditions, confirming that inquiry process competencies and conceptual literacy outcomes are interrelated dimensions. However, contrary to the hypothesized expectation, the correlation was stronger in the control group ($\rho = 0.486$) than in the experimental group ($\rho = 0.354$). This counterintuitive finding can be interpreted through two complementary lenses.

First, the stronger SPS-SL correlation in the control group may reflect a more linear benefit pattern in conventional inquiry classrooms, where students with pre-existing procedural and reasoning fluency benefited more consistently—and where fewer additional scaffolding resources were available to moderate the SPS-SL pathway. Second, the weaker correlation in the experimental group may reflect the WIML environment's introduction of additional pedagogical affordances—such as collaborative reasoning opportunities, multi-step task interaction, and cognitively demanding evidence synthesis – that diversified learning trajectories and reduced reliance on a single bivariate relationship. In other words, the multimedia environment may have created alternative pathways to SL improvement beyond direct SPS engagement, thereby attenuating the bivariate correlation while producing higher absolute SL gains overall. This interpretation supports the view that SPS and SL are interacting but non-deterministic dimensions: SPS serve as enabling mechanisms that facilitate literacy growth but are not its sole predictors (Georgiou & Kyza, 2023).

Limitations

Several limitations should be considered when interpreting these findings. First, the quasi-experimental design with non-randomized intact classrooms limits causal inference; although pretest equivalence was established, unmeasured confounding variables (e.g., prior technology experience, teacher effects) cannot be fully controlled. Second, the four-week intervention period, while sufficient to detect group differences, may not capture the full developmental trajectory of SPS cyclicality and SL growth. Third, the study was conducted in a specific geographic and cultural context (North Aceh, Indonesia), which may limit generalizability to other educational settings. Fourth, although inter-rater reliability for SPS observation was acceptable, observer effects (Hawthorne effect) cannot be entirely excluded. Fifth, the SL test comprised 10 dichotomously scored items, which may constrain the sensitivity of the measure to detect nuanced literacy differences. Future research should consider using polytomous scoring rubrics and larger item pools to improve measurement precision.

Conclusion

This study demonstrates that WIML, when integrated as a catalytic scaffold within inquiry-based chemistry instruction on GC topics, significantly strengthens students' observable SPS engagement (with a large effect) and produces meaningfully higher SL gains (with a moderate effect) compared to equivalent inquiry instruction without multimedia support. SPS engagement and SL improvement were positively correlated in both conditions, though the relationship was partially moderated by the multimedia environment's diversification of learning pathways.

Implications for Teaching Practice

The findings offer several actionable implications for chemistry educators. First, WIML should be positioned as a complement to—not a replacement for—structured inquiry pedagogy, with teachers maintaining their role as epistemic facilitators who bridge reasoning transitions that technology alone cannot scaffold effectively. Second, teachers should embed adaptive WIML scaffolds with sustained reasoning loops across inquiry phases, particularly at the predicting-to-hypothesizing and observing-to-interpreting transitions where students showed the weakest connections. Third, iterative feedback checkpoints should be incorporated into multimedia tasks to require students to revise hypotheses using GC sustainability warrants. Fourth, collaborative reasoning scripts within multimedia tasks can support diverse cognitive trajectories and promote more inclusive learning outcomes.

Implications for Future Research

Future research should consider: (a) randomized controlled designs with larger samples across diverse geographic and cultural contexts to strengthen causal inference and generalizability; (b) longitudinal designs to track the developmental trajectory of SPS cyclicality and SL growth over extended periods; (c) process-level analyses using learning analytics, screen recordings, or think-aloud protocols to examine how students interact with specific WIML features during inquiry phases; (d) investigation of teacher professional development models that integrate WIML with explicit epistemic prompting strategies; (e) expansion of SL measurement to include polytomous scoring and a broader range of GC sustainability competencies; and (f) comparative studies examining different levels of WIML scaffolding (unstructured, guided, and structured) to identify optimal support configurations for different learner profiles.

Acknowledgments

The authors express their sincere gratitude to all participating students, teachers, and school principals for their cooperation and support during the implementation of this research. This work was funded by the Non-Tax State Revenue (PNBP) of Universitas Malikussaleh for the 2023 fiscal year (Reference Code: 23.07.FKIP.03).

References

Ajayan, P., Cao, R., & Eichler, J. F. (2025). Promoting Conceptual Learning Using Scaffolded Activities That Incorporate Interactive Simulations. *Education Sciences*, *15*(5), 1-15. https://doi.org/10.3390/educsci15050566

Al-Duhani, F., Mohd Saat, R., & Abdullah, M. N. S. (2024). Effectiveness of web-based virtual laboratory on grade eight students' self-regulated learning. *Eurasia Journal of Mathematics, Science and Technology Education*, *20*(3), 1-20. https://doi.org/10.29333/ejmste/14282

Alvina, S., Mellyzar, M., & Rita Zahara, S. (2025). Hands-On, Minds-On: The Impact of Home-Based Experiments on Scientific Process Skills of Indonesian Middle School Students. *Science Education International*, *36*(3), 341–348. https://doi.org/10.33828/sei.v36.i3.9

Anastas, P. T., & Warner, J. C. (2000). Green chemistry: Theory and practice. Oxford University Press. https://doi.org/10.1093/oso/9780198506980.001.0001

Arantika, J., Saputro, S., & Mulyani, S. (2019). Effectiveness of guided inquiry-based module to improve science process skills. *Journal of Physics: Conference Series*, *1157*, 042019. https://doi.org/10.1088/1742-6596/1157/4/042019

Araripe, E., & Zuin Zeidler, V. G. (2024). Advancing sustainable chemistry education: Insights from real-world case studies. *Current Research in Green and Sustainable Chemistry*, *9*, 100436. https://doi.org/10.1016/j.crgsc.2024.100436

Aydin Gunbatar, S., Ekiz Kiran, B., Boz, Y., & Oztay, E. S. (2025). A systematic review of green and sustainable chemistry training research with pedagogical content knowledge framework: current trends and future directions. *Chemistry Education Research and Practice*, *26*(1), 34–52. https://doi.org/10.1039/D4RP00166D

Batamuliza, J., Habinshuti, G., & Nkurunziza, J. B. (2024). Integration of interactive computer simulations in teaching and learning chemical reaction: Students' performance and concept retention. *Journal of Technology and Science Education*, *14*(4), 1060. https://doi.org/10.3926/jotse.2682

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Houghton Mifflin.

Chen, C.-H., & Chu, Y.-R. (2024). VR-assisted inquiry-based learning to promote students' science learning achievements, sense of presence, and global perspectives. *Education and Information Technologies*, *29*(15), 19421–19441. https://doi.org/10.1007/s10639-024-12620-3

Chengere, A. M., Bono, B. D., Zinabu, S. A., & Jilo, K. W. (2025). Enhancing secondary school students' science process skills through guided inquiry-based laboratory activities in biology. *PLOS ONE*, *20*(4), e0320692. https://doi.org/10.1371/journal.pone.0320692

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Daniel, A., Gebeyhu, D., Assefa, S., & Abate, T. (2025). The holistic effect of nature of science and science process skills on students' conceptual and procedural knowledge and motivation within the context of modified guided discovery in physics laboratory. *Eurasia Journal of Mathematics, Science and Technology Education*, *21*(10), em2708. https://doi.org/10.29333/ejmste/17035

Dewi, N. W. R. U., & Kim, M. (2025). Basic Inquiry and Integrated Science Process Skills in Digital-Related Experiments Within Physics Textbooks of South Korea, The United Kingdom, and Indonesia. *Journal of Baltic Science Education*, *24*(5), 860–877. https://doi.org/10.33225/jbse/25.24.860

Georgiou, Y., & Kyza, E. A. (2023). Fostering Chemistry Students' Scientific Literacy for Responsible Citizenship through Socio-Scientific Inquiry-Based Learning (SSIBL). *Sustainability*, *15*(8), 6442. https://doi.org/10.3390/su15086442

Hacıeminoğlu, E., Yıldız, N. G., & Şeker, R. (2022). Factors Related to Cognitive Reasoning of Pre-Service Teachers' Science Process Skills: Role of Experiments at Home on Meaningful Learning. *Sustainability*, *14*(13), 7703. https://doi.org/10.3390/su14137703

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data. American Journal of Physics, *66*(1), 64–74. https://doi.org/10.1119/1.18809

Hamizi, M. A. A. M., & Mokmin, N. A. M. (2025). Enhancing secondary biology learning through guided and problem-based virtual reality: effects on interest, concentration, and performance. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13840-x

Hammer, M., & Avram, E. M. G. (2024). Online Interactive Activity: Using a Web-Based Multimedia Activity to Teach Balancing Chemical Equations. *Journal of Chemical Education*, *101*(10), 4510–4516. https://doi.org/10.1021/acs.jchemed.4c00786

Handoyo, K. J., Wisnuwardana, C. J., Austen, A., & Permana, F. (2024). Molecule World: Enhancing chemistry education through web-based augmented reality using Assemblr. *Procedia Computer Science*, *245*, 1249–1258. https://doi.org/10.1016/j.procs.2024.10.354

Hikmah, N., Yamtinah, S., Ashadi, & Indriyanti, N. Y. (2018). Chemistry teachers' understanding of science process skills in relation of science process skills assessment in chemistry learning. *Journal of Physics: Conference Series*, *1022*, 012038. https://doi.org/10.1088/1742-6596/1022/1/012038

Hinostroza, J. E., Armstrong-Gallegos, S., & Villafaena, M. (2024). Roles of digital technologies in the implementation of inquiry-based learning (IBL): A systematic literature review. *Social Sciences & Humanities Open*, *9*, 100874. https://doi.org/10.1016/j.ssaho.2024.100874

Hsu, C.-C., Chiu, C.-H., Lin, C.-H., & Wang, T.-I. (2015). Enhancing skill in constructing scientific explanations using a structured argumentation scaffold in scientific inquiry. *Computers & Education*, *91*, 46–59. https://doi.org/10.1016/j.compedu.2015.09.009

Hu, B., Zhu, L., & Bi, H. (2024). Effect of Computer Simulations on Student Ability to Translate Chemical Representations When Learning the Particulate Nature of Matter Concept. *Journal of Chemical Education*, *101*(10), 4149–4160. https://doi.org/10.1021/acs.jchemed.4c00964

Idul, J. J. A., Jaculbe, Q. M. I., Lucine, N. S., Canama, M. D., Galve, M. D., Sayson, C. L., Suico, M. P., Santiago, K. Z. D., & Walag, A. M. P. (2025). Green Modules: Integrating Green and Sustainable Chemistry Principles to Secondary Chemistry Modules through Process-Oriented Guided Inquiry Learning. *Journal of Chemical Education*, *102*(3), 1104–1116. https://doi.org/10.1021/acs.jchemed.4c01360

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. Journal of Computer Assisted Learning, 7(2), 75–83. https://doi.org/10.1111/j.1365-2729.1991.tb00230.x

Karpudewan, M., & Kulandaisamy, Y. (2018). Malaysian teachers' insights into implementing green chemistry experiments in secondary schools. *Current Opinion in Green and Sustainable Chemistry*, *13*, 113–117. https://doi.org/10.1016/j.cogsc.2018.06.015

Khuadthong, B., Rattanapong, T., Imjai, N., & Aujirapongpan, S. (2025). Sustainable tourism awareness in Thai Gen Z tourists: Roles and themes associated with scientific literacy and sociological and critical thinking skills. *Journal of Hospitality and Tourism Management*, *64*, 101316. https://doi.org/10.1016/j.jhtm.2025.101316

Lai, A., Chen, C., & Lee, G. (2019). An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *British Journal of Educational Technology*, *50*(1), 232–247. https://doi.org/10.1111/bjet.12716

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lembens, A., Heinzle, G., Tepla, A., Maulide, N., Preinfalk, A., Kaiser, D., & Spitzer, P. (2022). SpottingScience – a digital learning environment to introduce Green Chemistry to secondary students and the public. *Chemistry Teacher International*, *4*(2), 143–154. https://doi.org/10.1515/cti-2021-0025

Lin, X.-F., Hwang, G.-J., Wang, J., Zhou, Y., Li, W., Liu, J., & Liang, Z.-M. (2023). Effects of a contextualised reflective mechanism-based augmented reality learning model on students' scientific inquiry learning performances, behavioural patterns, and higher order thinking. *Interactive Learning Environments*, *31*(10), 6931–6951. https://doi.org/10.1080/10494820.2022.2057546

Linkwitz, M., & Eilks, I. (2022). An Action Research Teacher's Journey while Integrating Green Chemistry into the High School Chemistry Curriculum. *Sustainability*, *14*(17), 10621. https://doi.org/10.3390/su141710621

Listyarini, R. V., Pamenang, F. D. N., Harta, J., Wijayanti L.W., Asy'ari M., & Lee W. (2019). The Integration of Green Chemistry Principles into Small Scale Chemistry Practicum for Senior High School Students. *Jurnal Pendidikan IPA Indonesia*, *8*(3), 371–378. https://doi.org/10.15294/jpii.v8i3.19250

López-Fernández, M. D. M., Cano-Iglesias, M. J., & Franco-Mariscal, A. J. (2025). Chemistry inquiry conducted by secondary school students into material degradation in the context of sustainability. *RSC Sustainability*, *3*(9), 3997–4019. https://doi.org/10.1039/D5SU00176E

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382–386.

Maison, Darmaji, Astalini, Kurniawan, D. A., & Indrawati, P. S. (2019). Science Process Skills and Motivation. *Humanities & Social Sciences Reviews*, *7*(5), 48–56. https://doi.org/10.18510/hssr.2019.756

Mamlok-Naaman, R., & Mandler, D. (2020). *Education for Sustainable Development in High School through Inquiry-Type Socio-Scientific Issues* (p. 69–78). https://doi.org/10.1021/bk-2020-1344.ch006

Mashami, R. A., Ahmadi, & Pahriah. (2025). Green chemistry and cultural wisdom: A pathway to improving scientific literacy among high school students. *Social Sciences & Humanities Open*, *11*, 101653. https://doi.org/10.1016/j.ssaho.2025.101653

Mayer, R. E. (2021). The Cambridge handbook of multimedia learning (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781108894333

Moon, A., Stanford, C., Cole, R., & Towns, M. (2017). Analysis of inquiry materials to explain complexity of chemical reasoning in physical chemistry students' argumentation. *Journal of Research in Science Teaching*, *54*(10), 1322–1346. https://doi.org/10.1002/tea.21407

Mutlu, A. (2020). Evaluation of students' scientific process skills through reflective worksheets in the inquiry-based learning environments. *Reflective Practice*, *21*(2), 271–286. https://doi.org/10.1080/14623943.2020.1736999

Nedungadi, P., Malini, P., & Raman, R. (2015). *Inquiry Based Learning Pedagogy for Chemistry Practical Experiments Using OLabs* (p. 633–642). https://doi.org/10.1007/978-3-319-11218-3_56

OECD. (2023). PISA 2022 assessment and analytical framework. OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en

Orosz, G., Németh, V., Kovács, L., Somogyi, Z., & Korom, E. (2023). Guided inquiry-based learning in secondary-school chemistry classes: a case study. *Chemistry Education Research and Practice*, *24*(1), 50–70. https://doi.org/10.1039/D2RP00110A

OECD. (2023). PISA 2022 assessment and analytical framework. OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en

Pedaste, M., et al. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. Educational Research Review, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

Peperkorn, Y., Buschmann, J.-K., & Schwedler, S. (2024). Comparing drawing tasks and elaborate single-choice questions in simulation-based learning: how do they facilitate students' conceptual understanding on chemical equilibria? *Chemistry Education Research and Practice*, *25*(4), 1030–1051. https://doi.org/10.1039/D3RP00113J

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Potier, D. N. (2023). The Use of Guided Inquiry to Support Student Progress and Engagement in High School Chemistry. *Journal of Chemical Education*, *100*(2), 1033–1038. https://doi.org/10.1021/acs.jchemed.2c00793

Quinn, F., Elliott, S., Taylor, N., & Littledyke, M. (2015). Education for Sustainability in Primary Science Education. In *Educating for Sustainability in Primary Schools* (p. 91–119). SensePublishers. https://doi.org/10.1007/978-94-6300-046-8_6

Quiroz-Martinez, D., & Rushton, E. A. C. (2025). Teaching high school chemistry through education for sustainability: a collaborative case study from Chile. *Environmental Education Research*, 1–18. https://doi.org/10.1080/13504622.2025.2573145

Reyes, K. M. D., Bruce, K., & Shetranjiwalla, S. (2023). Green Chemistry, Life Cycle Assessment, and Systems Thinking: An Integrated Comparative-Complementary Chemical Decision-Making Approach. *Journal of Chemical Education*, *100*(1), 209–220. https://doi.org/10.1021/acs.jchemed.2c00647

Sailaubay, A., Myrzakhmetova, N., Tunçel, M., & Kishibayev, K. (2024). Analysis of the Relationships Between University Students' Laboratory Self-Efficacy Beliefs, Science Process Skills and Achievement in Chemistry Courses. *Journal of Curriculum Studies Research*, *6*(2), 36–51. https://doi.org/10.46303/jcsr.2024.1

Senisum, M., Susilo, H., Suwono, H., & Ibrohim. (2022). GIReSiMCo: A Learning Model to Scaffold Students' Science Process Skills and Biology Cognitive Learning Outcomes. *Education Sciences*, *12*(4), 228. https://doi.org/10.3390/educsci12040228

Stiller, C., & Wilde, M. (2025). Full-Structured or Supported by Incremental Scaffolds? Effects on Perceived Competence and Motivation. *The Journal of Experimental Education*, *93*(2), 239–260. https://doi.org/10.1080/00220973.2023.2269128
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, *12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7
Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. Educational Psychology Review, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5
Szalay, L., Tóth, Z., Borbás, R., & Füzesi, I. (2023). Scaffolding of experimental design skills. *Chemistry Education Research and Practice*, *24*(2), 599–623. https://doi.org/10.1039/D2RP00260D
Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional Development for Training Teachers of Exceptional Children: A Sourcebook*. Indiana University, Bloomington: Center for Innovation in Teaching the Handicapped.
Tseng, Y.-J., Hong, Z.-R., & Lin, H. (2022). Advancing students' scientific inquiry performance in chemistry through reading and evaluative reflection. *Chemistry Education Research and Practice*, *23*(3), 616–627. https://doi.org/10.1039/D1RP00246E
Vo, D. Van, & Simmie, G. M. (2025). Assessing Scientific Inquiry: A Systematic Literature Review of Tasks, Tools and Techniques. *International Journal of Science and Mathematics Education*, *23*(4), 871–906. https://doi.org/10.1007/s10763-024-10498-8
Widyantoro, C., Han, J. Y., Ong, J. S. H., Goh, K. H., & Fung, F. M. (2025). Teaching Sustainability through Green Chemistry: An Experiential Learning Approach. *Journal of Chemical Education*, *102*(7), 2743–2754. https://doi.org/10.1021/acs.jchemed.4c01476
Zidny, R., Laraswati, A. N., & Eilks, I. (2021). A Case Study on Students' Application of Chemical Concepts and Use of Arguments in Teaching on the Sustainability-Oriented Chemistry Issue of Pesticides Use Under Inclusion of Different Scientific Worldviews. *Eurasia Journal of Mathematics, Science and Technology Education*, *17*(7), em1981. https://doi.org/10.29333/ejmste/10979