

A deep learning-integrated RADEC model for higher-order thinking in preservice elementary teachers' environmental science learning

Idam Ragil Widiyanto Atmojo^{1*} , Ananda Ayu Rahmadhani¹ , Matsuri¹ , Roy Ardiansyah¹ ,
Dwi Yuniasih Saputri¹ , Fadhil Purnama Adi¹ 

¹ Universitas Sebelas Maret, Surakarta, INDONESIA

Received 07 July 2026 ▪ Accepted 17 October 2026

Abstract

In this study, deep learning (DL) refers to a pedagogical approach—not an artificial intelligence or machine learning technique—centered on meaningful, mindful, and joyful engagement with content (Kemendikdasmen, 2025). Higher-order thinking skills (HOTS)—analyzing, evaluating, and creating—are essential for preservice elementary teachers, who must develop HOTS themselves before they can teach it to pupils, yet HOTS remain underdeveloped in early teacher-education coursework. This study tested whether integrating Indonesia's DL principles (meaningful, mindful, and joyful learning) into the read-answer-discuss-explain-create (RADEC) model improves HOTS more than direct instruction. Using a true experimental, pre-/post-test control-group design with random assignment to condition, 172 third-semester primary school teacher education students at a public university in Indonesia (86 experimental, 86 control) completed a unit on environmental problems. HOTS was measured with a 20-item multiple-choice test spanning the analyze, evaluate, and create dimensions (Cronbach's alpha = .859). The experimental group's post-test scores (mean [M] = 77.56, standard deviation [SD] = 9.904) were significantly higher than the control group's (M = 69.42, SD = 10.72), $t(170) = 5.17$, $p < .001$, $d = 0.79$, with descriptive gains concentrated in evaluate (+16.94 points) and create (+12.21 points), suggesting where RADEC's discuss-explain-create sequence may contribute most. A one-way multivariate analysis of covariance confirmed that these dimension-level gains were statistically significant after controlling for pre-test scores, Wilks' lambda = .82, $F(3, 165) = 11.77$, $p < .001$, partial $\eta^2 = .18$, with significant univariate effects for analyze, evaluate, and create considered separately.

Keywords: science education, RADEC learning model, deep learning, higher-order thinking skills, preservice elementary teachers, environmental problem-solving

INTRODUCTION

Higher-Order Thinking in Science Education

21st century education places a premium on learners who can think analytically and reflectively, and on teachers who can facilitate that kind of thinking in the classroom (Tahim Bael et al., 2021). Higher-order thinking skills (HOTS)—the abilities to analyze, evaluate, and create (Andini & Fitria, 2021)—together with problem-solving, rational decision-making, and argument construction (Fauziana et al., 2021), sit at the center of this demand. Mastering HOTS allows learners

to approach problems critically, creatively, and logically, and to generate productive, innovative ideas (Noor & Abadi, 2022). This emphasis on higher-order cognitive skills (HOCS) echoes a longer line of research on constructivist and inquiry-oriented instruction: Aydın and Yılmaz (2010) found that a constructivist approach measurably raised students' HOCS relative to conventional teaching, while Zoller (2000) argued that interdisciplinary, systemic HOCS development is central to meaningful science-technology-environment-society (STES)-oriented science education. Zoller et al. (2002) further showed that students perform differently, and hold different attitudes, on lower-order cognitive skills

Contribution to the literature

- This study provides one of the first empirical tests of a deep learning (DL)-integrated read-answer-discuss-explain-create (RADEC) model on the HOTS of preservice elementary teachers in science learning, a population and pairing not previously examined together.
- Unlike prior RADEC research that reports HOTS as a single composite score, this study disaggregates outcomes by the three highest cognitive dimensions of the revised Bloom's taxonomy—analyze, evaluate, and create—revealing that the model's benefit is concentrated in evaluate and create rather than distributed evenly.
- By reporting pilot-stage, stage-by-stage implementation-fidelity observations alongside main-study dimension-level outcomes, the study offers preliminary, hypothesis-generating indications of which RADEC stages may be associated with which HOTS dimension, giving teacher educators a tentative instructional target rather than a general recommendation to “use more active learning.”

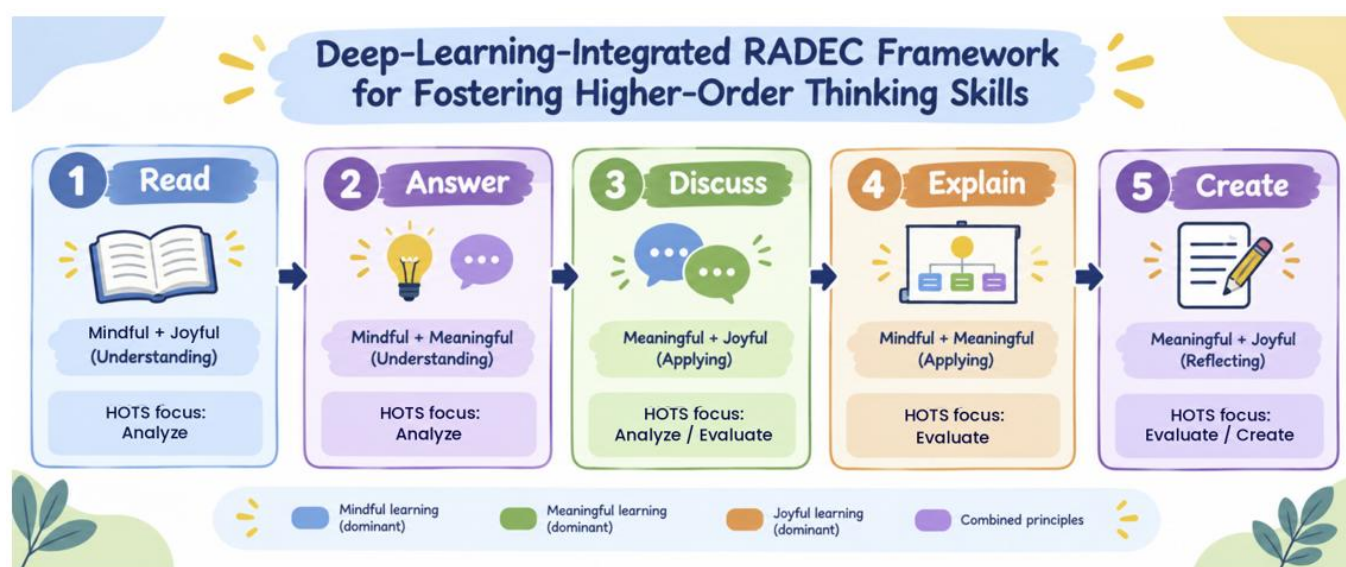


Figure 1. DL-integrated RADEC framework, showing the five stages, their dominant DL principle(s) and experience, and the HOTS dimension each stage primarily targets (the authors' own work created using Canva)

vs. HOCS examination questions, underscoring that HOCS cannot simply be assumed to follow from content mastery, and Renaud (2002) demonstrated that instruction built around higher-order questions can directly strengthen students' critical-thinking skills. Together, this literature supports the premise that HOTS/HOCS development requires instructional models—such as the DL-integrated RADEC model examined here—that deliberately and systematically target higher-order cognitive demand rather than leaving it to develop incidentally (Figure 1).

For preservice elementary teachers, HOTS occupies a dual position: it is a competency to be developed during coursework and, simultaneously, a capacity these future teachers will need in order to cultivate HOTS in their own pupils. Zhou et al. (2023) argue that HOTS is an essential competency for preservice teachers that must be measured systematically. Yet students—including preservice elementary teachers—who are unaccustomed to higher-order questions tend to struggle when problems demand contextual reasoning (Wardhani & Oktiningrum, 2022). Conversely, well-designed

instructional models can raise preservice teachers' HOTS while simultaneously modeling pedagogy they can later use themselves (Agustin et al., 2021), and learning experiences grounded in authentic problems can support both scientific literacy and pedagogical readiness (Cholily et al., 2022; Velazquez-Ulloa & Finkel, 2025).

HOTS deficits are not confined to K-12 settings; they persist in higher education, particularly around environmental problem-solving. Mualimin and Pamungkas (2024) reported that Indonesian undergraduates' holistic higher-order thinking averaged only 24.81 among biology majors and 27.48 among biology education majors (overall mean [M] = 26.23), a “very low” category. Budiarti et al. (2024) similarly found that among 16 third-semester students tested on analyze (C4), evaluate (C5), and create (C6), 75% fell in the moderate category and only 12.5% reached the high category. These findings indicate that HOTS development in the early stages of higher education still requires deliberate attention.

A pre-study conducted by the authors at the participating university reinforced this picture. Interviews suggested that students tend to default to practical, surface-level reasoning and need substantial lecturer scaffolding to develop solutions independently, with uneven performance across analyze, evaluate, and create. A pre-test administered as part of the same pre-study found 52% correct answers on analyze (C4), 51.67% on evaluate (C5), and 45% on create (C6), for an overall mean of 50.5—consistent with the broader literature’s characterization of HOTS among Indonesian preservice teachers as underdeveloped.

These difficulties are linked to learning experiences that provide too few opportunities for analytical, independent, contextual thinking. Yulianto et al. (2024) showed that habitually answering mechanical, decontextualized questions leaves students struggling once a question demands deeper analysis. Kurniawan et al. (2019) likewise linked low reading interest, little independent practice, and minimal exposure to contextual problems to the same difficulty. What is needed, therefore, is instruction that actively and meaningfully engages students while systematically pushing them to analyze, evaluate, and generate ideas or solutions.

Science learning offers a natural context for this kind of HOTS development, since it is not oriented solely toward mastering facts, concepts, and principles but also toward discovery through investigation and experimentation (Manalu et al., 2026). Science instruction should give students genuine opportunities to engage in scientific inquiry and build scientific literacy (Velazquez-Ulloa & Finkel, 2025); in higher education specifically, this means guiding students to analyze information, interpret findings, and discuss results in scientifically grounded ways (Din & MacInnis, 2024; Jin et al., 2023; Resendes, 2022).

Tertiary science instruction therefore needs to be designed not only to build conceptual mastery but also to provide investigative experiences that push students to analyze, evaluate, and generate solutions to scientific problems (Setyawarno & Kurniawati, 2022). This points to the need for an instructional model that provides systematic stages for building understanding, driving interaction and reasoning, and directing students toward generating ideas or solutions.

RADEC as an Active Learning Model

The RADEC model accommodates 21st century learning demands by actively involving students in constructing knowledge and developing HOTS. Tulljanah and Amini (2021) found that RADEC has the potential to raise HOTS in science learning, while Anggraeni et al. (2021) showed that inquiry-oriented RADEC can develop critical-thinking skills among preservice elementary teachers.

RADEC unfolds through five interlocking stages: read, in which students build initial understanding; answer, in which they process information through pre-learning questions; discuss, in which they exchange and compare ideas; explain, in which they communicate their understanding; and create, in which they generate ideas or solutions from what they have learned (Sopandi et al., 2023). This sequence positions students as active learners who acquire, process, discuss, explain, and produce knowledge in turn. Read and answer build initial understanding; discuss and explain drive elaboration, reasoning, and the exchange of ideas; and create provides space for generating ideas or solutions. RADEC thus offers a systematic instructional structure while giving students repeated opportunities for the cognitive activity that underlies HOTS development.

Deep Learning as a Pedagogical Approach

DL is a pedagogical approach that emphasizes thorough, comprehensive understanding rather than rote memorization (Suar Adnyana, 2024; Widyaningrum et al., 2025). As formulated in Indonesia’s national curriculum reform, it rests on three principles—mindful learning, meaningful learning, and joyful learning (Kemendikdasmen, 2025). Mindful learning cultivates students’ awareness of and engagement with their own learning process; meaningful learning connects new knowledge to prior experience and real-world context; and joyful learning creates a positive climate that supports motivation, creativity, and active participation (Diputera et al., 2024; Kemendikdasmen, 2025).

Although RADEC offers a systematic instructional structure, its implementation faces two recurring challenges: it requires a relatively long instructional time, and less confident students may participate less in discussion and presentation (Dahlia & Tirtoni, 2025; Suryana et al., 2025). DL can strengthen RADEC’s implementation precisely by supplying the meaningful, mindful, and joyful principles that address these gaps. Integrating RADEC with DL has also been reported to produce active, contextual, and meaningful learning (Lailiyati & Afandi, 2026). RADEC, in other words, supplies the instructional structure, while DL strengthens the quality of the learning experience within that structure. The integration of DL, RADEC, and HOTS is summarized in **Table 1**.

Integrating Deep Learning into RADEC

Integrating DL into RADEC means orienting the learning experience at every stage—read, answer, discuss, explain, and create—toward understanding, applying, and reflecting. Read and answer help students understand information independently; discuss drives the exchange of ideas; explain requires communicating understanding; and create generates ideas or solutions (Sopandi et al., 2023).

Table 1. Conceptual framework integrating DL, RADEC, and HOTS

RADEC stage	DL principle → experience	Student activity and HOTS orientation
Read	Mindful, joyful → understanding	Students read material on environmental problems independently and connect it to prior knowledge, in a focused, comfortable learning climate. Primarily supports analyze.
Answer	Mindful, meaningful → understanding	Students independently answer pre-learning questions based on their reading, logically connecting new information to prior knowledge. Primarily supports analyze.
Discuss	Meaningful, joyful → applying	Students discuss in groups to compare answers, exchange views, and apply concepts to a group task. Supports analyze and, especially, evaluate.
Explain	Mindful, meaningful → applying	Groups present their discussion results, explain concepts, and respond to questions from peers, supporting justification and evaluation of claims. Primarily supports evaluate.
Create	Meaningful, joyful → reflecting	Groups synthesize their understanding into a poster proposing a solution to an environmental problem, then reflect on the idea developed. Primarily supports evaluate and create.

Note. The RADEC syntax follows the understanding → applying → reflecting sequence used in Indonesia's DL pedagogy

At each stage, the mindful, meaningful, and joyful principles strengthen the quality of the learning process through the experiences of understanding, applying, and reflecting (Kemendikdasmen, 2025). This integration gives students room to analyze, evaluate, and generate ideas or solutions, consistent with RADEC's established support for HOTS development (Agustin et al., 2021) and with DL's orientation toward higher-order thinking (Zainuddin et al., 2026). The integration therefore preserves RADEC's systematic structure while deepening the quality of students' learning experience.

Research Gap

Prior research points to RADEC's potential for supporting active learning and HOTS development. Agustin et al. (2021) found that RADEC positively affects preservice elementary teachers' HOTS, while Rahayu et al. (2021) showed that problem-oriented RADEC affects preservice elementary teachers' critical thinking. At the tertiary level, Hayati et al. (2023) found that RADEC improved conceptual understanding in an organic chemistry course, indicating that the model adapts well to higher education.

On a separate but related track, Nurdiana and Pradana (2025) found DL effective for improving junior-secondary science learning outcomes, and Lailiyati and Afandi (2026) showed that integrating RADEC with DL can produce active, contextual, and meaningful learning that supports elementary pupils' concept understanding and motivation.

Beyond RADEC specifically, the broader active-learning literature offers converging evidence that structured, scaffolded instruction outperforms conventional teaching for HOTS. In physics education, an instructional module combining scaffolding strategies with guided discovery produced significantly larger HOTS gains among Saudi technical-college

students than conventional teaching, with a large analysis of covariance (ANCOVA) effect (partial $\eta^2 = .585$; Alanazi et al., 2024). On a broader scale, a systematic review and meta-analysis of 36 studies found that inquiry-based learning has a substantial positive effect on students' critical thinking (Hedges' $g = 1.27$, 95% confidence interval [0.78, 1.76]), although the effect varies considerably by education level, duration, and assessment type (Arifin et al., 2025). Structured, staged active-learning designs outside science have shown comparable benefits: a digital-book, project-based STEM intervention meaningfully improved junior-secondary students' mathematical critical thinking on number-pattern tasks (Pramasdyahsari et al., 2023). At the same time, not every active-learning gain is uniform: a quasi-experimental study of problem-based learning in physics found a significant, above-average effect on problem-solving skills but no significant change in motivation to learn physics (Argaw et al., 2017), a reminder that cognitive and affective outcomes do not always move together.

Despite this convergent evidence for active learning in general, research that integrates RADEC with DL specifically to test its effect on preservice elementary teachers' HOTS remains limited, particularly for third-semester students in science coursework. Prior RADEC and DL studies have also tended to report learning outcomes, conceptual understanding, or a single thinking construct as an undifferentiated whole, whereas studies that specifically examine HOTS through the analyze, evaluate, and create dimensions in preservice elementary teachers' science learning remain scarce.

This study was therefore conducted to examine the effect of integrating DL into the RADEC model on the HOTS of third-semester preservice elementary teachers in a science unit on environmental problems.

Table 2. Observed implementation fidelity across the five RADEC stages (pilot study experimental group, n = 20)

RADEC stage	Observed behavior	PM (%)	Category
Read	Reading, understanding, and connecting material to prior knowledge.	58.75	Fair
Answer	Answering questions independently, logically, and consistently with the material.	63.75	Good
Discuss	Discussing, exchanging views, and applying concepts in a group task.	66.25	Good
Explain	Presenting results, explaining concepts, and responding to questions.	76.25	Good
Create	Producing a poster, generating ideas, and reflecting on the proposed solution.	75.00	Good

Note. Data were collected during the pilot study, not the main study; Ratings were made on a 1-4 scale by a single observer; PM: Percentage of maximum = Observed score/(80 × 100)

Purpose and Research Questions

This study aimed to examine the effect of integrating a DL approach with the RADEC learning model on the HOTS of preservice elementary teachers in a science unit on environmental problems.

RQ1. Does integrating a DL approach with the RADEC model significantly affect preservice elementary teachers' HOTS in science learning on environmental problems?

RQ1a. Does integrating a DL approach with the RADEC model significantly affect preservice elementary teachers' analyze (C4) scores?

RQ1b. Does integrating a DL approach with the RADEC model significantly affect preservice elementary teachers' evaluate (C5) scores?

RQ1c. Does integrating a DL approach with the RADEC model significantly affect preservice elementary teachers' create (C6) scores?

MATERIALS AND METHODS

Research Design

This study used a quantitative, true experimental (randomized) pre-/post-test control-group design, with students individually randomly assigned to condition (see participants and context). The experimental group received instruction using the RADEC model integrated with the DL approach, while the control group received direct instruction with a scientific approach. Both groups completed a pre-test before treatment and a post-test afterward to measure HOTS; the difference in post-test results between groups was analyzed to test for a difference in HOTS following instruction.

Participants and Context

This study was conducted among third-semester students in the primary school teacher education (PGSD) program at a public university in Indonesia, from a population of 300 students. A sample of 172 students was selected from this population using *simple random sampling* and then divided into two equal groups of 86 students each. Assignment to the experimental and control conditions was carried out at random (*random assignment*), yielding an experimental group of 86 students and a control group of 86 students (N = 172 total). Condition was assigned at the individual level

through randomization rather than by pre-existing intact classes, which gives the design the structure of a randomized experiment; after assignment, each condition was taught in its own separate class, so class-level effects cannot be fully separated from the treatment effect (see limitations). The pre-/post-test control-group structure is retained throughout the manuscript.

Study Stages and Data Collection Timeline

The research was conducted in two stages. The first was a pilot study with 42 third-semester PGSD students from two universities in Indonesia (20 in the experimental group and 22 in the control group; a quasi-experimental two-group design using two intact university cohorts), carried out to prepare and refine the instruments and the implementation of the intervention for the main study. The implementation-fidelity observations reported in **Table 2** were collected during this pilot stage. The second was the main study, conducted with the 172 students described above. Pilot-study participants were not included in the main-study sample, and none of the analyses reported in the results section uses pilot-study data.

All main-study data (pre-test, intervention, and post-test; N = 172) were collected before this manuscript was first submitted to the journal; no participants were added and no new data were collected during the revision stages. The main study was conducted in the 2025/2026 academic year: the pre-test was administered in September 2025, the intervention was implemented between October and December 2025, and the post-test was administered in January 2026. The pilot study was conducted between June and August 2025, before the main study.

Following random assignment, the experimental and control groups of the main study were taught in separate classes by different lecturers. The lecturers were comparable in average teaching experience and educational background (all held doctoral degrees), and all had taken part in a meeting to align perceptions among the course instructors before the study.

Deep Learning-Integrated RADEC Intervention

DL was integrated into every RADEC stage by orienting student activity toward the experiences of understanding, applying, and reflecting. The intervention was implemented over eight weeks: a two-

hour face-to-face treatment session was held once a week and was followed by an analysis week, giving four treatment sessions and four analysis weeks, so that the total treatment time was eight hours. In each analysis week, the researchers and the lecturer prepared the lesson plan for the next treatment session and reviewed how the RADEC stages had been implemented in the preceding session. The implementation of each RADEC stage is described below.

Read

In the read stage, students independently read instructional material on environmental problems, then identified and connected the information to their prior knowledge. Reading was guided so that students focused on the material and remained aware of the learning objective; a comfortable learning atmosphere and purposeful reading activity were designed to make the experience enjoyable. The read stage thus provides an understanding-oriented learning experience reflecting the mindful and joyful principles.

Answer

In the answer stage, students independently answered pre-learning questions based on their reading and prior knowledge, composing logical, content-consistent answers. This activity pushes students to consciously process their understanding while connecting new information to what they already know. The answer stage therefore provides an understanding-oriented learning experience reflecting the mindful and meaningful principles.

Discuss

In the discuss stage, students discuss in groups to present and compare answers, exchange views, and apply the concepts they have understood to a group task. This interaction gives students a chance to use their knowledge in the context of the problem under discussion and to build understanding through the exchange of ideas; the collaborative activity also creates an interactive, enjoyable atmosphere. The discuss stage therefore provides an applying-oriented learning experience reflecting the meaningful and joyful principles.

Explain

In the explain stage, students present their group's discussion results, explain the concepts they have understood, and respond to questions or comments from other groups. Students consciously monitor their own understanding through the process of explaining and responding, connecting the concepts studied with their group's reasoning. This stage therefore provides an applying-oriented learning experience reflecting the mindful and meaningful principles.

Create

In the create stage, students work in groups to produce a poster related to an environmental problem, translating the understanding gained in earlier stages into a proposed solution. After producing this artifact, students reflect on the ideas and solutions they developed. The freedom to decide on both the idea and its presentation gives students room to express themselves creatively in an enjoyable learning atmosphere. The create stage therefore provides a reflecting-oriented learning experience reflecting the meaningful and joyful principles.

Intervention Fidelity

Implementation fidelity was assessed during the pilot study using an observation checklist completed by a single observer, who observed the 20 students of the pilot experimental group at each RADEC stage. Each fidelity indicator was rated on a four-point scale (1 = lowest, 4 = highest implementation), with every RADEC stage evaluated against indicators specific to that stage's DL orientation, as summarized in **Table 2**. In the main study, the intervention followed the same RADEC syntax and lesson-plan format as the pilot study, as described above, but implementation fidelity was not systematically observed: the weekly review in the analysis weeks did not use a structured rating instrument. The fidelity ratings therefore document how the procedure was implemented at the pilot stage and should not be interpreted as a measure of implementation fidelity for the 86 students in the main-study experimental group (see limitations).

Instruments

Higher-order thinking skills test

The main instrument was a HOTS test administered as a pre- and post-test. The test consisted of multiple-choice items built around a science unit on environmental problems and mapped onto the revised Bloom's taxonomy of Anderson and Krathwohl (2001), covering three higher-order cognitive processes: analyze (C4), evaluate (C5), and create (C6). Analyze covered differentiating, organizing, and attributing; evaluate covered checking and critiquing; and create covered generating and planning. Each correct answer was scored 1 and each incorrect answer 0.

Validity and reliability

The instrument was built around seven HOTS indicators—differentiating, organizing, attributing, checking, critiquing, generating, and planning. From an initial pool of 35 items, content validity was established by four experts using Aiken's V (criterion $V \geq .92$). The instrument was then piloted with 60 students outside the study sample to test construct validity via Pearson

Table 3. Descriptive statistics for pre- and post-test HOTS scores

Group	n	M (pre-test)	SD (pre-test)	M (post-test)	SD (post-test)
Experimental	86	64.42	10.53	77.56	9.90
Control	86	63.95	12.15	69.42	10.72

Note. Scores are expressed on a 0-100 scale

product-moment correlation, item difficulty ($0.30 \leq P \leq 0.70$), and item discrimination ($D > 0.30$) using SPSS version 22. This process yielded 20 valid, usable items with acceptable reliability (Cronbach's $\alpha = .862$).

Data Analysis

HOTS data were analyzed using descriptive statistics (minimum, maximum, mean, and standard deviation) for pre- and post-test scores in both groups. Prior to hypothesis testing, normality was assessed with Shapiro-Wilk and homogeneity of variance with Levene's test, using SPSS version 22. Baseline equivalence between groups was tested by comparing pre-test scores with an independent-samples t-test (Patimah & Widianjani, 2022). The main hypothesis was then tested with an independent-samples t-test comparing post-test HOTS means between the experimental and control groups at $\alpha = .05$ (Widodo et al., 2023). To aid interpretation, the authors additionally computed Cohen's d for the post-test comparison from the reported t-value and group sizes.

Ethical Considerations

All participants gave informed consent, and no individually identifying information is reported in this manuscript. This study was approved by the Research Ethics Committee at Universitas Sebelas Maret with approval number 3190/UN27.02.7.13/EC/2026. The ethics approval letter was issued on 17 April 2026, after data collection had been completed, owing to a prolonged administrative process; the research protocol had been submitted to the committee on 29 August 2025, before the main-study data collection began. The pilot study (June-August 2025) was conducted before this submission. Written informed consent was obtained from all participants before data collection in both the pilot and main studies. Participation was voluntary. Participants' names were replaced with code numbers in the analysis dataset, and the data are stored securely and accessible only to the research team.

RESULTS

Implementation of Deep Learning-Integrated RADEC

Implementation of the DL-integrated RADEC model was observed across the 20 students of the pilot-study experimental group at each stage of the syntax (**Table 2**). The maximum possible score at each stage was 80 (a maximum of 4 points per student). Observation showed that read scored 47 (58.75%, fair), answer scored 51

(63.75%, good), discuss scored 53 (66.25%, good), explain scored 61 (76.25%, good), and create scored 60 (75.00%, good). Overall, four of the five stages fell in the good category, while read fell in the fair category. The highest fidelity score was observed for explain (76.25%) and the lowest for read (58.75%) (**Table 2**). These ratings come from the pilot stage ($n = 20$) and are reported to document how the procedure was implemented, not as fidelity data for the main study.

Descriptive Statistics

Table 3 presents descriptive statistics for students' HOTS scores. Mean HOTS scores increased in both groups following instruction. In the experimental group, the mean score rose significantly from $M = 64.42$ ($SD = 10.53$) at pre-test to $M = 77.56$ ($SD = 9.90$) at post-test. In the control group, by contrast, the increase was more limited, from $M = 63.95$ ($SD = 12.15$) to $M = 69.42$ ($SD = 10.72$). In addition, the maximum score in the experimental group rose from 95 to 100, while the maximum score in the control group rose from 90 to 95 (**Table 3**).

Effect on HOTS

Before hypothesis testing, assumption checks were conducted. Shapiro-Wilk tests indicated that pre- and post-test scores were normally distributed in both groups: in the control group, $p = .072$ for pre-test and $p = .059$ for post-test; in the experimental group, $p = .097$ for pre-test and $p = .057$ for post-test. All values exceeded $\alpha = .05$, so the data were treated as normally distributed. A mean-centered Levene's test further indicated that the variances of the two groups were homogeneous at pre-test ($p = .095$). An independent-samples t-test on pre-test scores showed no significant difference between the experimental and control groups ($t(170) = 0.268$, $p = .789$, mean difference = 0.465), indicating that the two groups were equivalent at baseline before treatment. The independent-samples t-test on post-test scores, conducted using SPSS version 22, is presented in **Table 4**.

As shown in **Table 4**, the comparison of post-test scores under the equal-variances assumption revealed a significant difference between the experimental and control groups, $t(170) = 5.17$, $p < .001$, mean difference = 8.14 (standard error = 1.57). Because $p < .05$, H_0 was rejected, indicating a significant difference between the two groups. Based on the reported t value and group sizes ($n_1 = 86$, $n_2 = 86$), the effect size was Cohen's $d = 0.79$, indicating a large effect. This result shows that the

Table 4. Independent samples t-test for post-test HOTS scores

	F (Levene's)	Significance	t	df	Significance (2-tailed)	Mean difference
Equal variances assumed	0.385	0.536	5.172	170	< .001	8.140
Equal variances not assumed	-	-	5.172	168.941	< .001	8.140

Note. Levene's test was non-significant ($p = .536$), so the equal-variances-assumed row was used for hypothesis testing; Cohen's $d = 0.79$, computed from group means and pooled SD

Table 5. HOTS scores by cognitive dimension, pre- and post-test, by group

Dimension	Experimental pre-test (%)	Experimental post-test (%)	Control pre-test (%)	Control Post-test (%)
Analyze	81.56	91.69	77.41	86.54
Evaluate	59.47	76.41	62.13	66.45
Create	50.19	62.40	50.39	52.91

Note. Values are percentages of the maximum possible score on each dimension's items

group that received the DL-integrated RADEC model had significantly higher HOTS scores than the control group in science learning on environmental problems.

HOTS by Cognitive Dimension

To provide a more fine-grained account of the intervention's effect, HOTS scores were disaggregated by cognitive dimension—analyze, evaluate, and create—for each group at pre- and post-test (Table 5).

All three dimensions improved after instruction in both groups. **On analyze, the control group increased from 77.41% to 86.54% (+9.13 points), while the experimental group increased from 81.56% to 91.69% (+10.13 points). On evaluate, the control group increased from 62.13% to 66.45% (+4.32 points), while the experimental group showed a substantially larger increase, from 59.47% to 76.41% (+16.94 points). On create, the control group increased by 2.52 points (50.39% to 52.91%), while the experimental group increased by 12.21 points (50.19% to 62.40%).** Overall, the experimental group showed larger gains on evaluate and create and reached a higher post-test level on every HOTS dimension measured, indicating that the DL-integrated RADEC model was more effective than the control condition in raising higher-order thinking, with its comparative advantage concentrated in the evaluate and create dimensions.

Inferential Comparison of HOTS Sub-Dimensions (RQ1a-RQ1c)

To test RQ1a-RQ1c—whether the intervention's effect differed significantly by cognitive dimension—post-test analyze, evaluate, and create scores (expressed as a percentage of each dimension's maximum, to place the unequal item counts, 7, 7, and 6 items, respectively, on a common 0-100 scale) were analyzed with a one-way multivariate analysis of covariance (MANCOVA), with group (experimental vs. control) as the fixed factor and the corresponding pre-test analyze, evaluate, and create scores entered as covariates. Shapiro-Wilk tests indicated the raw post-test sub-dimension scores departed from normality in both groups (all $p < .01$), so a non-parametric Mann-Whitney U test on the

unadjusted post-test scores was also run as a robustness check for each dimension. The MANCOVA assumption of homogeneous within-group covariance matrices was supported by a non-significant Box's M test ($M = 5.44$, $\chi^2(6) = 5.33$, $p = .50$), and Levene's tests indicated homogeneous variances for each dimension's post-test scores considered individually (all $p > .30$).

The omnibus MANCOVA showed a statistically significant multivariate effect of group on the three post-test sub-dimension scores after adjusting for pre-test performance, Wilks' $\lambda = .824$, $F(3, 165) = 11.77$, $p < .001$, partial $\eta^2 = .18$. Follow-up univariate ANCOVAs, evaluated against a Bonferroni-adjusted α of $.05/3 = .0167$ to control the familywise error rate across the three dimensions, indicated that the experimental group scored significantly higher than the control group on all three dimensions after controlling for pre-test scores: analyze, $F(1, 169) = 6.82$, $p = .010$, partial $\eta^2 = .039$ (adjusted M: 91.59 experimental vs. 86.65 control); evaluate, $F(1, 169) = 12.19$, $p < .001$, partial $\eta^2 = .067$ (adjusted M: 76.68 vs. 66.17); and create, $F(1, 169) = 7.57$, $p = .007$, partial $\eta^2 = .043$ (adjusted M: 62.42 vs. 52.89) (Table 6). The Mann-Whitney U robustness check corroborated all three effects (analyze: $U = 2647.0$, $p < .001$; evaluate: $U = 2721.0$, $p = .002$; create: $U = 2886.5$, $p = .011$). Taken together, these results answer RQ1a-RQ1c affirmatively: the DL-integrated RADEC model produced a statistically significant advantage over the control condition on every HOTS sub-dimension measured, not only on the composite score, with the largest adjusted effect on evaluate (partial $\eta^2 = .067$), followed by create (.043) and then analyze (.039).

Table 7 shows the illustrative HOTS test items by cognitive dimension.

DISCUSSION

Deep Learning-Integrated RADEC Enhances HOTS

RADEC provides a systematic activity structure through the read, answer, discuss, explain, and create stages, while DL strengthens the quality of the learning process at each stage through three principles—mindful, meaningful, and joyful. In other words, RADEC keeps

Table 6. Follow-up ANCOVA for post-test HOTS sub-dimension scores, controlling pre-test scores

Dimension	F(1, 169)	p	Partial η^2	Adjusted mean experimental	Adjusted mean control	Mann-Whitney p (robustness)
Analyze (C4)	6.824	.010	.039	91.59	86.65	< .001
Evaluate (C5)	12.186	< .001	.067	76.68	66.17	.002
Create (C6)	7.566	.007	.043	62.42	52.89	.011

Note. N = 172 (86 experimental, 86 control); A one-way MANCOVA on the three posttest sub-dimension scores, with the corresponding pre-test sub-dimension scores as covariates, indicated a significant multivariate effect of group, Wilks' lambda = .824, $F(3, 165) = 11.77$, $p < .001$, partial $\eta^2 = .18$; Box's M = 5.44, $\chi^2(6) = 5.33$, $p = .50$, indicating the homogeneity-of-covariance-matrices assumption was met; The univariate ANCOVA results above are follow-up tests to this omnibus effect, evaluated against a Bonferroni-adjusted $\alpha = .05/3 = .0167$; all three remained significant at this stricter threshold; Adjusted means are estimated at the grand mean of each dimension's pretest covariate; Because Shapiro-Wilk tests indicated the raw post-test sub-dimension scores were not normally distributed in either group (all $p < .01$), a non-parametric Mann-Whitney U test on the unadjusted post-test scores was run as a robustness check; Results corroborated the ANCOVA conclusions for all three dimensions (analyze: $U = 2647.0$, $p < .001$; evaluate: $U = 2721.0$, $p = .002$; create: $U = 2886.5$, $p = .011$)

Table 7. Illustrative HOTS test items by cognitive dimension

Dimension	Sample item (translated)
Analyze (C4)	Two residents disagree about single-use plastic. Text A (a shopkeeper): "I prefer plastic bags for customers because they are cheap and strong; waste is the sanitation workers' problem, not mine." Text B (an environmental activist): "Every sheet of plastic we throw away will still be on Earth after we are gone; we must stop using single-use plastic now." From the shopkeeper's own perspective, why does he refuse to stop using plastic? (a) He cares about sanitation workers and wants to give them more work; (b) He believes his store's costs matter more than future environmental damage; (c) He does not know plastic is hard to decompose; (d) He wants to make it easier for customers to carry groceries.
Evaluate (C5)	A housing developer clears forest that was home to birds and squirrels, then installs iron cages on utility poles so birds stay dry, and bright garden lights so squirrels can forage safely at night. Evaluating this action against the principle of ecosystem balance, which judgment is most appropriate? (a) Appropriate, because bright lighting protects the animals from predators at night; (b) Not appropriate, because artificial cages and lighting cannot replace the food and shelter that natural trees provided; (c) Appropriate, because iron cages are safer than natural nests that can be damaged by strong wind; (d) Not appropriate, because the developer should have built a swimming pool for the animals as well.
Create (C6)	Plastic drink bottles pile up at school after break time and are rarely reused, leaving the schoolyard untidy. What is the most creative and functional sustainable solution students could implement? (a) Put up posters in every classroom urging students not to buy bottled drinks; (b) Provide many extra plastic-only bins with no further processing plan; (c) Design a 'hanging garden' program that turns used bottles into hanging plant pots to beautify the classroom corridor; (d) Collect all plastic bottles and burn them behind the school so the waste disappears quickly.

Note. Items are translated and condensed from the original 20-item instrument for illustration; The complete instrument is available from the corresponding author upon request

students actively engaged across a sequence of stages, while DL ensures that each stage unfolds more meaningfully, consciously, and enjoyably.

In the read stage, pre-learning questions gave students a focus before reading, so that reading became purposeful and meaningful rather than passive. This stage applies the mindful principle through the experience of understanding, as students become aware of themselves as active learners regulating their own learning process (Kemendikdasmen, 2025), while simultaneously building an initial knowledge schema consistent with Piagetian cognitive theory (Santrock, 2011). This is consistent with Thahir et al. (2021), who found a positive relationship between science literacy and higher-order thinking among biology students.

In the answer stage, students independently answered questions based on their reading, encouraging active knowledge construction before entering group discussion. This stage reflects the mindful principle, as students actively connect prior knowledge with new

information—a process consistent with the concept of assimilation in Piaget's theory (Santrock, 2011). The answer activity also gives an early picture of student understanding and serves as a bridge into the collaborative learning of the next stage (Rahayu et al., 2021).

In the discuss stage, students discuss in groups to apply the understanding they have built, exchange arguments, and weigh different viewpoints. This stage applies the mindful principle through the experience of applying, as students translate knowledge into the context of a problem through collaborative activity (Mustaghfirin & Zaman, 2025), reflecting a constructivist view of students as active agents building knowledge through social interaction. Discussion activity gives students a chance to exchange perspectives, question ideas, and construct arguments. This is consistent with Fandos-Herrera et al. (2023), who found that student discussion activity can drive critical thinking and debate through peer interaction, and with Virliana and Fauziah

(2025), who found collaborative learning—including group discussion—effective for training cooperation, idea exchange, and critical thinking. These findings reinforce the importance of the space for interaction and discussion that RADEC provides.

In the explain stage, each group presents its discussion results while other groups respond. This stage applies the meaningful and mindful principles through the experience of applying, as students connect the concepts studied with their own group's reasoning (Kemendikdasmen, 2025). This activity aligns with Vygotsky's (1978) social constructivism, which holds that knowledge develops through social interaction before being internalized as individual understanding. Agustin et al. (2021) showed that this stage not only trains preservice elementary teachers' verbal communication but also supports HOTS development.

In the create stage, students work together to produce a poster based on the material studied. This stage represents the joyful principle together with mindfulness through the experience of reflecting (Kemendikdasmen, 2025). Students do not merely recall material but use and reflect on their understanding to produce work relevant to the problem discussed, consistent with Piaget's (1970) view that knowledge is actively constructed rather than passively received. Agustin et al. (2021) affirm that this stage pushes preservice elementary teachers to generate ideas and solutions grounded in the understanding they have built, an experience aligned with the create (C6) cognitive process; the ANCOVA reported above confirms that this dimension showed a statistically significant, if comparatively modest, advantage for the experimental group.

Why Deep Learning Strengthens RADEC

RADEC drives active student engagement through a systematic sequence, from understanding concepts, through discussion and problem-solving, to collaboratively developing ideas. This structure scaffolds students' thinking process, while DL strengthens the quality of that engagement through the mindful, meaningful, and joyful principles and the experiences of understanding, applying, and reflecting (Kemendikdasmen, 2025). This is consistent with Gunawardena and Wilson (2021), who argue that higher-order thinking must be explicitly scaffolded through instructional support rather than assumed to emerge spontaneously. It is also consistent with evidence from outside the RADEC literature that explicit, structured scaffolding measurably raises HOTS beyond what conventional instruction achieves: Alanazi et al. (2024) found that a scaffolding-and-guided-discovery module produced significantly larger HOTS gains than conventional teaching among Saudi technical-college physics students (ANCOVA partial η^2

= .585), and Arifin et al.'s (2025) meta-analysis of 36 inquiry-based-learning studies found a large pooled effect on critical thinking (Hedges' $g = 1.27$). The effect observed here ($d = 0.79$, a large effect) falls within that broader distribution, below Arifin et al.'s (2025) pooled $g = 1.27$ but comparable in order of magnitude to Alanazi et al.'s (2024) ANCOVA effect, notwithstanding that the present intervention comprised only four treatment sessions (eight instructional hours) over eight weeks; the meta-analytic evidence indicates that intervention duration is itself a significant moderator of active-learning effects on higher-order thinking (Arifin et al., 2025), so a longer implementation might plausibly yield still larger gains.

Findings from Magfirah et al. (2024) show that RADEC is a student-centered approach built on understanding concepts, collaborating, solving problems, and generating creative ideas. Iwanda et al. (2022) likewise report that RADEC promotes active student participation, while Kara and Aslan (2024) found that actively engaging activities can raise preservice teachers' scientific reasoning. Separately, Mutmainnah et al. (2025) found that DL can build student confidence through group discussion, experimentation, and project work—an affective dimension that this study did not measure directly but that plausibly reinforces the cognitive gains observed here, especially given that preservice teachers who feel more confident engaging in discuss and explain may participate more fully in exactly the stages linked to the largest HOTS gains.

Theoretically, this integration is consistent with constructivism, in which learners actively build their own knowledge (Piaget, 1970), and with social constructivism, which emphasizes the role of social interaction in shaping understanding (Vygotsky, 1978). Both foundations are reflected in RADEC's discuss and explain activities, which DL then deepens further. RADEC thus supplies the instructional structure and scaffolding, while DL strengthens the quality of students' learning experience at every stage.

From Collaborative Learning to Higher-Order Thinking

The discuss and explain stages play a central role in training reasoning, comparison, and justification. In discuss, students exchange views, present arguments, and weigh perspectives in settling on an answer, supporting the comparison and evaluation of information. This is consistent with Appleby et al. (2025), who found that giving students room to voice disciplinary reasoning and integrate that reasoning into shared learning supports reasoning practices in science instruction. Virliana and Fauziah (2025) likewise affirm that collaborative learning through group discussion can train cooperation, idea exchange, and critical thinking. In

explain, students construct explanations in a coherent sequence, defend their arguments, and respond to questions—activities that support justification and the evaluation of information's accuracy (Agustin et al., 2021). The finding in **Table 5** that the largest gain occurred on evaluate (+16.94 points) is consistent with the role these argumentative activities play in supporting evaluative capacity.

The create stage subsequently demands more complex cognitive processes—transfer, synthesis, and generation. Students transform understanding from earlier stages into a new product, a poster, reflecting a process of synthesizing information and applying knowledge in the form of a concrete solution (Agustin et al., 2021). Conceptually, this sequence of activities reveals a shift in cognitive demand across RADEC: discuss and explain emphasize reasoning and justification through argumentative interaction, while create demands synthesizing understanding into a new product. Collaborative learning within RADEC therefore provides a sequence of activities that can move students from reasoning about and evaluating arguments toward the generative act of creation as a HOTS dimension in its own right.

Differential Gains Across HOTS Dimensions

A pattern worth examining directly is that, although the experimental group's advantage over the control group was statistically significant on all three HOTS dimensions once pre-test scores were controlled (**Table 6**), the size of that advantage was not uniform: partial η^2 was smallest for analyze (.039), intermediate for create (.043), and largest for evaluate (.067). One plausible explanation is a ceiling effect: both groups entered the study with comparatively high analyze scores (81.56% and 77.41% at pre-test, respectively), leaving less room for a differential gain regardless of instructional approach, whereas evaluate and create began from a markedly lower baseline (roughly 50-59% in both groups) and therefore had more room to move. A second, complementary explanation concerns which RADEC stages place the heaviest cognitive demand on which dimension. Analyze-type reasoning is exercised early, during read and answer, where students work largely independently before receiving peer or instructor input; evaluate-type reasoning is exercised more intensively during discuss and explain, where claims are exchanged, questioned, and defended in front of peers; and create-type reasoning is concentrated almost entirely in the single create stage. If DL's meaningful-mindful-joyful principles operate primarily by making social, argumentative exchange more productive—a possibility with which the pilot-stage discuss and explain fidelity ratings in **Table 2** (66.25% and 76.25%, both categorized “good”) are compatible, although, because they come from a 20-student pilot cohort rather than the main study, they cannot confirm it—then the

dimension most dependent on that exchange (evaluate) would be expected to show the largest adjusted effect, which is the pattern observed, although this remains a tentative interpretation. Create's smaller, though still significant, effect is consistent with it being supported by only one RADEC stage rather than two, and with create being the most cognitively demanding of the three levels in Bloom's revised taxonomy (Anderson & Krathwohl, 2001); that this dimension nonetheless reached significance, even at a Bonferroni-corrected threshold, suggests the intervention's benefit for generative, creation-level thinking is real but comparatively modest—a pattern broadly consistent with evidence elsewhere that generative, creation-level outcomes are typically the hardest to move in active-learning interventions of limited duration (Arifin et al., 2025).

Preservice Teachers as Both Learners and Future Facilitators

The findings carry implications for preservice elementary teachers' development, since they are not only learners in this study but also future teachers who will one day need to facilitate their own pupils' HOTS. Integrating RADEC with the meaningful, mindful, and joyful principles offers instructors a reference model for designing systematic, student-centered science instruction, including through the development of teaching modules and student worksheets that guide learners to understand, apply, and reflect. The HOTS instrument developed for this study can likewise serve as a reference for constructing assessments that measure higher-order thinking. In this sense, applying DL-integrated RADEC is relevant not only for developing preservice teachers' own HOTS during coursework but also for giving them a learning experience consistent with their future role as elementary science teachers—a dual function also highlighted in preservice-teacher-focused inquiry training research, where affective outcomes such as self-efficacy and emotion toward teaching were found to shift alongside cognitive or pedagogical-design outcomes over the course of a training program (García-Ruiz et al., 2022). Future work that measures such affective correlates alongside HOTS would help clarify whether the confidence-building mechanism suggested by Mutmainnah et al. (2025) is in fact operating here.

Comparison with Prior RADEC and Deep Learning Research

The large effect observed in this study ($d = 0.79$) is broadly consistent with, though somewhat more modest than, effects reported elsewhere in the RADEC and DL literature. Agustin et al. (2021) and Rahayu et al. (2021) both reported that RADEC positively affects preservice elementary teachers' HOTS and critical thinking, without disaggregating by cognitive dimension, so a direct effect-size comparison is not possible; the present

study extends that line of work by showing where within HOTS the RADEC advantage concentrates. Hayati et al. (2023) demonstrated RADEC's adaptability to tertiary science content in an organic chemistry course, supporting the plausibility of applying RADEC to third-semester preservice elementary teachers' environmental-science content, as done here. On the DL side, Lailiyati and Afandi (2026) found that RADEC integrated with DL produced active, contextual, meaningful learning that supported elementary pupils' concept understanding and motivation; because that study targeted school-age pupils on a mapping topic rather than preservice teachers on HOTS, it is complementary rather than directly comparable, but its finding that the integration supports motivation is consistent with the confidence-building mechanism proposed above. Set against the wider active-learning literature—where inquiry-based approaches show a pooled effect on critical thinking of $g = 1.27$ across 36 studies with substantial heterogeneity by duration and level (Arifin et al., 2025), and a structured digital STEM project produced meaningful mathematical critical-thinking gains through explicit scaffolding (Pramasdyahsari et al., 2023)—the present study's more modest effect size is consistent with its shorter, single-unit implementation rather than indicating that the RADEC-DL integration is a weaker instructional approach in principle.

Boundary Conditions: What Active Learning Does and Does Not Move

The present study measured only cognitive HOTS outcomes and did not assess motivation, self-efficacy, or attitude toward science, a limitation worth flagging explicitly given evidence that active-learning interventions do not always move cognitive and affective outcomes together. Argaw et al. (2017) found that problem-based learning produced a significant, above-average effect on physics students' problem-solving skills but no significant change in their motivation to learn physics, indicating that a cognitive HOTS gain of the kind reported here cannot be assumed to imply a parallel affective shift. This matters for interpreting the present findings: it is plausible that the DL-integrated RADEC model raised HOTS through the cognitive scaffolding that discuss and explain provide, independent of any change in participants' confidence or enjoyment, or it is equally plausible that a genuine affective shift (as Mutmainnah et al., 2025, suggest can occur with DL) partly explains the cognitive gain by increasing engagement in those same stages. The present data cannot distinguish between these possibilities, and future studies that pair a HOTS test with a validated motivation or self-efficacy measure would help clarify the mechanism.

Implications for Elementary Science Teacher Education

These findings carry implications for designing instruction for preservice elementary teachers, who are not only learners in this study but also future teachers responsible for facilitating their pupils' HOTS. Integrating RADEC with the meaningful, mindful, and joyful principles can serve as a reference for lecturers designing systematic, student-centered science instruction, including through developing teaching modules and student worksheets that guide learners to understand, apply, and reflect. The HOTS instrument developed here can also serve as a reference for building assessments that measure higher-order thinking. Applying DL-integrated RADEC is therefore relevant not only for developing preservice teachers' own HOTS during their coursework, but also for giving them a learning experience consistent with their future role as elementary science teachers.

CONCLUSIONS

Implementing the RADEC model integrated with a DL approach produced a significant difference in the HOTS of third-semester preservice elementary teachers compared with Direct Instruction using a scientific approach. An independent-samples t-test showed that the experimental group's post-test score was significantly higher than the control group's, $t(170) = 5.17$, $p < .001$, $d = 0.79$. Descriptively, the experimental group also scored higher at post-test on every HOTS dimension measured—analyze, evaluate, and create—with the largest descriptive gain on evaluate.

These findings indicate that integrating RADEC with the mindful, meaningful, and joyful learning principles can support HOTS development through a sequence of reading, answering, discussing, explaining, and generating ideas activities. A follow-up MANCOVA confirmed that this benefit extended to each individual HOTS dimension: after controlling pre-test scores, the experimental group scored significantly higher than the control group on analyze, evaluate, and create considered separately (all $p < .05$, Bonferroni-corrected), with the largest adjusted effect on evaluate. To our knowledge, this study is among the first to test a DL-integrated RADEC model specifically on preservice elementary teachers' HOTS in science learning, disaggregated and inferentially tested by cognitive dimension, offering methods-course instructors a tentative, stage-linked hypothesis to test rather than a general recommendation to "adopt more active learning."

Limitations and Directions for Future Research

This study has several limitations. First, it involved only third-semester students from a single PGSD program at one university in one city, so generalization

of the findings to other institutions or regions should proceed cautiously. Second, the study focused on only one science unit on environmental problems and was implemented over a relatively short intervention period. In addition, the dichotomously scored test used here may not fully capture the complexity of students' reasoning processes, including partial reasoning that emerges during problem-solving. A further limitation concerns implementation fidelity: it was observed only during the pilot study, on 20 students and by a single observer, and was not systematically observed in the main study. The ratings in **Table 2** therefore cannot be taken to represent the instructional experience of the 86 students in the main-study experimental group, and inter-rater reliability could not be established. Future research should observe fidelity systematically across the full experimental group, using a complete or randomly drawn set of sessions and at least two independent observers, and should report an inter-rater agreement statistic (e.g., Cohen's kappa). In addition, although students were randomly assigned to condition, each condition was taught in a single separate class, so class- and instructor-level effects cannot be fully separated from the treatment effect; although the lecturers were comparable in experience and qualifications and had aligned their instructional perceptions beforehand, replications should include multiple classes or instructors per condition. The use of multiple-choice items may also not fully capture the complexity of the create (C6) skill. Future research could therefore involve larger and more diverse samples, apply the intervention over a longer period and across a range of science topics, and complement the HOTS test with open-ended performance tasks and product-evaluation rubrics to capture students' reasoning processes qualitatively.

Author contributions: **IRWA:** conceptualization, data curation, methodology, resources, supervision, writing – original draft; **AAR:** investigation, visualization; **M:** project administration; **RA:** formal analysis; **DYS:** validation; **FPA:** writing – review & editing. All authors agreed with the results and conclusions.

Funding: This research received no external funding.

Ethical statement: This study was approved by the Research Ethics Committee at Universitas Sebelas Maret with approval number 3190/UN27.02.7.13/EC/2026. The ethics approval letter was issued on 17 April 2026, after data collection had been completed, owing to a prolonged administrative process; the research protocol had been submitted to the committee on 29 August 2025, before the main-study data collection began. The pilot study (June–August 2025) was conducted before this submission. Written informed consent was obtained from all participants before data collection in both the pilot and main studies. Participation was voluntary. Participants' names were replaced with code numbers in the analysis dataset, and the data are stored securely and accessible only to the research team.

AI statement: The authors used Claude (Anthropic) for language editing, translation, and drafting and revising parts of the manuscript and responses to reviewer comments. The authors reviewed and edited all AI-assisted content and took full responsibility for the article.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

- Agustin, M., Pratama, Y. A., Sopandi, W., & Rosidah, I. (2021). Pengaruh model pembelajaran RADEC terhadap keterampilan berpikir tingkat tinggi mahasiswa PGSD [The effect of the RADEC learning model on preservice elementary teachers' higher-order thinking skills]. *Jurnal Cakrawala Pendas*, 7(1), 110-117. <https://doi.org/10.31949/jcp.v7i1.2672>
- Alanazi, A. A., Osman, K., & Halim, L. (2024). Effect of scaffolding strategies and guided discovery on higher-order thinking skills in physics education. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(9), Article em2496. <https://doi.org/10.29333/ejmste/14980>
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Andini, S. R., & Fitria, Y. (2021). Pengaruh model RADEC pada pembelajaran tematik terhadap hasil belajar peserta didik sekolah dasar [The effect of the RADEC model in thematic learning on elementary students' learning outcomes]. *Jurnal Basicedu*, 5(3), 1435-1443. <https://doi.org/10.31004/basicedu.v5i3.960>
- Anggraeni, P., Sopandi, W., Septinaningrum, S., Hayati, A., Tursinawati, T., & Gumala, Y. (2021). Keterampilan berpikir kritis mahasiswa PGSD melalui pembelajaran read-answer-discuss-explain-and-create (RADEC) yang berorientasi penyelidikan [Preservice elementary teachers' critical-thinking skills through inquiry-oriented RADEC learning]. *Caruban: Jurnal Ilmiah Ilmu Pendidikan Dasar*, 4(1), 10-19. <https://doi.org/10.33603/cjiipd.v4i1.4398>
- Appleby, L., Caspari-Gnann, I., Gouvea, J., Hammer, D., & Tobin, R. (2025). Instructor listening during lecture: A case from introductory chemistry. *Journal of College Science Teaching*, 54(1), 8-15. <https://doi.org/10.1080/0047231X.2024.2398384>
- Argaw, A. S., Haile, B. B., Ayalew, B. T., & Kuma, S. G. (2017). The effect of problem based learning (PBL) instruction on students' motivation and problem solving skills of physics. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(3), 857-871. <https://doi.org/10.12973/eurasia.2017.00647a>

- Arifin, Z., Sukarmin, Saputro, S., & Kamari, A. (2025). The effect of inquiry-based learning on students' critical thinking skills in science education: A systematic review and meta-analysis. *Eurasia Journal of Mathematics, Science and Technology Education*, 21(3), Article em2592. <https://doi.org/10.29333/ejmste/15988>
- Aydın, N., & Yılmaz, A. (2010). The effect of the constructivist approach on students' higher-order cognitive skills. *Hacettepe University Journal of Education*, 39, 57-68.
- Budiarti, M. I. E., Kahar, M. S., & Jotley, E. M. (2024). Analisis kemampuan berpikir tingkat tinggi mahasiswa dalam menyelesaikan soal pemecahan masalah matematika pada mata kuliah kalkulus peubah banyak [Analysis of students' higher-order thinking skills in solving mathematics problem-solving tasks in multivariable calculus]. *Jurnal Riset dan Inovasi Pembelajaran*, 4(2), 1216-1229. <https://doi.org/10.51574/jrip.v4i2.1599>
- Cholily, Y. M., Eriyanti, R. W., & Masduki, M. (2022). Meningkatkan kreativitas guru dalam inovasi pembelajaran berbasis HOTS untuk mengembangkan berpikir kritis dan kreatif siswa [Enhancing teacher creativity in HOTS-based instructional innovation to develop students' critical and creative thinking]. *To Maega: Jurnal Pengabdian Masyarakat*, 5(3), 416-428. <https://doi.org/10.35914/tomaega.v5i3.1176>
- Dahlia, D., & Tirtoni, F. (2025). Penerapan model RADEC untuk meningkatkan literasi budaya dan kewarganegaraan siswa melalui pembelajaran pendidikan Pancasila di sekolah dasar [Applying the RADEC model to improve elementary students' cultural and civic literacy in Pancasila education]. *Edukasiana: Jurnal Inovasi Pendidikan*, 4(3), 784-792. <https://doi.org/10.56916/ejip.v4i3.1192>
- Din, C., & MacInnis, M. (2024). Reworking the recipe: Adding inquiry and reflection to college science labs. *Journal of College Science Teaching*, 53(5), 486-490. <https://doi.org/10.1080/0047231X.2024.2381418>
- Diputera, A. M., Zulpan, & Eza, G. N. (2024). Memahami konsep pendekatan deep learning dalam pembelajaran anak usia dini yang meaningful, mindful, dan joyful: Kajian melalui filsafat pendidikan [Understanding the deep learning approach in early-childhood learning through meaningful, mindful, and joyful principles: A philosophy-of-education review]. *Jurnal Bunga Rampai Usia Emas*, 10(2), 108-120. <https://doi.org/10.24114/jbrue.v10i2.65978>
- Fandos-Herrera, C., Herrando, C., Jiménez Martínez, J., & Pina, J. M. (2023). Discuss to learn: A student-focused learning strategy. *Education + Training*, 65(5), 697-710. <https://doi.org/10.1108/ET-02-2022-0049>
- Fauziana, Kastri Fani, & Rahmiaty. (2021). Analisis kemampuan siswa dalam menyelesaikan soal HOTS pada pelajaran IPA [Analysis of students' ability to solve HOTS questions in science]. *Genderang Asa: Journal of Primary Education*, 2(2), 66-75. <https://doi.org/10.47766/ga.v2i2.165>
- García-Ruiz, C., Lupión-Cobos, T., & Blanco-López, Á. (2022). Effects of an inquiry-based science education training program on pre-service teachers: A mixed-methods case study. *Eurasia Journal of Mathematics, Science and Technology Education*, 18(12), Article em2186. <https://doi.org/10.29333/ejmste/12578>
- Gunawardena, M., & Wilson, K. (2021). Scaffolding students' critical thinking: A process not an end game. *Thinking Skills and Creativity*, 41, Article 100848. <https://doi.org/10.1016/j.tsc.2021.100848>
- Hayati, N., Kadarohman, A., Sopandi, W., Martoprawiro, M. A., & Pratiwi, A. (2023). The effect of the RADEC model on conceptual understanding of polycyclic aromatic hydrocarbons (PAHs) topic. *Pertanika Journal of Social Sciences and Humanities*, 31(4), 1649-1667. <https://doi.org/10.47836/pjssh.31.4.15>
- Iwanda, N. C., Malika, H. N., & Aqshadigrama, M. (2022). RADEC sebagai inovasi model pembelajaran pendidikan agama Islam pasca pandemi COVID-19 di sekolah dasar [RADEC as an instructional innovation for Islamic religious education in elementary schools after the COVID-19 pandemic]. *Jurnal Ilmiah Wahana Pendidikan*, 8(24), 430-440. <https://doi.org/10.5281/zenodo.7494585>
- Jin, G., Wodika, A., Darner, R., & Lai, J. (2023). Fostering scientific literacy with problem sets that generate cognitive presence and fulfill basic psychological needs. *Journal of College Science Teaching*, 52(7), 85-95. <https://doi.org/10.1080/0047231X.2023.12315883>
- Kara, S., & Aslan, O. (2024). Evaluation of the effect of in-class activity-based practices on scientific reasoning skills of pre-service science teachers. *International Journal of Science Education*, 46(8), 773-794. <https://doi.org/10.1080/09500693.2023.2258253>

- Kemendikdasmen. (2025). Naskah akademik pembelajaran mendalam menuju pendidikan bermutu untuk semua [Academic paper on deep learning toward quality education for all]. *Pusat Kurikulum dan Pembelajaran*.
- Kurniawan, A., Juliangkary, E., Pratama, M. Y., & Saputra, I. (2019). Analisis kesulitan siswa dalam menyelesaikan soal fungsi [Analysis of student difficulties in solving function problems]. *Media Pendidikan Matematika*, 7(1), 72-82. <https://doi.org/10.33394/mpm.v7i1.1679>
- Lailiyati, & Afandi, M. (2026). Penerapan model pembelajaran RADEC terintegrasi deep learning materi peta kelas V terhadap pemahaman konsep ditinjau dari motivasi belajar [Applying a deep learning-integrated RADEC model on grade-V map material to concept understanding, viewed from learning motivation]. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 11(1), 78-91.
- Magfirah, Imran, M. E., & Amal, A. (2024). Pengaruh model RADEC (read, answer, discuss, explain, and create) terhadap kemampuan berpikir tingkat tinggi [The effect of the RADEC model on higher-order thinking ability]. *Jurnal Riset Guru Indonesia*, 3(3), 139-148. <https://doi.org/10.62388/jrgi.v3i3.454>
- Manalu, A., Simamora, A., Purba, I. D., & Sinaga, N. (2026). The effectiveness of problem-based learning model on understanding concepts and science process skills as implementation of the independent curriculum. *Jurnal Ilmu Sosial dan Humaniora*, 4(2), 633-641. <https://doi.org/10.58540/isihumor.v4i2.1679>
- Mualimin, M., & Pamungkas, R. (2024). Holistic higher-order thinking in solving environmental problems: A profile of Indonesian undergraduate students. *Jurnal Pendidikan Biologi Indonesia*, 10(3), 804-812. <https://doi.org/10.22219/jpbi.v10i3.36426>
- Mustaghfirin, U. A., & Zaman, B. (2025). Tinjauan pendekatan pembelajaran mendalam Kemendikdasmen perspektif pendidikan Islam [A review of the Ministry's deep-learning approach from an Islamic education perspective]. *Journal of Instructional and Development Researches*, 5(1), 75-85. <https://doi.org/10.53621/jider.v5i1.476>
- Mutmainnah, N., Adrias, & Zulkarnaini, A. P. (2025). Implementasi pendekatan deep learning terhadap pembelajaran matematika di sekolah dasar [Implementing the deep learning approach in elementary mathematics learning]. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 10(1), 858-871.
- Noor, P. P., & Abadi, A. P. (2022). Kemampuan berpikir tingkat tinggi dalam perkembangan pembelajaran matematika SMA [Higher-order thinking ability in the development of senior-secondary mathematics learning]. *Jurnal Educatio FKIP UNMA*, 8(2), 466-473. <https://doi.org/10.31949/educatio.v8i2.1986>
- Nurdiana, H., & Pradana, S. P. (2025). Efektivitas penerapan pendekatan deep learning untuk meningkatkan hasil belajar kognitif peserta didik pada pembelajaran IPA kelas VIII SMP Pandanaran Plupuh [Effectiveness of the deep learning approach in improving grade-VIII students' cognitive science learning outcomes]. In *Proceedings of the National Seminar on Science Education* (pp. 46-53).
- Patimah, L., & Widianjani, W. (2022). Pengaruh media pembelajaran dengan menggunakan aplikasi Quizizz terhadap hasil belajar mahasiswa pada mata kuliah matematika diskrit [The effect of Quizizz-based instructional media on student learning outcomes in discrete mathematics]. *Symmetry: Pasundan Journal of Research in Mathematics Learning and Education*, 7(2), 259-264. <https://doi.org/10.23969/symmetry.v7i2.6629>
- Piaget, J. (1970). *Genetic epistemology*. Columbia University Press. <https://doi.org/10.7312/piag91272>
- Pramasdyahsari, A. S., Setyawati, R. D., Aini, S. N., Nusuki, U., Arum, J. P., Astutik, I. D., Widodo, W., Zuliah, N., & Salmah, U. (2023). Fostering students' mathematical critical thinking skills on number patterns through digital book STEM PjBL. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), Article em2297. <https://doi.org/10.29333/ejmste/13342>
- Rahayu, A. H., Sopandi, W., Anggraeni, P., Tursinawati, T., & Septinaningrum, S. (2021). Keterampilan berpikir kritis mahasiswa PGSD melalui pembelajaran read-answer-discuss-explain-and-create (RADEC) yang berorientasi penyelidikan [Preservice elementary teachers' critical-thinking skills through inquiry-oriented RADEC learning]. *Caruban: Jurnal Ilmiah Ilmu Pendidikan Dasar*, 4(1), 10-19. <https://doi.org/10.33603/cjiipd.v4i1.4398>
- Renaud, R. (2002). *The effect of higher order questions on critical thinking skills* [PhD thesis, The University of Western Ontario].
- Resendes, K. (2022). A model for a data analysis- and literature-intensive undergraduate course. *Journal of College Science Teaching*, 51(5), 42-54. <https://doi.org/10.1080/0047231X.2022.12290580>
- Santrock, J. W. (2011). *Educational psychology* (5th ed.). McGraw-Hill.

- Setyawarno, D., & Kurniawati, A. (2022). Science learning oriented to higher order thinking in digital era. *AIP Conference Proceedings*, 2600(1), Article 070003. <https://doi.org/10.1063/5.0112450>
- Sopandi, W., Atep, S., Rendi, R. S., Sutinah, C., Yanuar, Y., & Imran, M. E. (2023). *Model pembelajaran RADEC: Teori dan implementasi di sekolah* [The RADEC learning model: Theory and implementation in schools]. UPI Press.
- Suar Adnyana, I. K. (2024). Implementasi pendekatan deep learning dalam pembelajaran bahasa Indonesia [Implementation of the deep learning approach in Indonesian language learning]. *Jurnal Retorika*, 5(2), 1-14.
- Suryana, N., Sopandi, W., Budimansyah, D., & Ruyadi, Y. (2025). Character transformation through the implementation of the RADEC learning model at the elementary school level. *Scaffolding: Jurnal Pendidikan Islam dan Multikulturalisme*, 7(1), 875-888. <https://doi.org/10.37680/scaffolding.v7i1.7268>
- Tahim Bael, B., Nachiappan, S., & Pungut, M. (2021). Analisis kesediaan guru dalam pelaksanaan kemahiran berfikir aras tinggi dalam pembelajaran, pengajaran dan pemudahcaraan abad ke-21 [Analysis of teacher readiness in implementing higher-order thinking skills in 21st-century teaching and learning]. *Muallim Journal of Social Science and Humanities*, 5(1), 100-119. <https://doi.org/10.33306/mjssh/115>
- Thahir, R., Magfirah, N., & Anisa. (2021). Hubungan antara high order thinking skills dan kemampuan literasi sains mahasiswa pendidikan biologi [The relationship between higher-order thinking skills and science-literacy ability among biology-education students]. *Biodik*, 7(3), 105-113. <https://doi.org/10.22437/bio.v7i3.14386>
- Tulljanah, R., & Amini, R. (2021). Model pembelajaran RADEC sebagai alternatif dalam meningkatkan higher order thinking skill pada pembelajaran IPA di sekolah dasar: Systematic review [The RADEC learning model as an alternative for improving higher-order thinking skill in elementary science learning: A systematic review]. *Jurnal Basicedu*, 5(6), 5508-5519. <https://doi.org/10.31004/basicedu.v5i6.1680>
- Velazquez-Ulloa, N., & Finkel, L. (2025). Assessing the impact of an undergraduate biology course for non-majors on students' scientific literacy, perceptions of science, and confidence in their ability to engage in scientific inquiry. *Journal of College Science Teaching*, 54(2), 100-107. <https://doi.org/10.1080/0047231X.2024.2401725>
- Virliana, A. I., & Fauziah, L. S. N. (2025). Pengaruh pembelajaran kolaboratif untuk meningkatkan cara berpikir kritis [The effect of collaborative learning on improving critical thinking]. *Jurnal Jendela Pendidikan*, 5(1), 1-7. <https://doi.org/10.57008/jjp.v5i01.1070>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wardhani, D. A. P., & Oktiningrum, W. (2022). Meningkatkan kemampuan berpikir kritis mahasiswa melalui pengembangan soal matematika dengan konteks COVID-19 [Improving students' critical-thinking ability through COVID-19-contextualized mathematics problems]. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 11(1), 69-79. <https://doi.org/10.24127/ajpm.v11i1.4377>
- Widodo, S., Ladyani, F., Asrianto, L. O., Rusdi, & Lestari, S. M. P. (2023). *Buku ajar metode penelitian* [Research methods textbook]. Science Techno Direct.
- Widyaningrum, P., Hendiyani, H., & Imamah. (2025). Pengembangan model pembelajaran mendalam (deep learning) untuk meningkatkan berpikir kritis pada anak usia 5-6 tahun: Studi deskriptif [Developing a deep-learning model to improve critical thinking in children aged 5-6: A descriptive study]. *Jurnal Anak Bangsa*, 3(1), 160-178. <https://doi.org/10.46306/jas.v1i1>
- Yulianto, D., Junaedi, Y., Juniawan, E. A., & Anwar, S. (2024). Kemampuan berpikir tingkat tinggi siswa SMP melalui pendekatan matematika realistik dengan model PBL dan CTL berbasis project-based learning pada penyelesaian soal AKM di kabupaten lebak banten [Junior-secondary students' higher-order thinking through a realistic mathematics approach combining PBL and CTL project-based learning]. *Teorema: Teori dan Riset Matematika*, 9(1), 57-76. <https://doi.org/10.25157/teorema.v9i1.13457>
- Zainuddin, A. H., Nazimuddin Al Kamil, M., & Japar, R. (2026). Penerapan pembelajaran mendalam (deep learning) di sekolah dasar: Studi literatur [Implementing deep learning in elementary schools: A literature study]. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 11(1), 20-32.
- Zhou, Y., Gan, L., Chen, J., Wijaya, T. T., & Li, Y. (2023). Development and validation of a higher-order thinking skills assessment scale for pre-service teachers. *Thinking Skills and Creativity*, 48, Article 101272. <https://doi.org/10.1016/j.tsc.2023.101272>

- Zoller, U. (2000). Interdisciplinary systemic HOCS development—The key for meaningful STES-oriented chemical education. *Chemistry Education: Research and Practice in Europe*, 1(2), 189-200. <https://doi.org/10.1039/A9RP90021G>
- Zoller, U., Dori, Y. J., & Lubezky, A. (2002). Algorithmic, LOCS and HOCS (chemistry) exam questions: Performance and attitudes of college students. *International Journal of Science Education*, 24(2), 185-203. <https://doi.org/10.1080/09500690110049060>

<https://www.ejmste.com>