Journal of Natural Science and Integration
P-ISSN: 2620-4967|E-ISSN: 2620-5092
Vol. 9, No. 1, April 2026, pp 70-83

Available online at:
http://ejournal.uin-suska.ac.id/index.php/JNSI
DOI: 10.24014/jnsi.v9i1.39231

Assessment of Pre-Service Chemistry Teachers' Laboratory Teaching Self-Efficacy and Lesson Plan Quality for Deep Learning Readiness

Setia Rahmawan[1*], Dea Santika Rahayu[2], Muhammad Zamhari[1,3]

[1] *Department of Chemistry Education, UIN Sunan Kalijaga, Indonesia*
[2] *Department of Science Education, Universitas Tidar, Indonesia*
[3] *Department of Chemistry, Faculty of Science, Chulalongkorn University, Thailand*

**Correspondence Author: setia.rahmawan@uin-suka.ac.id*

ABSTRACT
Laboratory-based instruction is fundamental in chemistry education because it supports conceptual understanding through direct engagement with phenomena, evidence generation, and scientific reasoning. At the same time, current curriculum reforms in Indonesia emphasise deep learning readiness, requiring learning designs that promote inquiry, higher-order thinking, and meaningful technology use. This study examines the alignment between pre-service chemistry teachers' laboratory-teaching self-efficacy and the quality of their laboratory-oriented lesson plans in fostering deep learning readiness. Using a quantitative descriptive design complemented by systematic document analysis, the study involved 46 pre-service chemistry teachers enrolled in a school-based practicum course at an Indonesian university. Data were collected using (1) an adapted Chemistry Laboratory Teaching Self-Efficacy Scale covering experimental processes, technology use, and laboratory safety and (2) an analytic rubric to evaluate lesson plans across deep learning orientation, inquiry and reasoning structure, practicum design, higher-order assessment alignment, technology integration, safety/risk documentation, and instructional clarity. Descriptive statistics summarised efficacy and lesson plan quality, while cross-tabulation explored patterns between perceived capability and planning competence. Findings indicate that most participants reported high to very high laboratory-teaching self-efficacy, particularly in experimental procedures and safety. However, most lesson plans were rated moderate, with recurring weaknesses in the design of open inquiry, explicit higher-order assessment tasks, the purposeful integration of digital tools, and detailed safety documentation. The results suggest a partial misalignment between strong self-beliefs and the demonstrated quality of deep-learning-oriented lesson planning. The study highlights the need for teacher education programmes to combine explicit rubric-based lesson planning instruction, iterative feedback cycles, and technology-rich laboratory pedagogy to strengthen deep learning readiness in chemistry education.

Keywords: *Self-efficacy, lesson plan assessment, laboratory-based learning, deep learning, chemistry education.*

INTRODUCTION

Laboratory-based instruction plays a central role in chemistry education because it connects macroscopic observations, submicroscopic explanations, and symbolic representations through experimentation and measurement. Well-designed laboratory experiences allow learners to collect and interpret data, build explanations, and refine conceptual understanding. When laboratory activities move beyond procedural "cookbook" routines, they can cultivate problem-solving, metacognitive regulation, and scientific reasoning (Agustian & Seery, 2017; van Brederode et al., 2020; Yuriev et al., 2017). Recent work also highlights the importance of structuring practical work at meaningful levels of inquiry, because inquiry depth influences how students engage with evidence and reasoning rather than merely following instructions (Raker et al., 2024).

In Indonesia, the discourse of deep learning *(pembelajaran mendalam)* emphasises learning designs that foster conceptual integration, reflective reasoning, and application across contexts. In chemistry classrooms, laboratory learning can strongly support this orientation if students are asked to generate questions or predictions, justify experimental decisions, evaluate evidence, and connect results to core ideas. However, deep-learning-oriented laboratory instruction is not automatic; it depends on teachers' pedagogical capacity to design inquiry, align assessment with higher-order thinking, and scaffold reasoning processes through coherent lesson planning (Großmann & Krüger, 2024; Li et al., 2024)

Among the psychological factors shaping instructional practice, teacher self-efficacy has consistently been linked to persistence, resilience, and openness to adopting learner-centred approaches. Syntheses and large-scale studies indicate that teacher self-efficacy is related to classroom processes, instructional quality, and teacher well-being(Fackler & Malmberg, 2016; Zee & Koomen, 2016). In science education specifically, self-efficacy has been associated with instructional practices and teacher outcomes, such as job satisfaction (Perera et al., 2022), while teacher education studies show that reflective practice and guided experiences can shape pre-service teachers' efficacy beliefs (Menon & Azam, 2021). At a theoretical level, recent reviews also clarify that the development of teaching self-efficacy depends on how mastery experiences, vicarious learning, and social persuasion are conceptualised and measured (Morris et al., 2017).

In laboratory contexts, self-efficacy becomes more complex because teachers must coordinate experimental planning, materials management, safety/risk control, and increasingly, technology integration. Conceptualised laboratory teaching self-efficacy as multidimensional, including confidence in experimental processes, technology-supported laboratory work, and safety (Turan-Oluk et al., 2022). Empirical studies also indicate that inquiry-oriented learning experiences can strengthen students' and teachers' sense of competence in chemistry learning environments (Nzomo et al., 2023). However, strong efficacy beliefs do not necessarily guarantee high-quality instructional design, particularly when teachers must transform curriculum expectations into detailed learning sequences, inquiry prompts, assessments, and safety procedures.

Lesson plans and teaching modules, therefore, function as critical professional artefacts to examine planning competence. A growing line of research argues that lesson plans provide observable evidence of how teachers translate goals into learning tasks, scaffolds, and assessments, and that rubric-based evaluation can distinguish levels of planning quality (Großmann et al., 2024; Ndihokubwayo et al., 2022). Scoping reviews also show that research on written lesson plans increasingly focuses on measurable indicators of professional competence but remains substantially varied in conceptual clarity and methodological approaches (Großmann et al., 2024). In science teaching contexts, design-based competence is particularly relevant for laboratory learning because teachers must align objectives, experimental designs, inquiry prompts, and assessment tasks into a coherent structure(Zhu et al., 2024).

Technology integration further complicates deep learning readiness in laboratory instruction. Digital tools can enrich inquiry through simulation, data logging, visualisation, and virtual experimentation, but benefits depend on pedagogical alignment rather than mere usage (Lee & Choi, 2017). Systematic reviews of virtual chemistry laboratories report growing technological options and instructional designs, including blended and inquiry-supported approaches (Chan et al., 2021; Kolil et al., 2020). Specific implementations, such as virtual worlds for laboratory learning, also suggest that digital environments can broaden access and engagement, though they still require careful scaffolding (Winkelmann et al., 2017). Teacher readiness to integrate technology is commonly examined through TPACK-related measures and self-efficacy; updated TPACK instruments and empirical work underscore that technology-related competence must be built systematically in teacher education (Nugraheni & Srisawasdi, 2025; Valtonen et al., 2017).

Laboratory safety (K3) is another non-negotiable dimension of laboratory-based deep learning readiness. Studies show that safety awareness, attitudes, and practices among learners can vary substantially and should be explicitly developed rather than assumed (Walters et al., 2017). Research also demonstrates that targeted interventions, such as video-based safety instruction, can improve safety learning outcomes (Pekdağ, 2020) and that structured safety training can shape a safer laboratory culture and professional practice (Viitaharju et al., 2021). In teacher education, this implies that lesson plans should explicitly document safety procedures, risk identification, and waste management, rather than relying solely on tacit knowledge.

Despite extensive research on self-efficacy, technology integration, and lesson planning competence, a key gap remains: relatively few studies examine whether pre-service chemistry teachers' laboratory teaching self-efficacy aligns with the quality of their written laboratory lesson plans designed for deep learning readiness. This gap matters because teacher candidates may feel confident conducting experiments and managing safety, while their lesson plans may still underrepresent open inquiry design, higher-order assessment alignment, and meaningful technology use (Großmann & Krüger, 2024; Perera et al., 2022; Raker et al., 2024). Preliminary observations indicate a potential mismatch between pre-service chemistry teachers' laboratory-teaching self-efficacy and the quality of their written lesson plans, particularly in domains that require explicit inquiry design, alignment with higher-order assessment, and safety documentation. Identifying such alignment or misalignment is essential for strengthening practicum-based teacher education.

Therefore, this study investigates laboratory-based chemistry deep learning readiness by integrating two complementary indicators: (1) pre-service teachers' perceived self-efficacy for laboratory teaching (Turan-Oluk et al., 2022)(Turan-Oluk et al., 2022), and (2) rubric-based quality of laboratory-oriented lesson plans/modules (Großmann & Krüger, 2024; Ndihokubwayo et al., 2022). The research addresses the following questions: (1) What are the levels and dimensions of pre-service chemistry teachers' laboratory teaching self-efficacy? (2) What is the quality of their laboratory lesson plans when evaluated using a deep-learning-oriented analytic rubric? (3) How do self-efficacy levels relate to lesson plan quality patterns, and what does this imply for deep learning readiness in chemistry teacher education?

METHODOLOGY

This study used a quantitative descriptive approach complemented by systematic document analysis to examine deep learning readiness in laboratory-based chemistry instruction from two angles: psychological readiness (self-efficacy beliefs) and pedagogical readiness (quality of lesson plans). This dual-focus design follows current directions in teacher education research that treat written lesson plans as assessable artefacts of professional competence (Großmann et al., 2024; Großmann & Krüger, 2024), while also recognising self-efficacy as a robust predictor of teacher motivation and instructional engagement (Perera et al., 2022; Zee & Koomen, 2016). The research stages were explicitly structured to ensure alignment between conceptual foundations, instruments, data collection procedures, and analysis techniques.

The objective of the research was to assess the degree of alignment between pre-service chemistry teachers' laboratory teaching self-efficacy and the quality of their laboratory lesson plans as indicators of deep learning readiness. The research subjects were 46 pre-service chemistry teachers enrolled in a School-Based Practicum (PLP) course at a public Islamic university in Indonesia. Participants were in advanced undergraduate teacher education and had completed core coursework in chemistry content, laboratory management, instructional design, and microteaching.

The research population consisted of all students enrolled in the PLP course during the data-collection semester. Because the population size was manageable, the study employed total sampling, meaning all 46 students served as the sample. The research population consisted of all 46 pre-service chemistry teachers enrolled in the School-Based Practicum (PLP) course during the data-collection semester. Given the manageable, finite population size, the study employed total sampling, in which all population members served as the sample. This approach eliminates sampling error and is considered appropriate when the entire population is accessible (Cohen et al., 2018). The sample size of 46 also exceeds the minimum threshold of 30 recommended for most quantitative research (Fraenkel et al., 2012). It is consistent with descriptive studies that do not require large samples for inferential modelling (Creswell & Creswell, 2018). The research objectives were (1) participants' laboratory teaching self-efficacy and (2) their written lesson plans/teaching modules integrating chemistry practicum activities.

Two categories of data were collected. The primary quantitative data were derived from a self-efficacy questionnaire measuring perceived capability to conduct laboratory-based chemistry teaching, including experimental processes, technology use, and safety (Turan-Oluk et al., 2022). These data were essential because efficacy beliefs can shape teachers' persistence and willingness to enact inquiry-oriented practice(Menon & Azam, 2021; Perera et al., 2022). The primary artefact-based data were 46 lesson plans/teaching modules submitted for the practicum course. Lesson plans were essential because they provide observable evidence of how teacher candidates operationalise inquiry depth, scaffolding, assessment, technology integration, and safety documentation (Großmann et al., 2024; Ndihokubwayo et al., 2022). Supporting data included practicum guidelines and curriculum references used to contextualise rubric criteria and interpret lesson plan structures.

The self-efficacy instrument was adapted from the Chemistry Laboratory Teaching Self-Efficacy Scale (Turan-Oluk et al., 2022). Adaptation procedures included translation, contextual alignment with Indonesian curriculum terminology, and expert validation involving chemistry education lecturers and an educational measurement specialist. Content validity was established through expert judgement evaluating item relevance, representativeness, and clarity. A pilot test with pre-service teachers outside the sample confirmed readability and item discrimination. Reliability was evaluated using Cronbach's alpha, with high internal consistency for the total scale and each subscale ($\alpha > .85$), consistent with expectations for efficacy measures. The instrument contained 28 items across three dimensions: (a) experimental processes, (b) technology use in laboratory instruction, and (c) laboratory safety (K3). Responses used a 4-point Likert scale (1 = strongly disagree to 4 = strongly agree). Scores were computed as total mean and subscale means, then categorised into low, moderate, high, and very high using equal-interval cut-offs on the 1–4 scale.

The lesson plan quality instrument was an analytic rubric developed by adapting rubric-based lesson plan assessment approaches (Großmann & Krüger, 2024) and lesson plan analysis protocols used in science education research(Ndihokubwayo et al., 2022). Rubric development was also informed by evidence on inquiry-oriented laboratory design and scaffolding for reasoning (Agustian & Seery, 2017; Raker et al., 2024; van Brederode et al., 2020; Yuriev et al., 2017), as well as technology integration principles drawn from TPACK-related research (Nugraheni & Srisawasdi, 2025; Valtonen et al., 2017)and virtual laboratory reviews (Chan et al., 2021; Kolil et al., 2020). Safety and risk management indicators were included based on laboratory safety studies emphasising explicit safety learning and documentation (Pekdağ, 2020; Viitaharju et al., 2021; Walters et al., 2017). The rubric contained eight dimensions rated on a 4-point scale (1 = weak, 4 = exemplary): (A) deep learning orientation of learning outcomes, (B) conceptual accuracy and coherence, (C) practicum design quality and alignment, (D) inquiry and reasoning structure, (E) alignment of assessment with higher-order thinking/deep learning, (F) safety and risk management (K3) documentation, (G) technology integration, and (H) clarity and scaffolding. Validity evidence was established via expert review; reliability was strengthened using two independent raters, rater training, and inter-rater agreement checks (Cohen's kappa > .75), indicating substantial agreement.

Data collection occurred in two phases over one practicum semester. First, participants completed the online self-efficacy questionnaire under supervised conditions to reduce collaboration bias and ensure independent responses. Second, lesson plans and modules were collected from official course submissions. Each artefact was anonymised and scored using the analytic rubric by two trained raters. Data analysis used descriptive statistics (means, standard deviations, frequency distributions) to profile self-efficacy and lesson plan quality across dimensions. Cross-tabulation was then used to examine patterns between efficacy categories and lesson plan quality categories. Given the sample size and the study's exploratory aim, the analysis focused on descriptive patterns and alignment profiles rather than inferential modelling. This approach was considered appropriate for identifying practical discrepancies between perceived capability and documented planning competence, which is central to deep learning readiness in laboratory-based chemistry teaching (Großmann & Krüger, 2024; Perera et al., 2022).

RESULT AND DISCUSSION

Laboratory Teaching Self-Efficacy

Across 46 pre-service chemistry teachers, overall laboratory teaching self-efficacy was high (M = 3.23, SD = 0.41; Table 1). Most respondents were classified as high (52.2%) or very high (43.5%), with only two respondents (4.3%) in the moderate category. At the subscale level, perceived capability was strongest for laboratory safety (K3) (M = 3.41, SD = 0.50) and experimental processes (M = 3.24, SD = 0.43). Technology integration efficacy was the lowest subscale (M = 3.06, SD = 0.55), indicating that participants felt less confident about implementing digital tools in laboratory instruction than about managing experiments and safety.

Item-level means show that the lowest-confidence statements were concentrated in advanced technology use: teaching students to experiment in virtual laboratories (M = 2.89), using virtual-lab applications when real laboratories lack sufficient materials/equipment (M = 2.91), and using augmented reality to support submicroscopic visualisation (M = 2.91). In contrast, the highest-rated items reflected routine laboratory competencies such as teaching equipment use and providing a safe laboratory environment (item means ≈ 3.46–3.48). This pattern suggests that participants' self-efficacy is anchored in familiar laboratory practices, while more specialised technology-enhanced laboratory pedagogies remain less established.

Lesson Plan Quality for Deep Learning Readiness

A total of 47 lesson plans/teaching modules integrating chemistry practicum activities were evaluated. Overall lesson plan quality was moderate-to-high (M = 2.77, SD = 0.52; Table 2), with 61.7% categorised as high, 17.0% as moderate, 14.9% as very high, and 6.4% as low. Dimension-level results reveal a clear profile of strengths and weaknesses (Figure 1). The strongest dimensions were learning outcomes and deep learning orientation (A; M = 3.30, SD = 0.69) and conceptual accuracy/coherence (B; M = 3.26, SD = 0.57), indicating that most plans stated appropriate outcomes and represented chemistry concepts correctly.

However, two dimensions emerged as persistent bottlenecks. First, practicum design quality (C) was comparatively low (M = 2.30, SD = 1.18), and 55.3% of plans scored ≤2, reflecting limited inquiry openness, incomplete data-collection logic, or weak alignment between practicum tasks and learning outcomes. Second, safety and risk management documentation (F) was the lowest dimension (M = 1.51, SD = 0.66), and 91.5% of plans scored ≤2, indicating that many plans lacked explicit hazard identification, controls/PPE specifications, waste handling, or emergency procedures. Moderate scores were observed for inquiry and reasoning structure (D; M = 2.91, SD

= 0.58), HOTS/deep learning assessment alignment (E; M = 2.74, SD = 0.71), and technology integration (G; M = 2.96, SD = 0.78), suggesting partial but inconsistent embedding of reasoning prompts, higher-order assessment tasks, and technology-supported learning activities.

Alignment Between Self-Efficacy and Lesson Plan Quality

For 27 participants whose survey responses could be matched to their lesson plan scores, the relationship between overall self-efficacy and overall lesson plan quality was weak (Pearson r = .17). Cross-tabulation shows that high/very high self-efficacy co-occurred with a range of lesson plan quality levels, including low and moderate categories (Table 3). Taken together, these findings indicate partial alignment in some areas, but also highlight areas where confidence does not readily translate into explicit planning.

Table 1. Descriptive Statistics for Laboratory Teaching Self-Efficacy (N = 46)

Scale	Mean	SD
Overall self-efficacy	3.23	0.41
Experimental processes	3.24	0.43
Technology integration	3.06	0.55
Laboratory safety (K3)	3.41	0.50

Note. Scale range = 1–4. Higher scores indicate higher self-efficacy.

Table 2. Category Distribution 0f Overall Self-Efficacy (N = 46)

Category	n	%
Low	0	0.0
Moderate	2	4.3
High	24	52.2
Very High	20	43.5

Table 3. Lesson Plan Quality by Rubric Dimension (N = 47 lesson plans/modules)

Dimension	Mean	SD	% plans ≤2
A. Learning outcomes & deep learning orientation	3.30	0.69	8.51
B. Conceptual accuracy & coherence	3.26	0.57	6.38
C. Practicum design quality	2.30	1.18	55.32
D. Inquiry and reasoning structure	2.91	0.58	21.28
E. HOTS/deep learning assessment alignment	2.74	0.71	31.91
F. Safety & risk management (K3)	1.51	0.66	91.49
G. Technology integration	2.96	0.78	23.4
H. Clarity & scaffolding	3.15	0.69	12.77

Note. Scores are on a 1–4 scale (1 = weak; 4 = exemplary). % plans ≤2 indicates the proportion of artefacts rated weak-to-basic for a given dimension.

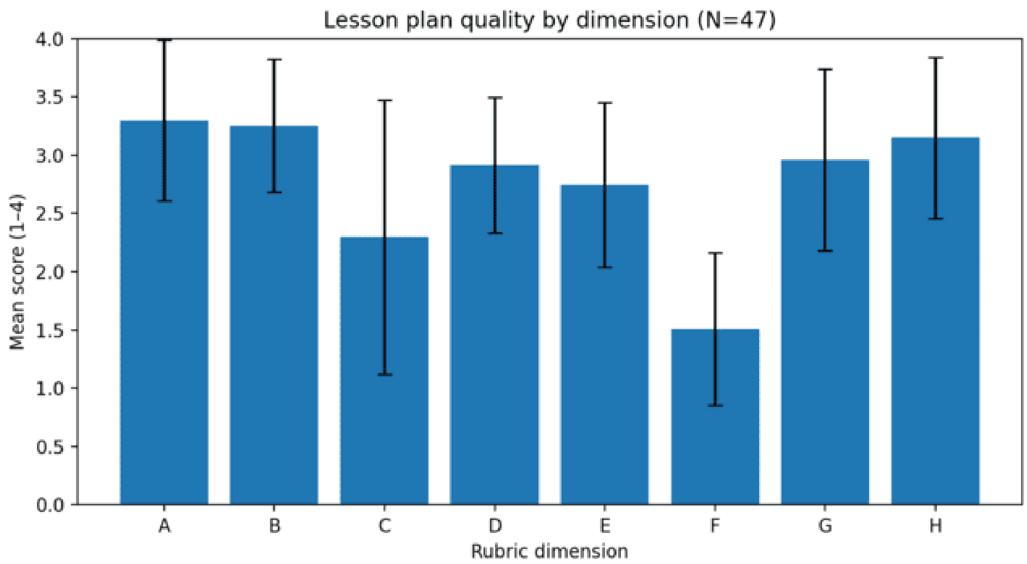

Figure 1. Mean Scores of Lesson Plan Quality Dimensions (A–H) with Standard Deviation Error Bars (N = 47).

Table 4. Cross-tabulation of Self-Efficacy and Lesson Plan Quality Categories (Matched N = 24; Pearson R = 0.15)

Self-efficacy category	Low	Moderate	High	Very High
Moderate	0	0	1	0
High	0	3	10	0
Very High	2	0	5	3

Note. Categories are based on equal-interval cut-offs on the 1–4 scale.

Discussion

This study aimed to (1) describe pre-service chemistry teachers' laboratory teaching self-efficacy, (2) evaluate the quality of their laboratory-oriented lesson plans/teaching modules for deep learning readiness, and (3) examine the alignment between perceived capability and documented planning competence. The importance of this research lies in its dual-lens approach: combining a belief-based indicator (self-efficacy) with an artefact-based indicator (lesson plan quality). This integration is particularly relevant to deep learning readiness because deep learning requires not only confidence in teaching but also explicit instructional design decisions that promote inquiry, higher-order assessment, purposeful technology use, and safe laboratory practices. This study contributes by integrating self-efficacy measures with rubric-based lesson plan analysis to reveal alignment patterns in deep learning readiness. This approach is relatively underrepresented in the existing literature on chemistry teacher education. Unlike prior work that has examined these dimensions in isolation, the current study demonstrates that perceived capability and documented instructional competence can diverge systematically, particularly in safety planning and inquiry design.

Overall, the results show a clear contrast between strong perceived readiness and uneven documented readiness. Most participants reported high to very high laboratory teaching self-efficacy, especially in experimental processes and safety. However, lesson plan quality, although generally moderate-to-high, revealed persistent bottlenecks in practicum design, particularly in safety and risk management documentation. In addition, the relationship between overall self-efficacy and overall lesson plan quality was weak, indicating that high confidence does not reliably predict high-quality written planning for deep learning-oriented laboratory instruction. This pattern matters for teacher education because lesson plans are often the primary basis for supervisor feedback and school accountability, and they function as a blueprint for consistent implementation.

Self-efficacy is High, But May Be Partially "Uncalibrated"

High self-efficacy among pre-service teachers is not unexpected during practicum settings, where repeated exposure to routine teaching tasks and supportive supervision can strengthen perceptions of capability. However, the weak alignment with lesson plan quality suggests possible calibration issues. Participants may feel confident about what they can do in practice, while lacking the skills to externalise that competence into detailed, assessable lesson plans. This interpretation aligns with evidence that pre-service teachers' self-perceptions can exceed supervisors' evaluations when criteria are inconsistently operationalised, increasing the risk of overconfidence rather than accurate self-assessment (Dassa & Nichols, 2019). Such calibration matters because deep learning demands careful orchestration of inquiry prompts, evidence use, and assessment alignment elements that must be planned explicitly rather than assumed.

From a developmental perspective, self-efficacy is known to change during professional experience placements and educational reform contexts. However, growth trajectories vary by the quality of mentoring, task demands, and opportunities to enact complex practices (Gordon et al., 2025). Importantly, self-efficacy is also linked to persistence in teacher education; if pre-service teachers perceive repeated failure without targeted support, efficacy may decline and even predict intentions to quit (Pfitzner-Eden, 2016). Therefore, teacher education should treat efficacy not as an endpoint but as a variable that must be supported and calibrated through explicit performance criteria and feedback.

Practicum Design and Deep Learning: The "Design Gap" Behind Moderate Lesson Plan Scores

Lesson plan analysis indicates that practicum design was a recurring weakness: many plans remained procedural or weakly aligned with deep learning goals. This finding suggests that pre-service teachers may understand deep learning rhetorically but struggle to translate it into inquiry structures that require students to make predictions, justify experimental choices, interpret anomalies, and connect evidence to chemical models. One plausible explanation is that designing inquiry-based laboratories requires specialised design knowledge (e.g., selecting variables, anticipating evidence patterns, and crafting reasoning prompts) that is not automatically acquired through conducting experiments. Even if teachers feel confident managing practical work, deep learning readiness requires purposeful design decisions that increase the epistemic demand of laboratory tasks and align those demands with assessment.

Rubric-based lesson plan coaching and iterative revision are promising solutions. Recent evidence on rubrics shows that they can improve learning outcomes, self-regulated learning, and self-efficacy when used with clear criteria and feedback cycles(Panadero et al., 2017). Studies also indicate that rubrics can enhance accuracy and reduce cognitive load in self-assessment, supporting better calibration of performance judgments(Krebs et al., 2022). In teacher education settings, structured self-assessment and instructor feedback based on rubrics can improve pre-service teachers' performance and perceptions of the assessment process (Karaman, 2024). These findings support the recommendation that lesson planning should be taught as a revision-based design practice: draft → critique → revise → re-evaluate, rather than a one-shot documentation task.

Technology Integration: Presence is Not The Same as Pedagogical Function

Although technology integration in lesson plans was not the lowest dimension, the study indicates inconsistent support for inquiry and deep learning. This pattern is consistent with research showing that technology integration in teacher education is shaped by multiple interacting factors: attitudes, beliefs, access, competency, and experience, and these factors influence whether technology becomes central to learning tasks or remains peripheral (Farjon et al., 2019). Multilevel evidence also shows that what matters in ICT competency development is not only individual knowledge but also the quality of training environments and opportunities to practice technology in authentic instructional contexts (Tondeur et al., 2018).

A key implication is that teacher education must focus on technology's epistemic function in laboratory learning: whether digital tools enable modelling, evidence generation, visualisation, or argumentation. Studies examining TPACK and technology use in lesson plans consistently find that self-reported knowledge does not guarantee meaningful integration; what matters is the translation of knowledge structures into design decisions within lesson artefacts (Schmid et al., 2021). Recent work also emphasises contextual effects. Technology effects: technology integration differs between pre-service and in-service teachers and is shaped by constraints, task demands, and instructional goals (Tschönhens et al., 2024). Therefore, "technology integration" should be assessed not by the presence of tools but by whether technology substantively advances inquiry and reasoning in laboratory learning.

From an assessment perspective, TPACK-related knowledge can be measured more precisely using validated instruments, enabling teacher education programs to diagnose which domains (e.g., technological pedagogical knowledge) require targeted support (Baier & Kunter, 2020). Intervention studies further show that structured programs can foster TPACK development and improve technology integration in lesson planning when teacher candidates are guided to design, test, and revise technology-enhanced tasks (Lachner et al., 2021). Taken together, these sources support redesigning practicum activities so that technology is used to enable deep learning: data collection, modelling, and explanation rather than presentation.

The Most Substantial Discrepancy: High Safety Self-Efficacy but Very Weak Safety Documentation

The most unexpected and important finding is the mismatch between high safety self-efficacy and very weak safety/risk management documentation in lesson plans. This suggests that safety knowledge may be treated as tacit routine rather than as explicit instructional content or a planning requirement. The chemical safety education literature emphasises that hazard recognition and risk assessment are foundational competencies that must be made explicit to improve safety culture and consistency (Hill, 2019). Moreover, safety education is increasingly framed not as a compliance add-on but as a teachable and assessable learning outcome within chemistry instruction (Goode et al., 2021).

Several plausible reasons may explain why safety documentation is absent even when confidence is high. First, teacher candidates may believe that safety is handled verbally or implicitly during implementation. Second, they may lack templates and prompts that operationalise risk management into lesson plans (hazards, controls, PPE, waste disposal, emergency procedures). Third, limited laboratory infrastructure in school settings may lead candidates to focus on feasibility and time constraints rather than systematic risk documentation. However, the literature is clear that safety expertise is not optional in chemistry teaching; insufficient explicit safety planning can produce inconsistent practices and increased risk (Sigmann, 2018).

Recent safety frameworks also offer concrete instructional strategies that can be integrated into lesson planning. For example, RAMP-based safety instruction has been used to improve student learning outcomes related to safety and risk management in organic chemistry laboratory contexts (Nyansa et al., 2024). Similarly, using risk-management-based safety case studies can strengthen safety reasoning and decision-making skills that can be embedded into laboratory lessons as part of deep learning (Bocwinski et al., 2021). Therefore, this study's findings strongly support making safety documentation a "must-pass" component in lesson-plan rubrics and requiring explicit risk-management routines in every practicum module.

Implications for Teacher Education Practice and Assessment

The findings suggest three practical implications. First, teacher education should institutionalise rubric-guided iterative planning. Rubrics should not only score products but function as tools for learning and calibration through repeated cycles of drafting and feedback (Karaman, 2024; Panadero et al., 2017). Second, practicum experiences should include structured technology-enhanced inquiry design tasks supported by TPACK development, ensuring that technology becomes a driver of epistemic practices rather than a presentation tool (Lachner et al., 2021; Schmid et al., 2021; Tondeur et al., 2018). Third, safety planning should be reframed as explicit learning design, supported by RAMP-based prompts and risk management case studies so candidates learn to document hazards, controls, and waste management systematically (Bocwinski et al., 2021; Goode et al., 2021; Nyansa et al., 2024).

Additionally, assessment systems used in practicum supervision can be improved by adopting consistent rubrics that target specific competencies. Emerging work suggests that assessment rubrics can shape practicum behaviours and promote implementation of complex teaching practices (Shamir-Inbal et al., 2025). This supports integrating lesson plan rubrics with supervision rubrics so that planning quality criteria align with observation-based evaluation.

Limitations and Future Research

This study has limitations. It used a descriptive design within a single teacher education context, limiting generalizability. Self-efficacy was measured through self-report, which may be influenced by social desirability or incomplete calibration (Dassa & Nichols, 2019). Lesson plan quality reflects intended instruction, not enacted classroom practice; some candidates may implement stronger inquiry or safety routines than they document. Future research should triangulate lesson-plan artefacts with classroom observations, interviews, and teaching rehearsals to explain why certain elements, especially safety documentation and inquiry-based assessment, are omitted. Quasi-experimental research could test whether rubric-guided revision cycles, TPACK-informed task-design interventions, or RAMP-based safety-planning prompts significantly improve both lesson-plan quality and enacted deep-learning practices (Lachner et al., 2021; Nyansa et al., 2024). Longitudinal studies could also examine how alignment between efficacy and planning competence evolves from practicum into induction years, particularly during curriculum reforms (Gordon et al., 2025), and how changes in efficacy relate to retention intentions (Pfitzner-Eden, 2016). Regarding contextual boundaries, this study was conducted within a single institutional setting in Indonesia, and the specific practicum structure, curriculum expectations, and student demographics may not be representative of other teacher education programmes, whether nationally or internationally. These institutional and cultural constraints, therefore, bound the applicability of the findings, and caution is warranted when interpreting them in different educational systems. Future research should replicate the study across diverse contexts, including different universities, regions, and national curriculum frameworks, to determine whether the observed misalignment between self-efficacy and lesson plan quality is a localised phenomenon or a more generalised pattern in pre-service chemistry teacher education.

CONCLUSION

This study assessed pre-service chemistry teachers' readiness to implement deep learning through laboratory-based instruction by examining laboratory teaching self-efficacy and the quality of practicum-oriented lesson plans/modules. The findings show that participants generally reported high to very high self-efficacy, particularly in experimental processes and laboratory safety.

However, the quality of written lesson plans was more uneven, with recurrent weaknesses in inquiry-oriented practicum design, explicit alignment of higher-order assessment with learning outcomes, purposeful technology integration to support inquiry, and especially the documentation of safety and risk management (K3). The weak association between self-efficacy and lesson plan quality suggests that strong confidence does not automatically translate into high-quality, deep-learning-oriented planning, highlighting a misalignment between perceived capability and documented professional competence. Therefore, chemistry teacher education programs should strengthen readiness through structured, rubric-guided lesson design practice with iterative feedback, explicit training in inquiry and HOTS assessment design, and mandatory safety-risk documentation routines (e.g., hazard identification, controls/PPE, waste handling, and emergency procedures) embedded in every laboratory lesson plan. Future research should triangulate document analysis with classroom enactment data (observations, teaching rehearsals, and interviews) and evaluate the effectiveness of targeted interventions, such as rubric-based coaching, technology-integrated task design, and RAMP-informed safety planning, on both lesson plan quality and enacted deep learning in school laboratory settings.

REFERENCES

Agustian, H. Y., & Seery, M. K. (2017). Reasserting the role of pre-laboratory activities in chemistry education: A proposed framework for their design. *Chemistry Education Research and Practice*, *18*, 518–532. https://doi.org/10.1039/C7RP00140A

Baier, F., & Kunter, M. (2020). Construction and validation of a test to assess (pre-service) teachers' technological pedagogical knowledge (TPK). *Studies in Educational Evaluation*, *67*, 100936. https://doi.org/10.1016/j.stueduc.2020.100936

Bocwinski, R., Finster, D. C., & Weizman, H. (2021). Framework for Teaching Safety Case Studies Using a Risk Management Approach. *Journal of Chemical Education*, *98*(12), 3824–3830. https://doi.org/10.1021/acs.jchemed.1c00625

Chan, P., Van Gerven, T., Dubois, J.-L., & Bernaerts, K. (2021). Virtual chemical laboratories: A systematic literature review of research, technologies and instructional design. *Computers and Education Open*, *2*, 100053. https://doi.org/10.1016/j.caeo.2021.100053

Dassa, C., & Nichols, J. (2019). Self-efficacy or overconfidence? Comparing the confidence and competence of preservice teachers. *The New Educator*, *15*(2), 156–174. https://doi.org/10.1080/1547688X.2019.1578447

Fackler, S., & Malmberg, L.-E. (2016). Teachers' self-efficacy in 14 OECD countries: Teacher, student group, school and leadership effects. *Teaching and Teacher Education*, *56*, 185–195. https://doi.org/10.1016/j.tate.2016.03.002

Farjon, D., Smits, A., & Voogt, J. (2019). Technology integration of pre-service teachers explained by attitudes and beliefs, competency, access, and experience. *Computers & Education*, *130*, 81–93. https://doi.org/10.1016/j.compedu.2018.11.010

Goode, S. R., Wissinger, J. E., & Wood-Black, F. (2021). Introducing the Journal of Chemical Education's Special Issue on Chemical Safety Education: Methods, Culture, and Green Chemistry. *Journal of Chemical Education*, *98*(1), 1–6. https://doi.org/10.1021/acs.jchemed.0c01459

Gordon, D., Bourke, T., Mills, R., & Blundell, C. N. (2025). From pre-service to beginning teacher: Understanding how teacher self-efficacy develops during educational reform. *Teaching and Teacher Education*, *168*, 105257. https://doi.org/10.1016/j.tate.2025.105257

Großmann, L., & Krüger, D. (2024). Assessing the quality of science teachers' lesson plans: Evaluation and application of a novel instrument. *Science Education*, *108*(1), 153–189. https://doi.org/10.1002/sce.21832

Großmann, L., Krüger, D., Campbell, T., Krepf, M., König, J., & Rothland, M. (2024). The use of written lesson plans in research on teacher education and teaching: A scoping review. *Studies in Science Education*, *61*(2), 329–379. https://doi.org/10.1080/03057267.2024.2415246

Hill, R. H. (2019). Recognizing and understanding hazards — The key first step to safety. *Journal of Chemical Health and Safety*, *26*(3), 5–10. https://doi.org/10.1016/j.jchas.2018.11.005

Karaman, P. (2024). Effects of using rubrics in self-assessment with instructor feedback on pre-service teachers' academic performance, self-regulated learning and perceptions of self-assessment. *European Journal of Psychology of Education*, *39*, 2551–2574. https://doi.org/10.1007/s10212-024-00867-w

Kolil, V. K., Muthupalani, S., & Achuthan, K. (2020). Virtual experimental platforms in chemistry laboratory education and its impact on experimental self-efficacy. *International Journal of Educational Technology in Higher Education*, *17*, 30. https://doi.org/10.1186/s41239-020-00204-3

Krebs, R., Rothstein, B., & Roelle, J. (2022). Rubrics enhance accuracy and reduce cognitive load in self-assessment. *Metacognition and Learning*, *17*, 627–650. https://doi.org/10.1007/s11409-022-09302-1

Lachner, A., Fabian, A., Franke, U., Preiß, J., Jacob, L., Führer, C., Küchler, U., Paravicini, W., Randler, C., & Thomas, P. (2021). Fostering pre-service teachers' technological pedagogical content knowledge (TPACK): A quasi-experimental field study. *Computers & Education*, *174*, 104304. https://doi.org/10.1016/j.compedu.2021.104304

Lee, J., & Choi, H. (2017). What affects learner's higher-order thinking in technology-enhanced learning environments? The effects of learner factors, interaction, and learning engagement. *Computers & Education*, *115*, 143–152. https://doi.org/10.1016/j.compedu.2017.06.015

Li, Y., Hu, X., & Oon, P.-T. (2024). Unveiling pre-service teachers' competency and challenges in designing 5E inquiry-based iSTEM lessons. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-024-10529-4

Menon, D., & Azam, S. (2021). Investigating preservice teachers' science teaching self-efficacy: An analysis of reflective practices. *International Journal of Science and Mathematics Education*, *19*, 1587–1607. https://doi.org/10.1007/s10763-020-10131-4

Ndihokubwayo, K., Uwamahoro, J., Ndayambaje, I., & Ralph, M. (2022). Lesson plan analysis protocol (LPAP): A tool for evaluating inquiry-based science lesson plans. *Heliyon*, *8*(2), e08780. https://doi.org/10.1016/j.heliyon.2022.e08780

Nugraheni, A. R. E., & Srisawasdi, N. (2025). Development of pre-service chemistry teachers' knowledge of technological integration in inquiry-based learning to promote chemistry core competencies. *Journal of Science Education and Technology*.

Nyansa, M. M. S., Burrows, N. L., Galerneau, A. J., Bekkala, A. P., & Hungwe, K. N. (2024). Investigating the Impact of RAMP-Based Safety Instruction on Student Learning in an Organic Chemistry Lab Course. *Journal of Chemical Education*, *101*(6), 2203–2214. https://doi.org/10.1021/acs.jchemed.3c00955

Nzomo, C. M., Rugano, P., Njoroge, J. M., & Gitonga, C. M. (2023). Inquiry-based learning and students' self-efficacy in chemistry among secondary schools in Kenya. *Heliyon*, *9*(1), e12672. https://doi.org/10.1016/j.heliyon.2022.e12672

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. In *Educational Research Review* (Vol. 22, pp. 74–98). Elsevier Ltd. https://doi.org/10.1016/j.edurev.2017.08.004

Pekdağ, B. (2020). Video-based instruction on safety rules in the chemistry laboratory: Its effect on student achievement. *Chemistry Education Research and Practice*, *21*, 953–968. https://doi.org/10.1039/D0RP00088D

Perera, H. N., Maghsoudlou, A., Miller, C. J., McIlveen, P., Barber, D., Part, R., & Reyes, A. L. (2022). Relations of science teaching self-efficacy with instructional practices, student achievement and support, and teacher job satisfaction. *Contemporary Educational Psychology*, *69*, 102041. https://doi.org/10.1016/j.cedpsych.2021.102041

Pfitzner-Eden, F. (2016). I feel less confident so I quit? Do true changes in teacher self-efficacy predict changes in preservice teachers' intention to quit their teaching degree? *Teaching and Teacher Education*, *55*, 240–254. https://doi.org/10.1016/j.tate.2016.01.018

Raker, J. R., Kucharski, M. M., Doidge, E. D., Yue, D., Tan, L., & Seery, M. K. (2024). Evaluating the level of inquiry in postsecondary chemistry laboratory experiments. *Chemistry Education Research and Practice*, *25*, 79–91. https://doi.org/10.1039/D3RP00154G

Schmid, M., Brianza, E., & Petko, D. (2021). Self-reported technological pedagogical content knowledge (TPACK) and technological pedagogical beliefs and their relationship to technology use in lesson plans. *Computers in Human Behavior*, *115*. https://doi.org/10.1016/j.chb.2020.106586

Shamir-Inbal, T., Amir, A., Avidov-Ungar, O., Hadad, S., & Blau, I. (2025). Can an Assessment Rubric Encourage Hybrid Teaching in Practicum Settings? Comparing Perspectives of Supervisors, Trainers, and Pre-service Teachers. *Journal of Formative Design in Learning*. https://doi.org/10.1007/s41686-025-00109-2

Sigmann, S. B. (2018). Playing with fire: Chemical safety expertise required. *Journal of Chemical Education*, *95*(10), 1736–1746. https://doi.org/10.1021/acs.jchemed.8b00152

Tondeur, J., Aesaert, K., Prestridge, S., & Consuegra, E. (2018). A multilevel analysis of what matters in the training of pre-service teacher's ICT competencies. *Computers & Education*, *122*, 32–42. https://doi.org/10.1016/j.compedu.2018.03.002

Tschönhens, F., Backfisch, I., Fütterer, T., & Lachner, A. (2024). TPACK in action: Contextual effects of pre-service and in-service teachers' knowledge structures for technology integration. *Computers and Education Open*, *7*, 100219. https://doi.org/10.1016/j.caeo.2024.100219

Turan-Oluk, N., Baran, A., & Ekmekci, G. (2022). Chemistry teachers' self-efficacy perception scale for teaching in chemistry laboratories. *Journal of Chemical Education*, *99*(9), 3114–3123. https://doi.org/10.1021/acs.jchemed.2c00240

Valtonen, T., Sointu, E., Kukkonen, J., Kontkanen, S., Lambert, M. C., & Mäkitalo-Siegl, K. (2017). TPACK updated to measure pre-service teachers' twenty-first century skills. *Australasian Journal of Educational Technology*, *33*(3), 15–31. https://doi.org/10.14742/ajet.3518

van Brederode, M. E., Zoon, S. A., & Meeter, M. (2020). Examining the effect of lab instructions on students' critical thinking during a chemical inquiry practical. *Chemistry Education Research and Practice*, *21*, 1173–1182. https://doi.org/10.1039/D0RP00020E

Viitaharju, A., Kiviniemi, K., & Kotiranta, S. (2021). Learning experiences from digital laboratory safety training. *Education for Chemical Engineers*, *34*, 87–93. https://doi.org/10.1016/j.ece.2020.11.009

Walters, A. U. C., Lawrence, W., & Jalsa, N. K. (2017). Chemical laboratory safety awareness, attitudes and practices of tertiary students. *Safety Science*, *96*, 161–171. https://doi.org/10.1016/j.ssci.2017.03.017

Winkelmann, K., Keeney-Kennicutt, W., Fowler, D., & Macik, M. (2017). Development, implementation, and assessment of general chemistry laboratory experiments performed in the virtual world of Second Life. *Journal of Chemical Education*, *94*(7), 849–858. https://doi.org/10.1021/acs.jchemed.6b00733

Yuriev, E., Naidu, S., Schembri, L. S., & Short, J. L. (2017). Scaffolding the development of problem-solving skills in chemistry: Guiding novice students out of dead ends and false starts. *Chemistry Education Research and Practice*, *18*(3), 486–504. https://doi.org/10.1039/C7RP00009J

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

Zhu, L., Cai, C., Tian, L., & Guo, S. (2024). Investigating pre-service science teachers' design performance in laboratory class: The inquiry-based design thinking approach. *Journal of Science Education and Technology*.