Eduproxima: Jurnal Ilmiah Pendidikan IPA
http://jurnal.stkippgritulungagung.ac.id/index.php/eduproxima

PBL-ESD INTERVENTION AND ITS EFFECT ON COGNITIVE OUTCOMES AND CRITICAL THINKING AMONG MALAYSIAN ISLAMIC PRIMARY SCHOOL STUDENTS

Alwi Afriansyah [1)], Hadi Purwanto [2)], Neng Sholihat [*3)], Eka Frima Asda [4)], Nurul Faatihah [5)]
[1,2,3,4)] Natural Science Education Study Program, Faculty of Teacher Training and Education, Universitas Muhammadiyah Riau, Riau, Indonesia.
[5)] Quranic Studies and Multimedia Program, Faculty of Quran and Sunnah Studies, Universiti Sains Islam Malaysia, Negeri Sembilan, Malaysia.
[*]*Corresponding author*
E-mail: afriansyahlwi7@gmail.com [1)], hadipurwanto@umri.ac.id [2)], nengsholihat@umri.ac.id [*3)]
frimaasdaeka@gmail.com [4)], faatihahzahari@gmail.com [5)]

Article history:
Submitted: May 09th, 2026; Revised: June 20th, 2026; Accepted: July 03rd, 2026; Published: July 12th, 2026

ABSTRACT

Critical thinking skills and cognitive learning outcomes in Islamic primary school Science remain low, as reflected in Malaysia's below-average TIMSS 2019 performance. This study examines the effect of the *Problem Based Learning* model integrated with *Education for Sustainable Development* (PBL-ESD) on Grade VI students of Sekolah Rendah Ugama Al-Hilal, Kota Bharu, Kelantan, Malaysia. A *quasi-experimental* method with a *Nonequivalent Pretest-Posttest Control Group Design* was applied (n = 48; experimental = 30, control = 18). Data were analyzed using Shapiro-Wilk, Levene's Test, Mann-Whitney U, and N-Gain. Results revealed significant differences in cognitive learning outcomes (U = 68,000; Z = -4.332; p = .000; N-Gain: experimental = 0.742 [high], control = 0.312 [moderate]) and critical thinking (U = 82.500; Z = -3.999; p = .000; N-Gain: experimental = 0.645 [moderate], control = 0.297 [low]). Beyond categorical classification, the practical magnitude of these gains is substantial: the experimental group outperformed the control group by a mean difference of 35.56 points in cognitive learning outcomes and 31.04 points in critical thinking on the posttest, corresponding to N-Gain differentials of 0.430 and 0.348, respectively—effect sizes that exceed typical benchmarks for large practical significance in educational interventions. These findings indicate that the PBL-ESD model is pedagogically effective in improving both cognitive learning outcomes and critical thinking skills in Islamic primary school Science, with gains large enough to be readily observable in classroom practice rather than merely statistically significant.

Keywords: PBL; ESD; critical thinking; cognitive learning; Islamic school

INTRODUCTION

In the entire educational process, learning activities are the most important activities. This means that the success or failure to achieve educational goals depends a lot on how the learning process is experienced by students as students, as well as in natural science education (Purwanto *et al*., 2020). Natural Sciences (Science) is one of the subjects that plays a strategic role in shaping the scientific and analytical thinking skills of students from the elementary education level. Science Learning is designed to develop curiosity, observation skills, and high-level thinking skills that are directly related to natural phenomena around students' lives(Utami & Desstya, 2021). In addition, effective Science learning can encourage learners to not only understand concepts, but also be able to apply them in real-life contexts, thus making them adaptive and responsive individuals to the challenges of the times (Arifin *et al*., 2024). Therefore, critical thinking skills are an inseparable aspect of

Science learning, because these skills equip students to analyze information, evaluate arguments, and make decisions logically and systematically (Rofi'ah & Rokhmaniyah, 2024).

The ability to think critically is one of the main competencies that must be mastered by students in the 21st century, but conditions in the field show that this ability is still far from optimal, especially at the elementary school level. Data *Trends in International Mathematics and Science Study* (TIMSS) 2019 shows that the achievement of Malaysian students in high-level thinking in Science subjects is below the international average (Wan Nor Sharhida *et al*., 2023;Mutakin *et al*., 2023). This condition reflects systemic weaknesses in the learning approach that is still dominated by conventional teacher-centered methods, where students are more passive and do not get enough space to practice analyzing, evaluating, and synthesizing information independently (Susilawati *et al*., 2020; Hasanah *et al*., 2025). In this case, learners need to be aware of how they understand basic concepts, recognize difficulties that may arise, and identify the most effective problem-solving strategies (Asda, 2024).

A similar problem was found at Al-Hilal Islamic Primary School, Kelantan, where more than 45% of students in grades IV–VI obtained scores below the minimum completeness criteria in Science subjects, while critical thinking skills have not developed adequately (Kelantan State Education Office, 2022). Teachers face obstacles in designing learning that accommodates the diverse needs of students, and are not fully able to direct students towards learning independence. This condition is in line with the findings Wahyudi & Purwanto, (2024) that the low active involvement of students directly contributes to the weakness of high-level thinking skills.

Theoretically, the urgency of developing critical thinking is based on the theory of constructivism that submitted by Vygotsky (1978), which affirms that knowledge is actively constructed by learners through meaningful social interactions within their proximal developmental zones. This theory suggests that it is not enough for students to just receive information, but need to be actively involved in the process of thinking, asking, discussing, and solving real problems. Moving from this theoretical foundation, various innovative learning models have been developed to facilitate critical thinking, including *Problem Based Learning* (PBL), *Project Based Learning* (PjBL), inquiry-based learning, and *Discovery Learning.* Among these, PBL has been most extensively studied and consistently shown to promote critical thinking by its inherent design of placing authentic problems at the center of the learning process (Mayasari *et al*., 2022; Kusnandar & Mirza 2025;Ardianti *et al*., 2021). However, a specific research gap exists: despite growing interest in integrating environmental sustainability values (ESD) into learning, empirical studies examining the combined effect of PBL and ESD in Islamic primary school settings—particularly in the unique socio-cultural context of Kelantan—remain scarce. This gap is significant because the effectiveness of problem-based approaches in Islamic schools is not merely a pedagogical issue but also a matter of aligning instructional design with the spiritual and communal values of the learner community (Mufid *et al*., 2024)

PBL is a learning approach that begins with presenting authentic problems from the real-life context of students, encouraging them to explore, analyze data, and formulate solutions collaboratively (Ardianti *et al*., 2021;Septiana & Kurniawan, 2018). Different from conventional learning that places the teacher as the center of information, PBL places students as independent problem solvers so that they naturally train the critical thinking dimension that is formulated Facione (2015), namely interpretation, analysis, evaluation, explanation, and self-regulation. Through a structured inquiry process, learners gain meaningful learning experiences while training the high-level thinking skills needed to face real-life challenges (Stuart & Scott, 2025). This study makes a distinct empirical contribution to the international literature by providing evidence from an Islamic primary school in Kelantan, Malaysia—a context that has been largely underrepresented in PBL-ESD research. Specifically, this study operationalizes critical thinking using Facione's (2015) multi-dimensional framework (interpretation, analysis, evaluation, explanation, and self-regulation) and examines cognitive learning outcomes across all six revised Bloom's Taxonomy levels (C1–C6), offering a more granular and theoretically grounded assessment than previous studies that have relied on generic achievement tests.

The partial implementation of PBL in Islamic elementary schools faces a number of limitations that need to be considered. First, problem scenarios that are generic or irrelevant to the context of a learner's life often fail to evoke the intrinsic motivation necessary for deep cognitive engagement (Cahyani & Ahmad, 2024). Second, without a context of personally and culturally meaningful values, Muslim learners tend not to see the relevance of the issues presented to their moral and religious responsibilities, so the problem-solving process takes place mechanically without involving the affective dimensions that support critical thinking (Mufid *et al*., 2024). Third, PBL without a sustainability orientation risks only resulting in short-term problem-solving without considering broader environmental and social impacts. Conceptually, the integration of Islamic values and ESD principles is not merely additive but synergistic: the Islamic concept of *khalifatullah fil ardh* (vicegerency of God on earth) establishes environmental stewardship as a religious obligation, directly aligning with ESD's core emphasis on intergenerational responsibility and sustainable action (UNESCO, 2020; Mufid *et al*., 2024). This conceptual convergence means that ESD-contextualized problems carry inherent motivational weight for Muslim learners: solving an environmental problem becomes not merely a cognitive exercise but an expression of religious duty. It is therefore expected that PBL-ESD integration will produce significantly deeper cognitive engagement and more authentic critical thinking compared to either approach implemented in isolation.

The integration of PBL with ESD results in a more comprehensive learning model because it places environmental sustainability issues experienced directly by learners—such as annual floods and forest degradation in Kelantan—as authentic problem scenarios (Wiek *et al*., 2022). UNESCO (2020) defines ESD as education that empowers students to make informed decisions and act responsibly for environmental integrity, economic feasibility, and social justice for current and

future generations. In the context of Islamic schools, this unity is further strengthened by the values of *Khalifatullah fil ardh* which makes environmental concern a religious obligation, not just a civic obligation, so that the affective involvement of students in the critical thinking process increases meaningfully (Mufid *et al*., 2024; Mohd Yusof, 2021). Thus, the PBL-ESD model synergistically develops students' cognitive competencies, sustainability values, and Islamic character at the same time in one integrated learning process.

Various studies have proven the effectiveness of PBL separately, but research integrating PBL with ESD in an integrated manner in Islamic primary schools, particularly in Kelantan, is still very limited (Rahman & Fuad, 2023; Ningtias *et al*., 2024). Based on these gaps, this study formulated two research questions: (1) is there a significant influence of the application of the PBL-ESD model on the cognitive learning outcomes of grade VI students in Science subjects? and (2) is there a significant influence of the application of the PBL-ESD model on the critical thinking skills of grade VI students in Science subjects, compared to the group that received conventional learning? This research aims to answer both questions through rigorous empirical testing, with the hope that the findings obtained can contribute to the achievement of the target *Higher Order Thinking Skills* (HOTS) in the Malaysian Education Development Plan (PPPM) 2013–2025 and provides practical guidance for Science teachers in planning meaningful and contextual learning (Ministry of Education Malaysia, 2023).

METHODS

Approaches and Types of Research

This study uses a quantitative approach with the type of research *quasi-experimental* (pseudo-experiment). This approach was chosen because the purpose of the study was to test the causal influence between the independent variable (PBL-ESD model) and the two bound variables (cognitive learning outcomes and critical thinking ability) through the collection of statistically analyzeable numerical data. The study compared two groups that received different treatments to determine whether there was a significant difference in the bound variables after treatment was given (Creswell & Creswell, 2023).

Research Design

The design used is *the Nonequivalent Pretest-Posttest Control Group Design*, which is a design that involves two groups that each receive pretest measurements before treatment and *posttest* after treatment, without full randomization because the class structure has been formed naturally. The research design is schematically presented in Table 1.

Table 1. Pretest-Posttest Control Group Design (Mutiara *et al*., 2019)

Classes		Pretest	Treatment	Posttest
Experiments	R	O_1	X_1	O_2
Controls	R	O_1	X_2	O_2

Description :

R : Cluster random assignment (class draw)

O_1: Pretest
O_2: Posttest
X_1 : Treatment using PBL-ESD
X_2 : Treatment using conventional learning

Population, Sample, and Research Location

The research population is all students of grade VI of Al-Hilal Islamic Primary School, Kota Bharu, Kelantan, Malaysia, in the 2026 school year, which consists of two classes with a total of 48 students. This school was chosen because it represents the characteristics of private Islamic elementary schools in Kelantan that face the problem of low critical thinking skills and real science learning outcomes. The sampling technique used *cluster random assignment* through a lottery, so that class VI Ibn Hibban (n = 30) was designated as the experimental group that received the PBL-ESD model, and class VI Ibn Rushd (n = 18) was determined as the control group that received conventional learning. This technique was chosen because class divisions are administrative and cannot be changed for individual randomization purposes. The cluster random assignment procedure was carried out as follows: (1) the two existing Grade VI classes were listed as eligible clusters; (2) a blind draw was conducted by the school vice-principal as an independent party, ensuring that neither the researcher nor the subject teachers influenced the allocation; (3) the drawn cluster (Ibn Hibban) was assigned to the experimental condition and the remaining cluster (Ibn Rushd) to the control condition. To minimize selection bias, the equivalence of initial ability between both groups was verified through pretest scores prior to treatment, which showed no statistically significant difference (independent t-test: $p > 0.05$ for both HB and BK variables). Additionally, both groups were taught by the same subject teacher with identical instructional time (8 sessions × 30 minutes), differing only in the learning model applied, thereby controlling for teacher effect and time-on-task as potential confounds.

Research Instruments

The research instrument consists of two types. First, The cognitive learning outcome test (HBK) is in the form of a multiple-choice test of 20 items arranged based on the revised Bloom taxonomy covering levels C1 to C6, equipped with four answer choices. Second, The critical thinking ability test (BK) is in the form of 10 reasoned multiple-choice tests that measure aspects of interpretation, analysis, evaluation, explanation, and drawing conclusions based on indicators Facione (2015). indicators, equipped with an assessment rubric to assess the quality of reasoning. Both instruments underwent content validation by two subject-matter experts using the *Content Validity Ratio* (CVR). Content validity was established through expert panel review (n = 2 experts in science education and educational measurement), with all items achieving CVR values exceeding the minimum threshold of 0.62 for a two-expert panel (average CVR = 0.90 for HBK; 0.88 for BK). Reliability was assessed through a pilot study with 30 students from a comparable school not included in the main sample. Reliability analysis using Cronbach's Alpha yielded $\alpha = 0.84$ for HBK and $\alpha = 0.81$ for BK, both exceeding the acceptable minimum threshold of $\alpha > 0.70$ (Field, 2018). Items that did not meet the validity threshold during piloting were revised prior to the main data collection.

Data Analysis Techniques

Data analysis was carried out in stages through four stages. First, descriptive statistics to describe the characteristics of the data through minimum, maximum, *mean*, and standard deviation. Second, the normality test used Shapiro-Wilk—chosen because the control group sample was relatively small (n = 48 < 50), where Shapiro-Wilk was more sensitive and accurate for small samples than Kolmogorov-Smirnov which was more suitable for large samples (Field, 2018). If the data are not normally distributed, hypothesis analysis cannot use parametric tests (t-tests or ANOVAs) because the basic assumptions are not met, so a nonparametric alternative is needed. Third, the variance homogeneity test using *Levene's Test* to confirm whether the variance between groups is homogeneous. Fourth, the hypothesis test uses the Mann-Whitney U as a valid nonparametric test when the assumptions of normality and homogeneity are not met. In addition, it is calculated *Normalized Gain* (N-Gain) with the formula: N-Gain = (posttest score − pretest) / (maximum score − pretest), with categories: high (> 0.70), medium (0.30–0.70), and low (< 0.30) (Hake, 1998). All analyses were conducted using IBM SPSS Statistics version 26 with a significance level of $\alpha = 0.05$.

RESULTS AND DISCUSSION

Results

Descriptive Statistics

Based on the results of descriptive statistical analysis, a comprehensive data picture of cognitive learning outcomes (HB) and critical thinking skills (BK) of the two groups before and after treatment is shown in Table 3. This data is a starting point to understand the initial conditions of the two groups and the magnitude of changes that occurred after the implementation of the PBL-ESD model, so that comparisons between the experimental and control groups can be made objectively and evidence-based.

Table 2. Descriptive Statistics of Research Data

Variable	N	Min	Max	Red	SD
HB Control Pretest	18	10	60	24,72	10,498
HB Control Posttest	18	20	95	46,94	23,772
Pretest Eksp. HB	30	15	60	31,00	11,402
Posttest Eksp. HB	30	45	95	82,50	10,887
BK Control Pretest	18	12	58	23,78	10,310
Posttest of BK Control	18	18	96	44,89	24,139
Pretest Eksp. BK	30	12	56	30,00	11,459
Posttest Eksp. BK	30	48	88	75,93	9,577

Table 2 shows two patterns that are very striking and pedagogically meaningful. First, the average increase in the experimental class was much greater than that of the control class, in both HB (82.50 vs. 46.94) and BK (75.93 vs. 44.89), indicating that the PBL-ESD treatment resulted in substantial changes compared to conventional learning. Second, the standard deviation *of the experimental* class posttest

(HB = 10.887; BK = 9.577) is much smaller than the control class (HB = 23.772; BK = 24.139), indicating that the PBL-ESD model results in a more homogeneous and consistent score distribution. The small standard deviation of the experimental class indicates that almost all students in the group experienced an even increase, not just certain students, so this model proved to be effective in an inclusive manner for all characteristics of students in the classroom.

Normality Test

The Shapiro-Wilk normality test is carried out on all research data as a prerequisite for determining the appropriate inferential analysis technique. The Shapiro-Wilk selection was based on the consideration that the control group had a small sample size (n = 48 < 50), so this method was more precise and sensitive to use than the more reliable Kolmogorov-Smirnov for large samples. The test results are presented in Table 3.

Table 3. Shapiro-Wilk Normality Test Results

Variable	Statistics	df	Sig.	Remarks
HB Control Pretest	0,744	18	0,000	Abnormal
HB Control Posttest	0,894	18	0,046	Abnormal
Pretest Eksp. HB	0,921	30	0,136	Normal
Posttest Eksp. HB	0,710	30	0,000	Abnormal
BK Control Pretest	0,776	18	0,001	Abnormal
Posttest of BK Control	0,886	18	0,033	Abnormal
Pretest Eksp. BK	0,953	30	0,474	Normal
Posttest Eksp. BK	0,944	30	0,343	Normal

Table 3 shows that not all research data are normally distributed. In the HB variable, *the pretest* and *posttest data* of the control class had significance values of 0.000 and 0.046 (< 0.05), respectively, so that it was declared not normally distributed, while *the HB experimental class pretest* was normally distributed (Sig. = 0.136 > 0.05) but the *posttest* was abnormal again (Sig. = 0.000 < 0.05). In the BK variable, *the pretest* and *posttest* of the control class were not normally distributed (Sig. = 0.001 and 0.033 < 0.05), while *the pretest* and *posttest* of the experimental class of BK were both normally distributed (Sig. = 0.474 and 0.343 > 0.05). Because there are data that do not meet the assumption of normality, hypothesis testing cannot use parametric tests and must be continued with nonparametric tests. The notably smaller standard deviation of the experimental group's posttest scores (HB = 10.887; BK = 9.577) compared to the control group (HB = 23.772; BK = 24.139) warrants specific explanation. This pattern is most plausibly attributed to a treatment-induced homogenization effect: the structured PBL-ESD model provided consistent scaffolding, clear problem scenarios, and collaborative group work that helped lower-performing students catch up with higher-performing peers, thereby compressing the score distribution. This interpretation is consistent with Vygotsky (1978) zone of proximal development theory, whereby structured social learning environments

reduce inter-individual performance gaps. This phenomenon is unlikely to reflect selection bias, as the pretest scores of the experimental group showed no statistically significant difference from the control group and exhibited comparable variance at baseline.

Homogeneity Test

Table 4. Levene's Test Homogeneity Test Results

Variable	Living Statistic (F)	Sig.	Remarks
Posttest Learning Outcomes	16.871	0.000	Not Homogeneous
Posttest Critical Thinking	23.766	0.000	Not Homogeneous

By Table 4, the results of Levene's Test showed an F value for the HB posttest of 16.871 (Sig. = 0.000 < 0.05) and the BK posttest of 23.766 (Sig. = 0.000 < 0.05). Both significance values were less than 0.05, so it can be concluded that the data variance between the control class and the experimental class was not homogeneous in both study variables. This inhomogeneity of variance is understandable given the difference in sample size (experiment n = 30 and control n = 18) which led to natural differences in variance between groups. The combination of the abnormal distribution of the data and the inhomogeneity of the variance together reinforces the decision to use the Mann-Whitney U test as a nonparametric test that does not require the assumption of normality or homogeneity, as recommended by the Field, (2018).

Hypothesis Test: Mann-Whitney U

Based on the results of the normality and homogeneity test that showed that the data did not meet the parametric assumptions, the hypothesis test was carried out using the Mann-Whitney U test. This test is a nonparametric test used as an alternative to the independent t-test when the data is not normally distributed and/or the variance is not homogeneous (Field, 2018;Siegel & Castellan, 2019). Unlike the t-test that compares average values, the Mann-Whitney U works on the basis of rank comparison *(rank)* of all the combined scores of the two groups, so that it remains valid even if the assumption of normal distribution is not met. The hypothesis tested was H_0: there was no significant difference between the two groups; and H_1: there was a significant difference between the two groups, with a significance level of $\alpha = 0.05$. The test results are presented in Table 5.

Table 5. Mann-Whitney U Test Results

Variable	Mann-Whitney U	Z	Sig. (2-tailed)	Remarks
Learning Outcomes (HB)	68,000	-4,332	0,000	Significant
Critical Thinking (BK)	82,500	-3,999	0,000	Significant

Based on Table 5, the significance value for both variables is 0.000 < 0.05, so that H_0 is rejected and H_1 is accepted for both study variables. This means that there are significant differences in cognitive learning outcomes (U = 68,000; Z = -4.332; p = 0.000) and critical thinking skills (U = 82,500; Z = -3,999; p = 0.000) between the experimental class that received the PBL-ESD model and the control class that received conventional learning. The negative and large Z-values (Z = -4.332 and -3.999) indicated that the experimental class score was consistently higher than the control class on both variables, so the differences found were not due to mere coincidence, but were a real result of the PBL-ESD treatment given.

N-Gain Analysis

To measure the magnitude of the increase that occurred in each group from *pretest* to *posttest* in a balanced manner by taking into account the initial ability, the N-Gain calculation was performed. The results are presented in Figure 1.

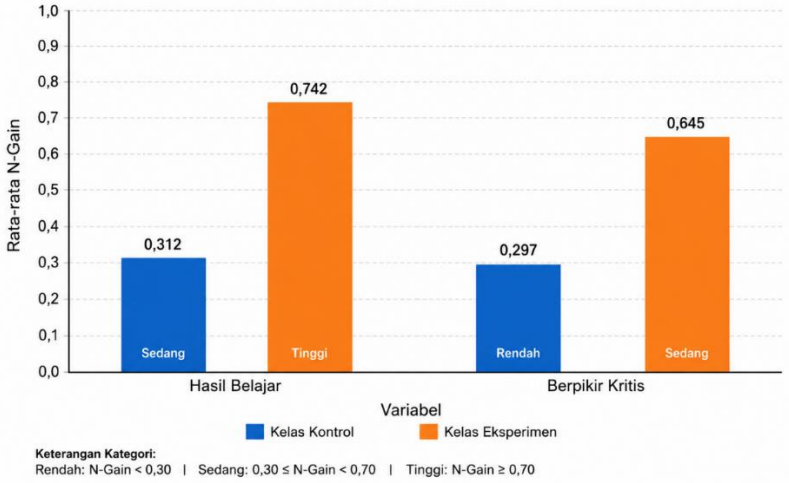


Figure 1. Average N-Gain

Figure 1 shows that the N-Gain of the experimental class in HB of 0.742 (high category) far exceeds the control class of only 0.312 (medium category). Similarly, in BK, the N-Gain of the experimental class of 0.645 (medium category) was significantly higher than that of the control class which was only 0.297 (low category). The N-Gain difference between the two groups at HB of 0.430 points and at BK of 0.348 points confirmed that the PBL-ESD model was much more effective in encouraging real learning improvements, rather than just trivial changes in scores.

Discussion

The findings that PBL-ESD significantly improved cognitive learning outcomes (N-Gain = 0.742, high category) were consistent with the meta-analysis (Merritt *et al*., (2021) which found the effect of PBL on elementary school cognitive learning outcomes by *effect size* d = 0.58. The increase in the high category in this study can be explained through the theory of constructivism Vygotsky, (1978): Kelantan's flood-based problem scenario and local forest degradation provides an authentic learning context that places learners in their proximal development zone, encouraging them to construct science knowledge meaningfully through group discussions and data investigation. Stuart & Stuart (2025) also found that contextualizing problems in the real lives of

learners in PBL not only improves conceptual mastery but also intrinsic motivation for learning which contributes to long-term knowledge retention.

The increase in critical thinking skills (N-Gain = 0.645, medium category) was also in line with the study Mayasari *et al*., (2022) and Ningtias *et al*., (2024). ESD integration provides a significant pedagogical dimension because the sustainability issue that is used as a problem scenario is *open-ended* and multidimensional. In the context of this study, the nature of *open-ended* It is evident when students are faced with the problem of river waste management in Kelantan that does not have a single answer: there are groups that propose biofilter technology solutions, some propose a community education approach, and some that combine the two. Its multidimensional nature is reflected in the need for students to consider the dimensions of natural science (ecological impact of pollution), the social dimension (community behavior), the economic dimension (management costs), and the dimension of Islamic values (the mandate to protect the environment as Allah's creation) simultaneously in formulating solutions. This process forces students to activate the ability to analyze information from various perspectives, evaluate the strengths and weaknesses of each solution, and draw balanced conclusions—the three core dimensions of critical thinking according to Facione (2015).

More specifically, the improvement of students' critical thinking can be attributed to the role of each PBL-ESD syntax applied. In the *problem orientation*, students were introduced to the 2022 Kelantan flood documentation video and river water quality data, which directly activates interpretation skills and authentically identifies scientific problems. In the *collaborative research,* Structured group discussions train the ability to analyze and evaluate peer arguments, an activity identified by Wiek *et al*., (2022) as the most effective in developing critical thinking through ESD. In the *presentation of works,* Students are required to explain their solutions scientifically in front of the class, practicing the dimensions of explanatory and argumentation which are indicators of advanced critical thinking. This entire learning process can be seen in Figure 2.



Figure 2. PBL-ESD Model Learning Process in Class VI of Al-Hilal Islamic Primary School

The N-Gain BK value of the experimental class which is in the medium category (0.645)—lower than HB (0.742)—can be understood not from the difference in sample size, but from the inherent complexity of critical thinking skills as a competency that requires longer habituation. Facione (2015) and Budiarti & Airlanda, (2019) affirms that critical thinking, especially in the dimension of self-evaluation and regulation, develops gradually and requires repeated exposure to complex problem situations. In this study, the duration of 4 meeting sessions may not be enough to mature all dimensions of

critical thinking, especially for students who were previously familiar with memorization learning. Beyond statistical significance, these N-Gain differences carry concrete practical implications for everyday teaching. The high N-Gain for cognitive outcomes (0.742) suggests that teachers can meaningfully substitute conventional lecture-based Science lessons with PBL-ESD problem scenarios without sacrificing content coverage—students in the experimental group demonstrated mastery across all six Bloom's Taxonomy levels despite spending more time on problem-solving activities. For critical thinking, the medium N-Gain (0.645) implies a gradualist implementation strategy: teachers should begin with one or two ESD-contextualized PBL units per semester, initially focusing on lower-order critical thinking dimensions (interpretation and analysis) before progressing to higher-order dimensions (evaluation and self-regulation) as students develop inquiry habits. In practice, the first month of implementation should prioritize structured group discussions with explicit scaffolding; subsequent months can gradually reduce teacher guidance to promote autonomous critical thinking. Professional development programs for Science teachers should incorporate at least 16 hours of PBL-ESD facilitation training, with emphasis on designing local environmental problem scenarios resonant with students' daily experiences in Kelantan. Nevertheless, the role of ESD is strongly felt in driving the dimension of analysis and evaluation: when learners have to consider the long-term environmental impact of each solution they propose, they unconsciously practice high-quality critical thinking because every decision must be scientifically and morally accountable. Further research is recommended to use a longer duration of treatment so that all dimensions of critical thinking can develop more optimally.

Overall, consistent improvements in both variables confirm that the PBL-ESD model works through three mutually reinforcing mechanisms. First, *Cognitive mechanisms:* The PBL inquiry structure forces learners to go through the entire high-level cognitive cycle of Bloom, from analyzing environmental data, evaluating alternative solutions, to creating original problem-solving products. Second, *Motivational mechanisms:* the context of Kelantan's environmental issues experienced directly by students, strengthened by the dimension of Islamic values as a mandate to protect the earth, creates a deep personal relevance so that students' affective involvement in the critical thinking process is much stronger than when facing abstract problems that are not related to their lives (Mufid *et al*., 2024). Third, *Social Mechanisms:* Intensive group discussions in the PBL syntax create a space for verification and rebuttal of arguments between learners, a social practice that is central to the development of critical thinking according to the perspective of Vygotsky's social constructivism. Ismail *et al*., (2024) confirms that environmental issue-based PBL that integrates these three mechanisms simultaneously produces mutually reinforcing effects on problem-solving, critical thinking, and environmental awareness.

CONCLUSION

This study concludes that the PBL-ESD learning model is proven to have a significant and effective influence in improving the cognitive learning outcomes and critical thinking skills of grade VI

students of Al-Hilal Islamic Primary School, Kelantan, Malaysia, compared to conventional learning. The improvement in cognitive learning outcomes of the experimental class was in the high category, while the improvement in critical thinking skills was in the moderate category, both of which were significantly better than the control class. The effectiveness of this model is driven by the synergy between the contextualization of Kelantan's local environmental issues as authentic problem scenarios, PBL's inquiry structure that facilitates high-level thinking, and the integration of Islamic values that strengthen students' motivation and affective engagement. Further research is suggested to use *a true experimental* design with a longer duration, a wider cross-school sample coverage, and add affective variables such as pro-environmental attitudes as a follow-up impact of ESD. For policymakers, this study offers three specific and measurable recommendations. First, the Kelantan State Education Department (JPNK) should mandate the integration of at least two PBL-ESD units per academic year in Grade V–VI Science curricula across Islamic primary schools in Kelantan, supported by standardized ESD problem scenario banks built around locally relevant environmental issues (flood risk, river water quality, deforestation). Second, the Malaysian Ministry of Education (KPM) should incorporate PBL-ESD as a recognized module within the Continuous Professional Development (CPD) framework for primary Science teachers, requiring a minimum of 16 certified training hours per teacher annually, with measurable competency indicators aligned with HOTS targets in the Malaysian Education Blueprint (PPPM) 2013–2025. Third, Islamic primary school administrators should establish termly peer-learning communities (komuniti pembelajaran profesional) in which Science teachers collaboratively design, implement, and evaluate PBL-ESD lessons, enabling systematic monitoring of student critical thinking development across grade levels. These recommendations are directly grounded in this study's empirical evidence and aligned with Malaysia's commitments under SDG 4 (Quality Education) and SDG 13 (Climate Action) of the United Nations Sustainable Development Goals.

THANKS

The author would like to thank Al-Hilal Islamic Primary School, Kota Bharu, Kelantan, Malaysia, for the permission and support during the data collection process, as well as to all students and teachers who have participated in this study.

REFERENCES

Ardianti, R., Sujarwanto, E., & Surahman, E. (2021). Problem-based learning: Apa dan bagaimana. *DIFFRACTION: Journal for Physics Education and Applied Physics*, *3*(1), 27–35. http://jurnal.unsil.ac.id/index.php/Diffraction

Arifin, M., Yunira, Y., Harahap, S. E., & Marbun, E. (2024). Penerapan model PBL dalam pembelajaran IPAS untuk meningkatkan motivasi dan hasil belajar siswa. *Journal of Education Research*, *5*(4), 6109–6121. https://doi.org/10.55187/jer.v5i4.8451

Asda, E. F. (2024). ANALISIS PERBEDAAN METAKOGNISI SISWA PADA PEMBELAJARAN KIMIA DITINJAU DARI GENDER DAN PRESTASI AKADEMIK. *Dalton : Jurnal Pendidikan Kimia Dan Ilmu Kimia*, *7*(2). http://

dx.doi.org/10.31602/dl.v7i2.15886

Budiarti, I., & Airlanda, G. S. (2019). *Jurnal Riset Teknologi dan Inovasi Pendidikan Penerapan model problem based learning berbasis kearifan lokal untuk meningkatkan keterampilan berpikir kritis*. *2*(1), 167–183. https ://doi.org/ 10.29407/jrtip.v2i1.12743

Cahyani, V. P., & Ahmad, F. (2024). Efektivitas problem based learning terhadap keterampilan berpikir kritis, hasil belajar dan motivasi siswa. *Jurnal Pendidikan Sains*, *12*(1), 76–82.https://doi.org/10.53696/venn.v3i2.155

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications. https://lccn.loc.gov/2022032270

Facione, P. A. (2015). *Critical Thinking : What It Is and Why It Counts*. 1–30.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th editio). SAGE Publications.

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

Hasanah, S. U., Akasyah, D., & Luthfi, A. (2025). *REVITALIZATION OF THE MODERN MOVEMENT IN ISLAM AS A*. 1299–1310. https://doi.org/ 10.30868/im.v8i02.8729

Ismail, R., Imawan, O. R., Inayah, S., Kau, M. S., Sulfiati, Y., Nasrullah, A., & Yuliati, Y. (2024). *Pembelajaran dengan Problem Based Learning, Stratei dan Implementasi*. Edupedia Publisher.

Kementerian Pendidikan Malaysia. (2023). *Laporan tahunan PPPM 2023: Transformasi sistem pendidikan.* KPM.

Mayasari, A., Arifudin, O., & Juliawati, E. (2022). *IMPLEMENTASI MODEL PROBLEM BASED LEARNING (PBL)*. *3*(2), 167–175. https://doi.org/10.38035/jmpis.v3i2.1234

Merritt, J., Lee, M. Y., Rillero, P., Kinach, B. M., Merritt, J., Lee, M. Y., Rillero, P., Kinach, B. M., & State, A. (2021). *Problem-Based Learning in K – 8 Mathematics and Science Education : A Literature Review The Interdisciplinary Journal of Problem-based Learning Special Issue On Competency Orientation in Problem-BasED Learning Problem-Based Learning in K – 8 Mathematics a*. *11*(2), 5–17. https://doi.org/ 10.7771/ 1541-5015.1972

Mufid, F., Nugraha, A. R., & Shobaruddin, D. (2024). *Islamic Education and Sustainable Bridging Faith and Global Goals Development : 01*(3), 173-180.https://doi.org/ 10.29408/jicp.v1i3.20341

Mutakin, T. Z., Tola, B., & Hayat, B. (2023). *Analisis Kemampuan Siswa Sekolah Dasar Dalam Menyelesaikan Soal Matematika Menggunakan Framwork TIMSS 2019*. *58*, 225–236. https://doi.org/10.55215/jprima.v58i1.9876

Mutiara, F. B., Komikesari, H., & Asiah, N. (2019). *EFEKTIVITAS MODEL KOOPERATIF TIPE COURSE REVIEW HORAY (CRH) TERHADAP HASIL BELAJAR FISIKA SISWA EFFECTIVENESS OF COOPERATIVE MODEL TYPE COURSE REVIEW HORAY (CRH) TOWARDS STUDENTS '*

LEARNING. *02*(1), 116–122. https://doi.org/10.31539/jfpk.v2i1.598

Ningtias, Y., Panjaitan, D. J., Nasution, A. E., & Indayanti, Y. (2024). *Penerapan Model Problem Based Learning Untuk Meningkatkan Kemampuan Berpikir Kritis Matematis Siswa Di SMA Negeri 2 Medan*. *06*(04), 21756–21764.https://doi.org/10.31004/basicedu.v6i4.12345

Purwanto, H., Aminah, S., Ramadhani, W., & Azim, F. (2020). *Penerepan Model Pembelajaran Problem Based Learning untuk Meningkatkan Sikap Ilmiah dan Hasil Belajar Kognitif Mahasiswa pada Mata Kuliah Ekologi Tumbuhan*. *3*(2), 151–162. https://doi.org/10.29408/bioeduca.v3i2.2891

Rahman, R., & Fuad, M. (2023). Implementasi kurikulum merdeka belajar dalam pembelajaran IPAS di sekolah dasar. *DISCOURSE: Indonesian Journal of Social Studies and Education*, *1*(1), 75–80. https://doi.org/10.69875/djosse.v1i1.103

Rofi'ah, S., & Rokhmaniyah. (2024). Analisis Kemampuan Berpikir Kritis Siswa dalam Memecahkan Masalah pada Mata Pelajaran IPAS kelas V Sekolah Dasar. *Social, Humanities, and Educational Studies SHEs:*, *7*(3), 1763–1770.https://doi.org/10.23887/shes.v7i3.65432

Septiana, T. S., & Kurniawan, M. R. (2018). *PENERAPAN MODEL PROBLEM BASED LEARNING UNTUK MENINGKATKAN BERPIKIR KRITIS SISWA KELAS 5 PADA MATA PELAJARAN PKN DI SD MUHAMMADIYAH KAUMAN*. *1*(1), 94–105. https://doi.org/10.12928/fundadikdas.v1i1.74

Siegel, S., & Castellan, N. J. (2019). *Nonparametric statistics for the behavioral sciences* (2nd ed). McGraw-Hill.

Susilawati, E., Agustinasari, A., Samsudin, A., & Siahaan, P. (2020). Analisis tingkat keterampilan berpikir kritis siswa SMA. *Jurnal Pendidikan Fisika Dan Teknologi*, *6*(1), 11–16. https://doi.org/10.29303/jpft.v6i1.1453

UNESCO. (2020). *Education for sustainable development: A roadmap*. UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000374802

Utami, R. T., & Desstya, A. (2021). Analisis Cakupan Literasi Sains dalam Buku Siswa Kelas V Tema 4 Karya Ari Subekti di Sekolah Dasar Ririn Tri Utami 1 □ , Anatri Desstya 2. *JURNALBASICEDU*, *5*(6), 5001–5013.https://doi.org/10.31004/basicedu.v5i6.1556

Vygotsky, L. S. (1978). *The Development of Higher Psychological Processes . The Development of Higher*. Harvard University Press.

Wahyudi, A., & Purwanto, H. (2024). Menganalisis aspek peningkatan kemampuan berpikir kritis siswa dalam pembelajaran IPA. *Jurnal Pendidikan Sains*, *1*(4), 112–125. https://doi.org/10.62383/edukasi.v1i4.640

Wiek, A., Bernstein, M. J., Foley, R. W., Cohen, M., Forrest, N., Kuzdas, C., & Kay, B. (2022). *COMPETENCIES IN HIGHER EDUCATION FOR SUSTAINABLE DEVELOPMENT*. 241–260. https://doi.org/ 10.4324/9780429202964-18

Wuriningsih, F. R., & Agung, A. (2025).

Reforming Pedagogy in Indonesian Primary Schools : A Five-Year Systematic Review of Differentiated Instruction , Project-Based Learning , and Formative. *10*, 1694–1705. https://doi.org/10.24247/jpmpkdec.2025.105