International Journal of Ethnoscience and Technology in Education (IJETE)
https://e-journal.undikma.ac.id/index.php/ijete/index
September 2025 Vol. 2, No. 2
pp. 259-280

Remote Inquiry and Virtual Simulation in a Fourier Transform Course: Effects on Critical Thinking Among Prospective STEM Teachers

Ni Nyoman Sri Putu Verawati[1]*, & Hafsemi Rapsanjani[2]

[1] Physics Education Department, University of Mataram, Mataram, INDONESIA
[2] Department of Primary Teacher Education, Universitas Musamus Merauke, Merauke, INDONESIA

*Corresponding author e-mail: veyra@unram.ac.id

Article Info

Article History
Received: July 2025
Revised: August 2025
Published: September 2025

Keywords
Remote Inquiry;
Virtual Simulation;
Critical Thinking Skills;
Fourier Transform Course;
Prospective STEM Teachers

10.33394/ijete.v2i2.18252

Abstract

Critical thinking is a stated goal in STEM teacher education but often underdeveloped in courses that emphasize procedures over reasoning. This study examined whether virtual simulation-assisted remote inquiry improves critical thinking among prospective STEM teachers in a Fourier Transform course. We conducted a randomized pretest–posttest control-group design with two intact classes at one university (experimental n = 20, control n = 20). Both groups received the same content, instructor, timing, and assessments. The intervention embedded prediction, observation, explanation, and decision steps inside an LMS using a PhET Fourier simulation. Critical thinking was measured with an eight-item essay test aligned to analysis, inference, evaluation, and decision making, scored 0–4 per item. All students completed pretest and posttest. The experimental mean rose from 10.90 (SD 2.30) to 26.60 (SD 2.10) with high normalized gain (g = 0.74), while the control mean increased from 11.20 (SD 2.10) to 15.10 (SD 2.40) with low gain (g = 0.19). Gain scores met normality, and an independent-samples t-test showed a significant between-group difference, $t_{(38)} = 10.94$, $p < .001$. Category shifts mirrored these results, with the experimental group moving to critical and very critical at posttest. Findings indicate that simulation-supported remote inquiry can meaningfully elevate critical thinking in abstract topics and offers a feasible model for teacher preparation.

How to Cite: Verawati, N. N. S. P., & Rapsanjani, H. (2025). Remote Inquiry and Virtual Simulation in a Fourier Transform Course: Effects on Critical Thinking Among Prospective STEM Teachers. *International Journal of Ethnoscience and Technology in Education*, *2*(2), 259–280. https://doi.org/10.33394/ijete.v2i2.18252

INTRODUCTION

Thinking, reflecting, and taking in new knowledge are distinctly human, and the common core tying them together is reasoning—our capacity to judge claims, connect evidence, and decide how to act (Byrnes, 2012). Reform agendas often promise to center learning on critical thinking, yet progress remains uneven when courses stay content heavy and procedure first (Hasemi, 2011). Teachers sit at this pressure point because they are asked to cultivate critical thinking while operating inside syllabi and assessments that sometimes reward recall more than reasoning (Fuad et al., 2017). STEM classrooms feel this tension acutely: reports of weak opportunities to analyze evidence, draw inferences, evaluate competing claims, and make justified decisions suggest many students advance without the cognitive habits that higher education says it values (Pursitasari et al., 2020). If critical thinking is to function as more than a slogan, teacher preparation must make it a concrete, assessable goal rather than an assumed by-product of advanced coursework.

It is simplistic to attribute poor educational quality only to teachers, because resources and structures constrain practice. Still, teachers design the proximal experiences through which students think, so their choices shape what learners practice and value (Rubini et al., 2016; Wahidin & Romli, 2020). A practical response is to work upstream with prospective STEM teachers and to treat critical thinking as a learnable target during college so they can later engineer classroom tasks that foster it in schools (Lam et al., 2003). This stance implies a responsibility for higher education institutions to provide systematic opportunities and tools to practice critical thinking and to evaluate it with appropriate measures rather than waiting for it to emerge spontaneously (Innabi & Sheikh, 2007). The claim that such preparation will translate into better schooling deserves to be tested with careful designs situated in demanding STEM topics, not assumed.

A standard account defines critical thinking as solving problems through reflective analysis and evaluation of information or knowledge (Ennis, 2011). A complementary view frames it as rational and reflective decision making grounded in systematic analysis, sound inference, and appropriate evaluation, whether deductive or inductive (Ennis, 2011; Hassard, 2005; E. R. Lai, 2011). Empirically, critical thinking correlates with academic success, reinforcing its status as a core outcome of undergraduate education rather than a luxury skill to be developed only in capstone experiences (Abrami et al., 2008; Halpern, 2014). For the present study, four facets are central because they capture the cognitive operations common in complex STEM tasks: analytical thinking, inference, evaluation, and decision making (Prayogi et al., 2019; Verawati et al., 2021; Wahyudi et al., 2019). Institutions that claim to promote these facets should be able to show growth with credible measures and plausible instructional mechanisms (Tiruneh et al., 2017; Guo & Wang, 2021). The question is not whether these skills matter, but which designs actually move them.

Inquiry-based learning is frequently proposed as a vehicle for strengthening the four facets. It organizes activity around questions, evidence, and argumentation so that students engage in the very processes the facets index (Arends, 2012). The logic is straightforward: if learners repeatedly analyze situations, draw inferences, evaluate explanations, and decide next steps in response to feedback, improvement should follow. Several studies report gains in critical thinking under inquiry approaches, though effect sizes vary with task design, scaffolding quality, and fidelity of implementation (Llewellyn, 2001; Thaiposri & Wannapiroon, 2015). Bailin's analysis emphasizes practices that align directly with the four facets, including identifying assumptions, weighing current scientific knowledge, judging evidence quality, and checking argument coherence (Bailin, 2002). That said, not all inquiry is equal. Tasks that permit procedural completion without explicit justification can fail to exercise analysis, inference, evaluation, or decision making in meaningful ways, which is why careful task design and clear roles matter.

The difficulty compounds in topics that are abstract and representation heavy. Fourier transform is a prime example. Students must coordinate time- and frequency-domain reasoning, track linearity and convolution, and manage assumptions about sampling, windowing, and spectral leakage. Unsurprisingly, learners often describe the material as difficult, and instructors observe persistent gaps in the links that connect formal manipulations to conceptual explanations (Kohaupt, 2015). At the same time, Fourier transform operates as a unifying theme across STEM because it connects signals, systems, imaging, and data analysis; if students can reason well here, benefits plausibly transfer to adjacent areas (Shoenthal, 2014). The instructional challenge is to create experiences where students predict spectral behavior, interpret discrepancies, evaluate claims, and decide on parameter settings with explicit justification rather than rely on rote algorithm following. This challenge sets a high bar for any design claiming to cultivate critical thinking.

Virtual simulation is a plausible scaffold for that challenge. Simulations can visualize hidden quantities, compress time, and expose learners to parameter regimes that would be impractical in physical labs. In Fourier topics, students can manipulate sampling frequency, window type, and signal composition, then inspect amplitude and phase spectra as immediate feedback. Advocates argue that such manipulability supports hypothesis testing and explanation building, which can catalyze analysis, inference, evaluation, and decision making when prompts demand justification rather than simple parameter tuning (Wietecha et al., 2021). Skeptics raise two cautions. First, simulations can become black boxes that encourage trial-and-error "spectral shopping" without conceptual grounding. Second, interface complexity can add extraneous cognitive load that distracts from reasoning. These concerns do not invalidate simulation; they point to the need to embed simulation in inquiry tasks that require prediction, comparison between expected and observed outcomes, and principled reconciliation supported by feedback.

Delivery context shapes what learners actually do. Covid-19 disrupted face-to-face models and forced rapid redesigns of courses that relied on labs and spontaneous instructor coaching (Silva et al., 2022). Programs had to realign strategies to distance learning while still cultivating higher-order skills, not merely shifting to content transmission (Sangster et al., 2020). One response was to adapt inquiry for remote delivery by leveraging information and communication technologies to sustain questioning, evidence gathering, and argumentation at a distance (Novitra et al., 2021). Done well, this shift is not just a stopgap; it can help students practice the same analytic, inferential, evaluative, and decision-oriented moves in online spaces that modern STEM practice increasingly requires. But the quality of remote inquiry depends on structure, clarity of expectations, and integration of tools with prompts that make reasoning visible.

Traditional inquiry frameworks offer a usable script for that structure. Hanson's sequence of orientation, exploration, concept formation, application, and closure provides a path from curiosity to consolidation (Hanson, 2005). In our setting these phases are orchestrated within a learning management system that hosts resources, prediction prompts, simulation tasks, and feedback cycles. As illustrated in Figure 1, the LMS-mediated flow places e-orientation to activate prior knowledge, e-exploration to analyze contextual problems with ICT resources, e-conceptual formation to articulate emerging ideas, e-application to extend concepts to new situations, and e-closure to reflect and receive feedback, which together anchor remote inquiry in a transparent sequence (Bilad et al., 2022). Mentioning the architecture does not claim that structure alone produces gains; it clarifies how opportunities for analysis, inference, evaluation, and decision making are distributed and documented across the learning cycle.

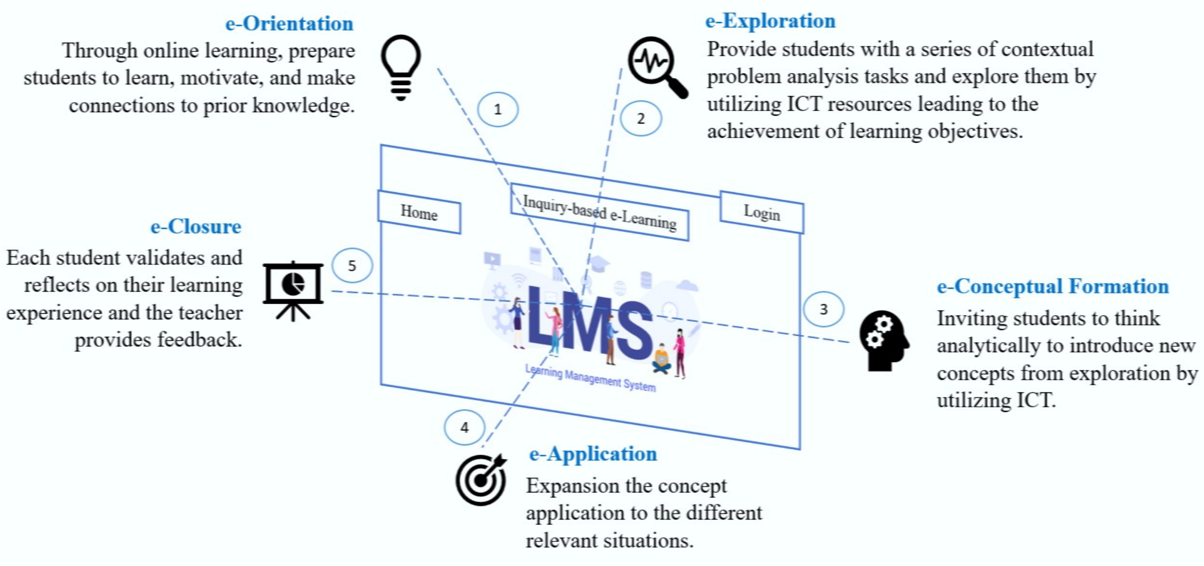

Figure 1. Remote inquiry learning within the LMS framework (Bilad et al., 2022)

Feasibility depends on access and mobility. Remote inquiry reaches more students when activities run on common devices, function under modest bandwidth, and integrate with an institution's platform so learners can participate synchronously or asynchronously. Framed this way, remote inquiry belongs within digital and mobile learning as a normal feature of higher education rather than a temporary workaround (Prahani et al., 2022). Evidence suggests that ICT-supported remote inquiry can enhance higher-order thinking, but the size and durability of effects depend on implementation fidelity and on prompts that tie action to explanation and choice (Novitra et al., 2021). A variety of simulations can serve as resources. PhET, for example, has been used as an accessible library that fits inquiry cycles, with the caveat that tasks must demand reasoning about why outcomes occur rather than simple reproduction (Chinaka, 2021). Tools can amplify or dilute the four facets depending on how they are woven into argumentation and reflection; the pedagogy matters more than the interface.

Focusing on prospective STEM teachers raises the stakes because these learners are mastering content while apprenticing into professional practice. A design that requires them to analyze spectral structures, infer parameter effects, evaluate competing explanations, and decide next steps with explicit warrants doubles as rehearsal for the orchestration they will later perform with adolescents. Reflection prompts that ask them to critique task variants or anticipate student misconceptions make the teacher-education link explicit. Critics might argue that such meta-teaching emphasis reduces time on content. The counterpoint is that separating pedagogy from substance risks brittle knowledge that fails under classroom complexity. The more relevant test is whether the design elicits content-specific reasoning and pedagogical justification and whether assessments capture growth on both fronts. That is an empirical matter, not a claim to be granted in advance.

Assessment choices are consequential. A single total score can hide uneven development across analysis, inference, evaluation, and decision making. Discipline-embedded performance tasks can reveal how students marshal spectral evidence and critique claims in context, but they require careful rubrics and scorer training. Standardized measures offer efficiency and comparability, yet they risk construct underrepresentation if divorced from the context of learning. A pragmatic compromise is triangulation: pair a validated measure with embedded prompts scored for analytic decomposition, inferential warrant, evaluative critique, and decision rationale so that general gains can be contrasted with discipline-specific reasoning (Abrami et al., 2008; Guo & Wang, 2021). For measurement clarity, this study treats analytical thinking as decomposing problems and identifying relevant variables, inference as drawing warranted conclusions from spectral evidence, evaluation as judging the plausibility and coherence of claims and representations, and decision making as selecting among parameter settings or solution paths with explicit justification; these operational definitions guide task prompts and scoring.

Replication also matters because single studies can be shaped by cohort features, instructor expertise, or one-time alignments between tasks and assessments. Replicating a design with comparable participants and content can show whether effects persist, shrink, or shift, and can surface elements that are essential. For example, tighter alignment between prompts and the four facets can reduce noise, and clearer fidelity checks can separate a design enacted as intended from one that drifted. Fourier transform is a stringent testbed for such replication: if a design claims to improve analytical thinking, inference, evaluation, and decision making here, the evidence is less likely to be dismissed as an artifact of easy content or lenient tasks (Kohaupt, 2015; Shoenthal, 2014). Conversely, null results would still be informative if they identify where the chain from prediction to explanation to justified choice breaks, guiding revisions to prompts, scaffolds, or feedback cycles rather than encouraging abandonment of inquiry approaches.

Equity and engagement require planning rather than after-the-fact fixes. Access to devices, bandwidth, and quiet study space is uneven. Group work can slide into social loafing if roles are vague and accountability weak, and simulations can privilege students with stronger digital fluency. These risks can depress participation and learning regardless of pedagogy. Sensible guardrails include offline-capable materials where possible, structured collaboration with rotating roles and clear deliverables, and prompts that reward explanation and critique rather than perfunctory agreement. Without such supports, a coherent design can underperform for reasons unrelated to analytical quality, inferential accuracy, evaluative judgment, or decision rationale (Sangster et al., 2020; Novitra et al., 2021; Prahani et al., 2022). Addressing these constraints upfront is not cosmetic; it is part of the causal story about how remote inquiry plus simulation might plausibly affect the four facets.

Taken together, the literature supports a cautious and testable claim. Well-scaffolded remote inquiry that leverages virtual simulation can support critical thinking in demanding STEM topics if tasks require analytic decomposition, defensible inference, careful evaluation of claims and evidence, and explicit decisions about next steps, and if assessments make those moves visible. The present study treats this claim as a hypothesis rather than a conclusion and examines it in the demanding context of a Fourier Transform course for prospective STEM teachers. Specifically, the study aims to examine the effects of virtual simulation-assisted remote inquiry on prospective STEM teachers' critical thinking. Accordingly, the research poses one general question:

- What is the impact of virtual simulation-assisted remote inquiry in a Fourier Transform course on prospective STEM teachers' critical thinking skills?

METHODS

Design and Setting

This study is a direct replication from previous work (Bilad et al., 2022). The replication preserves the course content, LMS workflow and inquiry phases, CT instrument, scoring rubric, and analysis plan; the only planned deviation is a smaller sample size. We employed a randomized pretest–posttest control-group design (Table 1). Two intact classes of prospective STEM teachers (PSTs) at the Mandalika University of Education were randomly assigned to two groups: experimental (E, n = 20) and control (C, n = 20). Both groups were enrolled in the same Fourier Transform course during the same semester and completed identical assessments before (pretest, O_1) and after (posttest, O_2) instruction using an essay-based CT measure aligned to four facets (analytical, inference, evaluation, decision making). Assessments were administered online under proctored conditions. Instructor, topic sequence, and total instructional time were held equivalent across groups.

Table 1. Experimental design (randomized pretest–posttest control-group)

Group	Obs.-1	Treatment	Obs.-2
Experimental (E)	Pretest	Virtual simulation-assisted remote inquiry	Posttest
Control (C)	Pretest	Online learning without inquiry and simulation	Posttest

Intervention and Control Conditions

The intervention integrated a remote inquiry sequence within the university's LMS, organized around adapted phases of e-orientation, e-exploration, e-concept formation, e-application, and e-closure (see Figure 1). Students engaged with virtual simulations to make and test predictions, compare expected and observed spectra, and justify revisions using Fourier principles such as linearity and superposition. Prompts and rubrics were explicitly aligned to four CT facets—analytical thinking, inference, evaluation, and decision making—so that each session required learners to decompose problems, draw warranted conclusions from spectral evidence, appraise competing explanations, and select parameter settings or solution paths with stated reasons. Figure 2 illustrates a representative simulation screen (e.g., PhET "Fourier: Making Waves") used to support these prediction–observation–explanation–justification cycles.

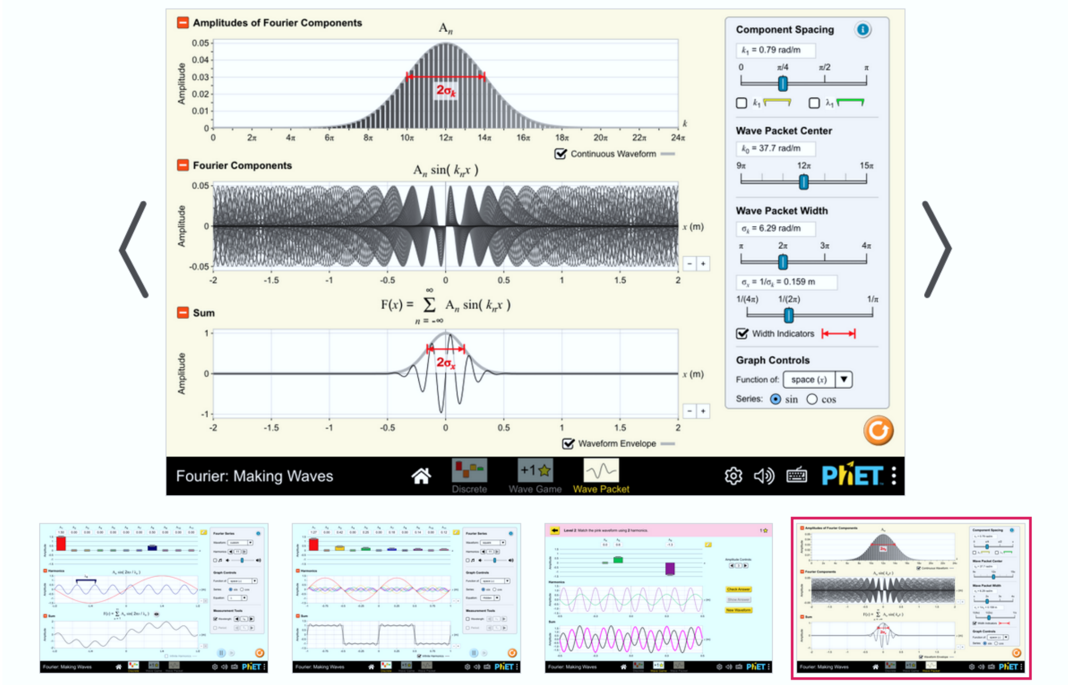

Figure 2. Virtual simulation used as learning content for the the experimental group

Figure 2 illustrates the simulation used in the experimental group: PhET's "Fourier: Making Waves." Students worked inside the LMS with this module to run prediction–observation–explanation–justification cycles. Before interacting, they predicted the spectrum of a target waveform. They then manipulated the Fourier Series panel (selecting waveform type such as square, adjusting the number of harmonics, and tuning each coefficient via sliders labeled A_1, A_2, …), observed changes in the Harmonics plot, and inspected the resulting composite in the Sum plot. Using Graph Controls (space vs. time function, sine/cosine series, equation visibility) and Measurement Tools (e.g., wavelength, period), they gathered evidence to compare with their predictions. Prompts required them to analyze which harmonics dominated a shape (analytical), infer how adding or removing terms altered spectral features (inference), evaluate competing explanations when observations diverged from expectations (evaluation), and decide on parameter settings or next steps with explicit justification grounded in linearity and superposition (decision making).

The control condition consisted of online instruction delivered through the same LMS without inquiry scripting and without simulation. Students received brief readings and mini-lectures followed by worked examples and routine problem sets that matched the intervention's topic sequence, instructor, and total time-on-task. Activities emphasized solution demonstration and end-of-topic quizzes but did not require prior predictions, parameter manipulation, or written justifications tied to the four CT facets. This contrast isolates the added value of inquiry cycles and simulation while holding content coverage and instructional exposure equivalent across groups.

Participants

Forty PSTs participated and were evenly assigned to the experimental (n = 20) and control (n = 20) groups. Demographic characteristics are summarized in Table 2. Ethical approval was obtained from the Ethics Committee of the Faculty of Applied Science and Engineering, Mandalika University of Education. All participants provided informed consent and were assured that their course grades would not be affected by participation.

Table 2. Demographic of the samples

Group	Sample characteristics, n (%)				
	Gender		Age (year)		
	Male (%)	Female (%)	< 18 (%)	18 to 19 (%)	> 19 (%)
Experimental, n = 30	11 (55)	9 (45)	0 (0)	17 (85)	3 (15)
Control, n = 30	10 (50)	10 (50)	1 (5)	17 (85)	2 (10)

Based on Table 2, the experimental group comprised 20 PSTs (11 males, 55%; 9 females, 45%) and the control group also had 20 PSTs (10 males, 50%; 10 females, 50%). Ages were concentrated in the 18–19 range for both groups (17 students each, 85%); only the control group included a participant younger than 18 (1 student, 5%). Students older than 19 accounted for 3 individuals (15%) in the experimental group and 2 individuals (10%) in the control group. Overall, gender distribution was near balanced (21 males, 52.5%; 19 females, 47.5%), and the age profiles were similar across groups, indicating comparable baseline characteristics.

Procedures

Both groups completed the CT pretest in week 1 under proctored online conditions. Over five 90-minute meetings, the experimental group followed an LMS-orchestrated remote inquiry sequence each session: e-orientation to activate prior knowledge and set goals; e-exploration using the PhET "Fourier: Making Waves" simulation to construct/decompose waveforms, adjust harmonics, and observe spectra; e-concept formation to explain how parameter changes mapped to Fourier principles; and e-application to transfer reasoning to novel signals or critique peer explanations. Each meeting concluded with e-closure, where students reflected and received feedback specifically targeting the four CT facets (analytical, inference, evaluation, decision making). Written prompts required predictions before manipulation, comparisons between expected and observed results, and justified revisions.

The control group covered the same topics, instructor, and total time-on-task via the same LMS but without inquiry scripting or simulation. Sessions consisted of brief readings and mini-lectures followed by worked examples and routine problem sets; no prediction–observation–explanation cycles or parameter experimentation were assigned, and written justifications tied to the CT facets were not required. In week 6, both groups completed the CT posttest under the same proctored online conditions as the pretest.

Research Instruments and Analysis

Critical thinking (CT) was measured using an eight-item essay test aligned to four facets targeted in this study: analytical thinking, inference, evaluation, and decision making. Each item elicited written justifications rather than short answers, allowing evidence of reasoning to be scored directly. Prior to implementation, three expert validators reviewed the instrument for content relevance, clarity, and alignment to the four facets; revisions were made accordingly, and the instrument was deemed suitable for use. Each item was scored on a 0–4 scale (0 = no relevant evidence; 4 = complete, well-justified reasoning), yielding a maximum total CT score of 32 per participant.

Scoring guidelines specified observable indicators for each facet to support consistent judgments across items (e.g., identification of relevant variables for analytical thinking; warranted conclusions from given spectral evidence for inference; criteria-based appraisal of competing explanations for evaluation; and explicit rationale for chosen parameters or solution paths for decision making). The resulting total scores were interpreted using the categorical bands shown in Table 3, which classify CT performance from “Not critical” to “Very critical” based on score intervals. These categories were used to describe the distribution of students across levels at pretest and posttest.

Table 3. Categorization of critical thinking skills

Score intervals of CT skill	Category
CTs > 25.60	Very critically
19.20 < CTs ≤ 25.60	Critically
12.80 < CTs ≤ 19.20	Sufficient
6.41 < CTs ≤ 12.80	Less critically
CTs ≤ 6.41	Not critically

For improvement analyses, we computed Hake’s normalized gain (Hake, 1999). Descriptive statistics (mean, standard deviation, and 95% confidence intervals) were reported for pretest, posttest, and (g) in each group. Normality of (g) was examined using the Shapiro–Wilk test ($p > .05$ as the criterion). When the normality assumption was met, between-group differences in gain were tested with an independent-samples t-test at $\alpha = .05$, with hypotheses specified as H_0: no significant difference in CT improvement between groups, and H_a: a significant difference exists.

All analyses were conducted in SPSS 25.0. In addition to significance testing, we reported group means in the Table 3 categories to aid interpretability of practical shifts in performance levels from pretest to posttest. Where relevant, assumptions, test statistics, and exact p-values were presented to ensure transparency of the analytic decisions.

Fidelity and Compliance

To document treatment fidelity in the experimental group, the instructor completed a checklist after each session covering adherence to the five inquiry phases, use of simulation tasks, time allocation, and prompt delivery. LMS logs verified student engagement with simulation activities and submission of prediction and explanation prompts. In the control group, fidelity checks confirmed the absence of inquiry scripting and simulation, as well as parity in total instructional time and topic coverage. Any deviations were recorded and considered in sensitivity analyses.

Ethics and Data Protection

Participation was voluntary. Students could withdraw at any time without penalty. Identifiers were replaced with codes prior to analysis, and de-identified datasets were stored on a secure university server accessible only to the research team. Ethical approval details are noted above; the procedures complied with institutional and national research guidelines.

RESULTS AND DISCUSSION

All participants completed both the pretest and posttest (experimental n = 20; control n = 20). Using the categorization thresholds in Table 3, Table 4 shows the distribution of critical thinking (CT) performance across categories at pretest and posttest for each group. At baseline, both groups clustered in the lower bands (less critical and sufficient), indicating comparable starting points on a challenging topic. After instruction, the experimental group shifted markedly into the upper bands (critical and very critical), while the control group concentrated in sufficient with a modest reduction in less critical.

Table 4. Frequency distribution of CT skill by interval category (n = 20 per group)

Category	Interval	Exp. Group, n (%)		Cont. group, n (%)	
		Pretest	Posttest	Pretest	Posttest
Very critically	CTs > 25.60	0 (0.0)	14 (70.0)	0 (0.0)	0 (0.0)
Critically	19.20 < CTs ≤ 25.60	0 (0.0)	6 (30.0)	0 (0.0)	0 (0.0)
Sufficient	12.80 < CTs ≤ 19.20	4 (20.0)	0 (0.0)	3 (15.0)	18 (90.0)
Less critically	6.41 < CTs ≤ 12.80	15 (75.0)	0 (0.0)	17 (85.0)	2 (10.0)
Not critically	CTs ≤ 6.41	1 (5.0)	0 (0.0)	0 (0.0)	0 (0.0)
Total		20 (100)	20 (100)	20 (100)	20 (100)

At pretest, both groups clustered in the lower bands: the experimental group had 15 students (75%) in Less critical, 4 (20%) in Sufficient, and 1 (5%) in Not critical, with none in the upper categories; the control group showed a similar pattern with 17 (85%) in Less critical and 3 (15%) in Sufficient. After instruction, distributions diverged sharply. The experimental group shifted entirely into the upper bands, with 14 students (70%) in Very critical and 6 (30%) in Critical, indicating a wholesale movement out of the lower categories. By contrast, the control group concentrated in Sufficient at posttest (18 students, 90%), with a small remainder in Less critical (2 students, 10%) and no entries in the upper bands. In short, while both groups improved, only the experimental cohort transitioned from predominantly lower to predominantly upper CT categories. Descriptive statistics are consistent with these shifts (Table 5).

Table 5. Measurement results from CT skill (n = 20 per group)

Group	CT skill (pretest-posttest score)				n-gain	Category
	Pretest	Category	Posttest	Category		
Experimental, n = 30	10.90	Less critically	26.60	Very critically	0.74	High
Control, n = 30	11.20	Less critically	15.10	Sufficient	0.19	Low

The experimental group's mean increased from 10.90 (SD 2.30, Less critical) to 26.60 (SD 2.10, Very critical), yielding a high normalized gain (g = 0.74). The control group increased from 11.20 (SD 2.10, Less critical) to 15.10 (SD 2.40, Sufficient), corresponding to a low gain (g = 0.19). Given the non-overlapping category distributions at posttest and the large mean difference, the practical contrast between groups is substantial. The results in Table 5 are visualized as in Figure 3.

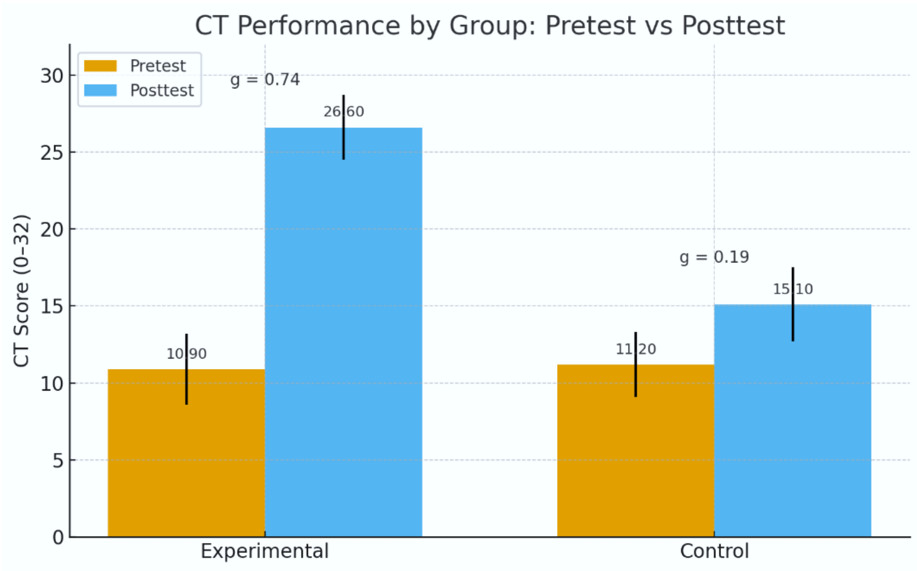

Figure 3. The CT skill measurement results from the experimental and control groups

Figure 3 shows a clear separation in outcomes: the experimental group's mean CT score rose from 10.90 (SD 2.30) to 26.60 (SD 2.10), corresponding to a high normalized gain (g = 0.74), while the control group increased from 11.20 (SD 2.10) to 15.10 (SD 2.40) with a low gain (g = 0.19). Given the identical course, timing, and assessments, this pattern is consistent with the study's aim: virtual simulation–assisted remote inquiry is associated with substantially larger improvements in critical thinking than online instruction without inquiry or simulation. That said, the inference should be read in light of the small sample (n = 20 per group) and near-term posttest; stronger claims would require facet-level analyses and follow-up retention data. Therefore, further analysis was carried out statistically, preceded by a normality test and continued by a difference test (t-test).

The results of the normality test of the data in the two sample groups based on the n-gain parameter are presented in Table 6, and the results of the different test are presented in Table 7.

Table 6. Normality test results based on the n-gain parameter, $p > .05$

Group	Statistic	df	Sig.	Annotation
Experimental	0.957	20	0.391	Normal distribution
Control	0.969	20	0.721	Normal distribution

Table 7. Different test results using independent sample t-test, $p < .05$

Variable	Var. Assumption	Levene's Test		t-test for Equal. of Means		
		F	Sig.	t	df	Sig. (2-tailed)
CT Skills (g)	Equal var. assumed	0.080	0.780	10.94	38	<.001
	Equal var. not assumed	-	-	10.94	36.700	<.001

The Shapiro–Wilk normality checks on gain scores show that both groups meet the normality assumption. In the experimental group the test statistic was 0.957 with a significance of 0.391, and in the control group it was 0.969 with a significance of 0.721. Because both p-values are greater than 0.05, the distributions do not deviate from normal. This supports the use of a parametric comparison for the next step and indicates that any difference we test between groups is unlikely to be an artifact of non-normal data. In short, the data are suitable for an independent-samples t-test, which aligns with the planned analysis and the study's hypothesis about group differences in improvement.

The subsequent independent-samples t-test confirms a clear between-group difference in gains. Levene's test for equality of variances was not significant ($F = 0.080$, $p = 0.780$), so the equal-variances result is the appropriate reference. Under that assumption, the t-test yielded a very large and statistically reliable difference favoring the experimental group ($t = 10.94$, $df = 38$, $p < .001$). Substantively, the group receiving virtual simulation–assisted remote inquiry improved in critical thinking much more than the group receiving online instruction without inquiry or simulation. Interpreted against the hypothesis pair for this study, these findings lead to rejection of the null hypothesis of no difference in improvement and support the alternative hypothesis that the intervention produces a significantly greater gain in critical thinking.

The comparison with the control class indicates a consistent advantage for remote inquiry supported by virtual simulation in a Fourier Transform course, as visible in Tables 4–5 and Figure 3. Students in the intervention moved from lower bands at pretest to the Critical and Very critical bands at posttest, while the control class largely remained in the Sufficient band. Both classes shared the same instructor, content, schedule, and assessments, which narrows alternative explanations linked to exposure or coverage. The distinctive sequence of prediction, evidence checking, explanation, and reasoned decision appears to account for the observed difference in outcomes. Because this sequence aligns directly with what was scored, the gains have a credible mechanism rather than chance alignment. Closing the loop with the statistical results, this pattern is consistent with the tested hypothesis and supports rejection of H0 in favor of the alternative that the intervention yields greater improvement in critical thinking.

Tasks required a stated prediction, focused observation, a testable explanation, and a decision with explicit reasons, creating repeated practice on analysis, inference, evaluation, and decision making. When instructors facilitate questioning and evidence gathering rather than deliver finished answers, students' critical thinking strengthens through visible construction and critique of arguments, which supports designs that make reasoning public and accountable (Khaeruddin & Bancong, 2022). Inquiry classrooms that center questions and evidence report parallel gains when students must show how claims follow from data rather than merely provide results, a feature mirrored in the present intervention (Scott et al., 2018). Facilitation that directs attention to criteria and evidence quality helps learners judge and synthesize information in technology-rich settings, which is vital for dependable growth in critical thinking across courses and cohorts (Suhirman & Prayogi, 2023).

Fourier Transform content is abstract and representation heavy, which makes it difficult for learners to connect symbolic manipulation with conceptual meaning. Virtual simulation brings hidden quantities into view so students can manipulate waveform composition, sampling, and windowing and then watch amplitude and phase respond in real time, which supports hypothesis testing and structured comparison (Lynch & Ghergulescu, 2017). Simulation use is also associated with deeper conceptual understanding and growth in higher-level cognition when tasks demand explanation and justification rather than unsystematic trial and error, a condition met by the present prompts and rubrics (Husnaini & Chen, 2019). These affordances matter in Fourier topics where dynamic relations are easier to see than to infer from static diagrams or isolated algebra, especially for novices who are still learning to map formalism onto physical or informational structure (Husnaini & Chen, 2019).

Within remote formats, technology need not be an obstacle to higher-order thinking when it functions as part of the task environment rather than a delivery pipe. Studies of distance inquiry show that digital tools can help students formulate questions, gather data, and build explanations if those steps are embedded in the workflow and are evaluated with clear expectations, which reduces drift toward passive consumption (Novitra et al., 2021). The present intervention used an LMS to orchestrate inquiry phases alongside simulation so that prediction, observation, explanation, and decision steps were visible, time-stamped, and tied to scoring, which supports accountability without requiring physical presence (Novitra et al., 2021). That structure made reasoning moves inspectable and reviewable, an outcome that is hard to achieve in unstructured forums where contributions lack built-in prompts or evaluative anchors (Novitra et al., 2021).

When inquiry is merged with laboratory-style simulation and both scaffolds and assessments align with cognitive goals, gains in thinking performance and learning outcomes are commonly observed across settings. A recent review reported broad benefits for designs that integrate inquiry with virtual labs and that tie measurement to the intended reasoning moves rather than to procedural speed or recall, which strengthens claims about mechanism and generality (T. L. Lai et al., 2022). In highly abstract domains, simulation can even serve functions once reserved for physical labs because students gain precise control of parameters and immediate feedback, which makes structural relations more visible to novices who otherwise struggle to infer them from sparse signals (T. L. Lai et al., 2022). These conditions match Fourier topics in which spectra change quickly with parameter tweaks that are difficult to stage reliably with limited equipment (T. L. Lai et al., 2022).

Evidence focused on Fourier Transform strengthens the local claim that simulation is well matched to abstract signal topics where traditional demonstrations often struggle to expose underlying structure. Prior work found that virtual simulation nurtures critical thinking in STEM settings and that the approach is particularly suitable for Fourier Transform instruction in which learners must coordinate time- and frequency-domain reasoning under parameter constraints (Bilad et al., 2022). The present results echo that pattern because the jump from lower to higher performance bands occurred only in the class that used simulation within an inquiry sequence, not in the control class that used readings, mini-lectures, and problem sets (Bilad et al., 2022). This alignment between topic demands and representational tools provides a coherent account of why the effect emerged in this context and supports cautious generalization to adjacent signal-and-systems units (Bilad et al., 2022).

Benefits for instructors and students also appear in the management of online instruction, where consistency and traceability matter for fidelity. Simulation platforms can simplify the setup of structured activities, enable monitoring of participation, and capture artifacts for targeted feedback across sessions and cohorts, which reduces drift from intended designs and supports efficient iteration (Cook, 2022). Students often report higher engagement when representations respond directly to their choices, which encourages explanation and critique rather than passive note taking during worked examples or static slides (Hovardas et al., 2018). A synthesis of empirical studies concluded that virtual simulation can lift outcomes that include content understanding, inquiry skills, analytical performance, scientific communication, and social skills, all relevant to teacher preparation where multidimensional competence is required (Brinson, 2015).

Program sustainability is a practical concern in many institutions, and simulation technology broadens what can be taught and assessed when time, cost, or safety limit physical labs. Designers have argued that simulation helps programs maintain investigation experiences while keeping attention on reasoning outcomes that matter for transfer, a priority in resource-constrained environments and hybrid timetables (Delgado & Krajcik, 2010). Others link simulation to inquiry transformations that make goal attainment more likely when instruction must operate at scale across large cohorts or split modalities, conditions common in modern STEM programs (Radhamani et al., 2021). The present design fits these recommendations since it delivered multiple inquiry cycles within ordinary course constraints and kept workload manageable through structured prompts and templated rubrics for scoring and feedback (Radhamani et al., 2021).

Guidance quality remains a boundary condition because unguided exploration can devolve into unsystematic tinkering that leaves inference, evaluation, and decision making underdeveloped. Facilitators who model standards of evidence, press for criteria-based judgments, and ask for reconciliation when predictions fail help ensure that reasoning develops within the simulation-rich workflow and that gains are durable and transferable (Suhirman & Prayogi, 2023). Under that model, simulation functions as a tool that activates problem solving tied to expected STEM outcomes rather than a novelty that distracts from core goals, which addresses a common concern about surface-level engagement in interactive environments (Suhirman & Prayogi, 2023). Future work should sample instructor prompts and student explanations to map which interactions drive change most reliably and to refine guidance for remote cohorts (Suhirman & Prayogi, 2023).

The teacher-preparation context adds practical weight because prospective teachers are mastering content while rehearsing professional orchestration. Integrating simulation into STEM curricula can deepen understanding and empower the processes of critical thinking that matter for future instructional roles in which teachers must press students for reasons rather than mere answers during laboratory-style tasks (Khaeruddin & Bancong, 2022). Emphasis on applying concepts in authentic problem solving strengthens decision making when students must move beyond definitions to workable solutions that can be defended under scrutiny, mirroring the decision prompts used here (Prayogi et al., 2024). Related designs report gains when tasks require pragmatic application in contexts that mirror classroom demands rather than isolated recall, which supports the development of instructional judgment in teacher candidates (Sutoyo et al., 2023).

Equity considerations also emerge through virtual labs because candidates encounter variability in access and readiness inside realistic scenarios that can shape plans for diverse classrooms. Experiencing such constraints during preparation can inform strategies for scaffolding and differentiation that matter for inclusive practice in schools with uneven resources or connectivity across regions and cohorts (Makamure & Tsakeni, 2020). The increasing diversity of teaching contexts demands mastery of content and pedagogy that prioritize critical thinking, which makes the convergence of technology and education a practical necessity rather than an optional enhancement for well-resourced institutions (Salvetti et al., 2023). Designing for inclusion at the outset is consistent with the present approach where structure and expectations were made explicit through the LMS to keep reasoning steps transparent and supportable for all learners (Salvetti et al., 2023).

Limitations temper the claim and keep it aligned with the study aim that guided the statistical tests. The posttest was near in time to instruction, so durability and transfer remain unknown and require delayed measurement once the interface is removed to estimate retention and generalization. The essay instrument and facet rubric were expert-validated, yet reliability would be stronger with double scoring and agreement indices to document consistency across raters in future replications. The sample is smaller than in earlier work, which widens uncertainty even though the pattern of between-group differences is strong. Within these boundaries, the evidence supports the hypothesis-consistent conclusion that virtual simulation-assisted remote inquiry improves critical thinking in this setting.

CONCLUSION

This study tested whether virtual simulation–assisted remote inquiry improves critical thinking in a Fourier Transform course for prospective STEM teachers. With content, instructor, time, and assessments held constant across groups, the intervention class shifted from lower bands at baseline to critical and very critical at posttest, while the control class clustered in sufficient. Mean differences and the t test on normalized gains pointed in the same direction, indicating a larger improvement for the intervention than for online instruction without inquiry or simulation. Read against the stated hypothesis pair, the evidence supports rejection of the null of no difference and acceptance of the alternative that remote inquiry with simulation produces greater growth in critical thinking in this domain.

Evidence from the learning process clarifies why the effect is plausible. Each session began with a prediction that committed students to a claim, followed by observation within the simulation, comparison of expectation and result, an explanation that reconciled discrepancies, and a decision justified by explicit criteria. These moves repeatedly exercised the four targeted facets—analysis, inference, evaluation, and decision making—and the LMS workflow recorded them in ways that could be scored. Practice and measurement were therefore aligned, reducing the gap between what students were asked to do and what counted as evidence of learning. In abstract, representation-heavy topics like Fourier Transform, this alignment made structure visible and gave students repeated, feedback-rich opportunities to reason with spectra rather than follow procedures by rote.

LIMITATIONS

This study has clear constraints that qualify interpretation. The sample was small and clustered, consisting of two intact classes of twenty participants each from a single institution, which narrows external validity and limits generalization to other settings or cohorts.

Assignment occurred at the class level; although baseline measures appeared comparable, unmeasured preexisting differences between sections may persist. Outcomes were collected immediately after instruction, so the durability of gains over time and transfer to related topics, including filtering or convolution, remain untested. These design features argue for cautious claims about both scope and persistence of effects.

Measurement and implementation also introduce uncertainty. The essay instrument and four-facet rubric underwent expert review, yet scripts were not double-scored and interrater agreement indices were not reported, constraining inferences about scoring reliability. Fidelity checks verified adherence to phases but did not analyze how variation in facilitation quality or engagement intensity related to effect sizes. Only one simulation platform and a single inquiry script were employed, limiting inferences to alternative tools or sequencing. Subgroup analyses by gender or age were not conducted, and covariates such as prior GPA or digital fluency were not controlled, leaving potential moderators unexamined.

RECOMMENDATIONS

Future studies should test this approach across multiple institutions using larger, cluster-randomized cohorts, add delayed posttests and transfer tasks in related signal topics to estimate retention and generalization, and report preregistered analyses with open materials. Measurement ought to include double scoring with interrater agreement, facet-level scoring that separately tracks analysis, inference, evaluation, and decision making, and checks of measurement invariance across groups and time. Design experiments should vary simulation platforms, guidance density, and prompt structure to identify minimal reliable configurations, and compare simulation-driven inquiry with active methods such as problem-based learning. Mixed-methods work should analyze student explanations, instructor prompts, and LMS traces to clarify mechanisms and model interaction patterns that predict improvement. Moderation by prior achievement, digital fluency, and demographics should be tested, accompanied by equity audits of device and bandwidth access. Program-level studies should estimate cost, workload, and facilitator training needs, and follow graduates into practicum and early teaching.

Author Contributions

The authors have sufficiently contributed to the study, and have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgment

This study is a cross-university partnership research. The researcher would like to gratitude acknowledged to the university, and the participants involved, they are prospective STEM teachers at the Mandalika University of Education.

Conflict of Interests

The authors declare no conflict of interest.

REFERENCES

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research, 78*(4), 1102–1134. https://doi.org/10.3102/0034654308326084

Arends, R. (2012). *Learning to teach* (9th ed.). McGraw-Hill.

Bailin, S. (2002). Critical thinking and science education. *Science & Education, 11*(4), 361–375. https://doi.org/10.1023/A:1016042608621

Bilad, M., Anwar, K., & Hayati, S. (2022). Nurturing prospective STEM teachers' critical thinking skill through virtual simulation-assisted remote inquiry in Fourier Transform courses. *International Journal of Essential Competencies in Education, 1*(1), 1–10. https://doi.org/10.36312/ijece.v1i1.728

Brinson, J. R. (2015). Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research. *Computers & Education, 87*, 218–237. https://doi.org/10.1016/j.compedu.2015.07.003

Byrnes, K. (2012). A portrait of contemplative teaching: Embracing wholeness. *Journal of Transformative Education, 10*(1), 22–41. https://doi.org/10.1177/1541344612456431

Chinaka, T. W. (2021). The effect of PhET simulation vs. phenomenon-based experiential learning on students' integration of motion along two independent axes in projectile motion. *African Journal of Research in Mathematics, Science and Technology Education, 25*(2), 185–196. https://doi.org/10.1080/18117295.2021.1969739

Cook, J. (2022). Bridging the gap: Computer simulations and video recordings for remote inquiry-based laboratory activities in mechanics. *AEE Journal, 10*(2). https://doi.org/10.18260/3-1-1153-36026

Delgado, C., & Krajcik, J. (2010). Technology supports for science learning. In *International encyclopedia of education* (pp. 197–203). Elsevier. https://doi.org/10.1016/B978-0-08-044894-7.00729-6

Ennis, R. H. (2011). The nature of critical thinking: An outline of critical thinking dispositions and abilities. *Inquiry: Critical Thinking Across the Disciplines, 26*(2), 1–8. https://doi.org/10.5840/inquiryctnews201126214

Fuad, N. M., Zubaidah, S., Mahanal, S., & Suarsini, E. (2017). Improving junior high schools' critical thinking skills based on test three different models of learning. *International Journal of Instruction, 10*(1), 101–116. https://doi.org/10.12973/iji.2017.1017a

Guo, L., & Wang, J. (2021). Relationships between teacher autonomy, collaboration, and critical thinking focused instruction: A cross-national study. *International Journal of Educational Research, 106*, 101730. https://doi.org/10.1016/j.ijer.2020.101730

Hake, R. R. (1999). *Analyzing change/gain scores*. Indiana University.

Halpern, D. F. (2014). *Thought and knowledge: An introduction to critical thinking*. Psychology Press. https://www.routledge.com/Thought-and-Knowledge-An-Introduction-to-Critical-Thinking/Halpern/p/book/9781848726291

Hanson, D. (2005). Designing process-oriented guided-inquiry activities. In *Intellectual development: Instructional design* (pp. 281–284). Pacific Crest.

Hasemi, S. A. (2011). The use of critical thinking in social science textbooks of high school: A field study of Fars Province of Iran. *International Journal of Instruction, 4*(1), 63–78.

Hassard, J. (2005). *The art of teaching science: Inquiry and innovation in middle school and high school*. Oxford University Press.

Hovardas, T., Pedaste, M., Zacharia, Z., & de Jong, T. (2018). Model-based inquiry in computer-supported learning environments: The case of Go-Lab. In M. E. Auer, A. K. M. Azad, A. Edwards, & T. de Jong (Eds.), *Cyber-physical laboratories in engineering and science education* (pp. 241–268). Springer. https://doi.org/10.1007/978-3-319-76935-6_10

Husnaini, S. J., & Chen, S. (2019). Effects of guided inquiry virtual and physical laboratories on conceptual understanding, inquiry performance, scientific inquiry self-efficacy, and enjoyment. *Physical Review Physics Education Research, 15*(1), 010119. https://doi.org/10.1103/PhysRevPhysEducRes.15.010119

Innabi, H., & Sheikh, O. E. (2007). The change in mathematics teachers' perceptions of critical thinking after 15 years of educational reform in Jordan. *Educational Studies in Mathematics, 64*(1), 45–68. https://doi.org/10.1007/s10649-005-9017-x

Khaeruddin, K., & Bancong, H. (2022). STEM education through PhET simulations: An effort to enhance students' critical thinking skills. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni, 11*(1), 35–45. https://doi.org/10.24042/jipfalbiruni.v11i1.10998

Kohaupt, L. (2015). Introduction to the discrete Fourier series considering both mathematical and engineering aspects: A linear-algebra approach. *Cogent Education, 2*(1), 1064560. https://doi.org/10.1080/2331186X.2015.1064560

Lai, E. R. (2011). *Critical thinking: A literature review* (pp. 1–50). Pearson. https://images.pearsonassessments.com/images/tmrs/CriticalThinkingReviewFINAL.pdf

Lai, T.-L., Lin, Y.-S., Chou, C.-Y., & Yueh, H.-P. (2022). Evaluation of an inquiry-based virtual lab for junior high school science classes. *Journal of Educational Computing Research, 59*(8), 1579–1600. https://doi.org/10.1177/07356331211001579

Lam, M.-Y. L., Lim, S. E., Ma, J. C., & Adams, L. (2003). What Hong Kong teachers and parents think about thinking. *Early Child Development and Care, 173*(1), 147–158. https://doi.org/10.1080/0300443022000022495

Llewellyn, D. (2001). *Inquire within: Implementing inquiry-based science standards*. Corwin Press.

Lynch, T., & Ghergulescu, I. (2017). Review of virtual labs as the emerging technologies for teaching STEM subjects. *INTED2017 Proceedings*, 7681–7690. https://doi.org/10.21125/inted.2017.1422

Makamure, C., & Tsakeni, M. (2020). COVID-19 as an agent of change in teaching and learning STEM subjects. *Journal of Baltic Science Education, 19*(6A), 1078–1091. https://doi.org/10.33225/jbse/20.19.1078

Novitra, F., Festiyed, F., Yohandri, Y., & Asrizal, A. (2021). Development of online-based inquiry learning model to improve 21st-century skills of physics students in senior high school. *Eurasia Journal of Mathematics, Science and Technology Education, 17*(9), em2004. https://doi.org/10.29333/ejmste/11152

Prahani, B. K., Amiruddin, M. Z. B., Jatmiko, B., Suprapto, N., & Amelia, T. (2022). Top 100 cited publications for the last thirty years in digital learning and mobile learning. *International Journal of Interactive Mobile Technologies, 16*(8), 18–33. https://doi.org/10.3991/ijim.v16i08.29803

Prayogi, S., Bilad, M., Verawati, N., & Asy'ari, M. (2024). Inquiry vs. inquiry-creative: Emphasizing critical thinking skills of prospective STEM teachers in the context of STEM learning in Indonesia. *Education Sciences, 14*(6), 593. https://doi.org/10.3390/educsci14060593

Prayogi, S., Muhali, M., Yuliyanti, S., Asy'ari, M., Azmi, I., & Verawati, N. N. S. P. (2019). The effect of presenting anomalous data on improving student's critical thinking ability. *International Journal of Emerging Technologies in Learning, 14*(6), 133–141. https://doi.org/10.3991/ijet.v14i06.9717

Pursitasari, I. D., Suhardi, E., Putra, A. P., & Rachman, I. (2020). Enhancement of students' critical thinking skill through science context-based inquiry learning. *Jurnal Pendidikan IPA Indonesia, 9*(1), 97–105.

Radhamani, R., Kumar, D., Nizar, N., Achuthan, K., Nair, B., & Diwakar, S. (2021). What virtual laboratory usage tells us about laboratory skill education pre- and post-COVID-19: Focus on usage, behavior, intention and adoption. *Education and Information Technologies, 26*(6), 7477–7495. https://doi.org/10.1007/s10639-021-10583-3

Rubini, B., Ardianto, D., Pursitasari, I. D., & Permana, I. (2016). Identify scientific literacy from the science teachers' perspective. *Jurnal Pendidikan IPA Indonesia, 5*(2), 299–303. https://doi.org/10.15294/jpii.v5i2.7689

Salvetti, F., Rijal, K., Owusu-Darko, I., & Prayogi, S. (2023). Surmounting obstacles in STEM education: An in-depth analysis of literature paving the way for proficient pedagogy in STEM learning. *International Journal of Essential Competencies in Education, 2*(2), 177–196. https://doi.org/10.36312/ijece.v2i2.1614

Sangster, A., Stoner, G., & Flood, B. (2020). Insights into accounting education in a COVID-19 world. *Accounting Education, 29*(5), 431–562. https://doi.org/10.1080/09639284.2020.1808487

Scott, D., Smith, C., Chu, M., & Friesen, S. (2018). Examining the efficacy of inquiry-based approaches to education. *Alberta Journal of Educational Research, 64*(1), 35–54. https://doi.org/10.55016/ojs/ajer.v64i1.56439

Shoenthal, D. (2014). Fourier series as a unifying topic in Calculus II. *PRIMUS, 24*(4), 294–300. https://doi.org/10.1080/10511970.2013.870265

Silva, H., Lopes, J., Cruz, G., Dominguez, C., & Morais, E. (2022). Does university attendance affect students' critical and creative thinking skills? A longitudinal research with pre-service teaching and psychology undergraduates. *Higher Education Research & Development*, 1–11. https://doi.org/10.1080/07294360.2022.2057448

Suhirman, S., & Prayogi, S. (2023). Overcoming challenges in STEM education: A literature review that leads to effective pedagogy in STEM learning. *Jurnal Penelitian Pendidikan IPA, 9*(8), 432–443. https://doi.org/10.29303/jppipa.v9i8.4715

Sutoyo, S., Sanjaya, I., Susantini, E., Cahyono, E., Abdullah, N., & Allamin, S. (2023). Improving students' critical thinking skills and self-efficacy through implementation of integrated guided inquiry model with science, technology, engineering and mathematics (STEM). In *Proceedings* (pp. 65–74). https://doi.org/10.2991/978-2-38476-078-7_9

Thaiposri, P., & Wannapiroon, P. (2015). Enhancing students' critical thinking skills through inquiry-based learning activities using social network and cloud computing. *Procedia – Social and Behavioral Sciences, 174*, 2137–2144. https://doi.org/10.1016/j.sbspro.2015.02.013

Tiruneh, D. T., De Cock, M., Weldeslassie, A. G., Elen, J., & Janssen, R. (2017). Measuring critical thinking in physics: Development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education, 15*(4), 663–682. https://doi.org/10.1007/s10763-016-9723-0

Verawati, N. N. S. P., Hikmawati, H., Prayogi, S., & Bilad, M. R. (2021). Reflective practices in inquiry learning: Its effectiveness in training pre-service teachers' critical thinking viewed from cognitive styles. *Jurnal Pendidikan IPA Indonesia, 10*(4), 505–514. https://doi.org/10.15294/jpii.v10i4.31814

Wahidin, D., & Romli, L. M. (2020). Students' critical thinking development in National Sciences and Mathematics Competition in Indonesia: A descriptive study. *Jurnal Pendidikan IPA Indonesia, 9*(1), 106–116. https://doi.org/10.15294/jpii.v9i1.22240

Wahyudi, W., Verawati, N. N. S. P., Ayub, S., & Prayogi, S. (2019). The effect of scientific creativity in inquiry learning to promote critical thinking ability of prospective teachers.

International Journal of Emerging Technologies in Learning, 14(14), 122–131. https://doi.org/10.3991/ijet.v14i14.9532

Wietecha, T., Kurzydło, P., & Bogacz, B. (2021). Introduction to Fourier analysis using IT tools of Laboratory of Technical Aids of Teaching IF UJ. *Science, Technology and Innovation, 12*(1), 10–15. https://doi.org/10.5604/01.3001.0014.8499