JOURNAL OF HUMAN AND EDUCATION

E-ISSN: 2776-5857 | P-ISSN: 2776-7876
Website: https://jahe.or.id/index.php/jahe
Vol. 6, No. 5 (2026)

Problem-Based Learning and Final Cooling-System Competency in Heavy-Duty Diesel Engine Education: A Quasi-Experimental Pretest-Posttest Study

Ahmat Nuron Makki[1*], Erwin Komara Mindarta[2], Besar Abadi Prayogi[3], Sahirul Alim[3], Novia Riza[3]

[1]Teacher Professional Education, Graduate School, State University of Malang, Malang, Indonesia
[2]Automotive Engineering Education, Department of Mechanical and Industrial Engineering, Faculty of Engineering, State University of Malang, Malang, Indonesia
[3]Automotive Engineering Program, SMK Negeri 6 Malang, Malang, Indonesia
**Email: ahmat.nuron.2531547@students.um.ac.id*

ABSTRAK

Problem-Based Learning (PBL) sesuai untuk pendidikan teknik karena mengorganisasi pembelajaran melalui masalah autentik dan penalaran diagnostik. Penelitian ini menguji apakah PBL berkaitan dengan kompetensi akhir cooling system yang lebih tinggi pada pembelajaran diesel engine alat berat serta apakah keunggulan tersebut tercermin pada peningkatan skor yang lebih besar atau pada capaian akhir. Desain quasi-experimental pretest-posttest melibatkan 61 siswa kelas XI Teknik Alat Berat (PBL, $n = 32$; kontrol, $n = 29$). Analisis meliputi statistik deskriptif, uji berpasangan, perbandingan Welch, normalized gain, effect size, analisis sensitivitas non-parametrik, dan ANCOVA dengan standard error robust HC3. Kedua kelompok meningkat secara substansial (PBL: $50{,}00 \pm 9{,}84$ menjadi $89{,}06 \pm 6{,}02$; kontrol: $45{,}69 \pm 8{,}53$ menjadi $84{,}83 \pm 8{,}40$; keduanya $p < 0{,}001$), sedangkan raw gain hampir identik (39,06 vs. 39,14; $p = 0{,}979$). Kelompok PBL memperoleh skor post-test lebih tinggi (selisih = 4,23 poin; 95% CI 0,44–8,03; Hedges' $g = 0{,}58$), dan selisih setelah penyesuaian baseline tetap 3,98 poin (HC3 $p = 0{,}048$). Dengan demikian, PBL berkaitan dengan capaian akhir yang lebih tinggi, bukan peningkatan absolut yang lebih besar. Evaluasi pedagogi vokasi perlu membedakan perubahan skor dari tingkat kompetensi akhir.

Kata Kunci: *problem-based learning; pendidikan vokasi; cooling system; diesel engine; capaian Kompetensi*

ABSTRACT
Problem-Based Learning (PBL) is suited to technical education because it organizes learning around authentic problems and diagnostic reasoning. This study examined whether PBL was associated with higher final cooling-system competency in heavy-duty diesel engine learning and whether any advantage reflected larger learning gain or higher endpoint attainment. A quasi-experimental pretest-posttest design involved 61 Grade XI Heavy Equipment Engineering students (PBL, n = 32; control, n = 29). Descriptive statistics, paired tests, Welch comparisons, normalized gain, effect sizes, non-parametric sensitivity analyses, and ANCOVA with HC3 robust standard errors were applied. Both groups improved substantially (PBL: 50.00 ± 9.84 to 89.06 ± 6.02; control: 45.69 ± 8.53 to 84.83 ± 8.40; both p < 0.001), while raw gains were virtually identical (39.06 vs. 39.14; p = 0.979). The PBL group nevertheless achieved a higher post-test score (difference = 4.23 points, 95% CI 0.44–8.03; Hedges' g = 0.58), and the baseline-adjusted difference remained 3.98 points (HC3 p = 0.048). PBL was therefore associated with higher final attainment, not a larger absolute gain. Evaluating vocational pedagogy through both change and endpoint performance provides a more precise interpretation of instructional effects in intact-class designs.
Keywords: *problem-based learning; vocational education; cooling system; diesel engine; competency attainment*

INTRODUCTION

Technical and vocational education is expected to develop more than declarative knowledge. In occupational domains such as heavy-equipment maintenance, students must learn to interpret technical systems, connect symptoms with underlying causes, reason across interacting components, and select actions that are defensible in relation to the evidence available. Cooling-system learning within a heavy-duty diesel engine course illustrates this challenge. A learner may be able to name a radiator, thermostat, water pump, cooling fan, coolant passage, and pressure cap while still being unable to reason through overheating, abnormal coolant circulation, pressure loss, or heat-transfer problems as an integrated system. For this reason, a meaningful learning outcome in this domain should reflect the ability to organize technical knowledge around functional relationships rather than isolated facts.

Problem-Based Learning (PBL) is conceptually compatible with this type of learning requirement. In a PBL environment, a problem is not merely an exercise placed at the end of a lesson; it structures the learning process itself. Students are expected to identify what is known, determine what needs to be understood, investigate relevant information, propose explanations, test the coherence of those explanations, and justify a solution. In a technical context, this sequence resembles the reasoning expected during fault diagnosis. The educational attraction of PBL therefore lies in the possibility that students learn the cooling system as a network of causes, functions, and observable consequences instead of as a collection of components that must be memorized separately.

Engineering-education research supports movement from transmission toward problem-centered learning, but PBL is not a uniform intervention. A review of 108 studies documented implementation at project, course, cross-course, and curriculum levels, with recurring student, instructor, institutional, and cultural challenges (Chen et al., 2021). Reviews and framework studies likewise emphasize student readiness, teacher preparation, curriculum alignment, assessment, and institutional support when shifting from lectures to PBL (Das, 2026; Devika, Siddapuram, Kaur, & Bollampally, 2024; Patnawar, 2023). PBL is therefore better treated as a structured pedagogy whose effects depend on implementation quality. This view is consistent with practice-based engineering frameworks centered on authentic professional activity and assessment alignment (Mann et al., 2021), and with cooperative PBL evidence showing positive acceptance alongside workload and assessment challenges (Alaloul & Qureshi, 2022).

Empirical studies across engineering disciplines provide a plausible basis for expecting PBL to influence higher-order learning, although outcomes vary by context and measure. In-home laboratory PBL was associated with improved performance, motivation, and self-efficacy, while a freshman engineering case reported gains in performance, teamwork, and engagement (Ariza, 2022; Naik et al., 2024). Biomedical and circuit-theory studies respectively linked PBL with confidence in higher-order tasks and improved critical-thinking disposition (Long et al., 2022; She-Ping et al., 2023). Electrical and civil engineering studies also report positive student experiences and workable implementation in demanding courses (Katageri & Raikar, 2022; Naveh et al., 2022; Tan & Shen, 2022). Technical-education and antennas-course studies further report improved autonomous work and stronger academic or program outcomes (Bilbao et al., 2024; Devika, Siddapuram, Kaur, & Kondala Rao, 2024).

Nevertheless, the claim that a pedagogical approach is "effective" requires a precise empirical definition. A common approach in classroom research is to compare pre-test and post-test scores and interpret a significant increase after instruction as evidence that the treatment worked. This logic is incomplete when a control group also receives instruction. Students in both groups are expected to improve because both participate in learning activities, encounter the subject matter, and complete the instructional period. A statistically significant pre-to-post change in the experimental group therefore establishes that learning occurred, but it does not establish that the experimental method produced a larger improvement than the alternative condition. The effect attributable to the instructional model must be isolated through an explicit between-group comparison.

This distinction becomes especially important in a pretest-posttest design using intact classes. If the groups do not start at exactly the same level, a direct comparison of final scores can reflect both baseline differences and instructional effects. Conversely, a comparison based only on raw gain can obscure meaningful differences when the outcome scale has an upper bound. Students who begin at a higher level have less numerical room to improve before reaching the maximum score, so the same absolute gain can represent a different proportion of the improvement that was available. A rigorous analysis therefore benefits from examining multiple estimands: baseline status, post-test status, raw change, normalized change, and baseline-adjusted final performance.

The dataset used in the present study makes this analytical problem visible. It contains pre-test and post-test scores for two Grade XI Heavy Equipment Engineering classes studying the cooling system within the Diesel Engine Alat Berat subject. The records identify XI TAB 1 as the experimental group and XI TAB 2 as the control group, and they identify PBL as the learning model associated with the experimental condition. The experimental class contains 32 students, whereas the control class contains 29 students. All observed assessment scores are available for both time points.

At baseline, the PBL group had an average score of 50.00 and the control group an average of 45.69. The difference of approximately 4.31 points was not statistically significant under a conventional two-sided Welch test at the 0.05 level, but it was large enough to matter analytically. The experimental group therefore entered the study with a somewhat higher observed level of performance. At post-test, both groups improved markedly, and both distributions shifted upward. The key question is not whether learning occurred, because the dataset clearly shows that it did. The key question is how the patterns of improvement and final attainment differ between the PBL and control conditions after the starting position of each group is considered.

Problem-centered learning is particularly relevant to vocational education, where knowledge must support job-related action. TVET studies position PBL among pedagogical principles that connect theory, active participation, and workplace learning, while stressing teacher development and institutional support (Okolie et al., 2021; Okolie et al., 2023). Problem-based blended learning and integrated inquiry-PBL designs have shown favorable effectiveness or learning-outcome patterns in vocational settings (Maksum et al., 2023; Rafiq et al., 2023). Task-based problem scenarios are positively associated with practical skills in vocational electrical technology, and quasi-experimental STEM-PBL evidence shows improved problem-solving ability among higher-vocational engineering students (Orji, 2025; Wang & Somasundram, 2025). Automotive-vocational studies also link PBL-e-learning integration with academic performance, workplace competencies, perceived value, satisfaction, engagement, and motivation (Maksum et al., 2025; Wagino et al., 2025; Wulansari et al., 2024).

Mechanically oriented studies are especially pertinent because they require reasoning about interacting physical systems. A PBL intervention in turbomachinery improved higher-order outcomes and performance consistency, while integrated PBL in mechanical vibration improved technical and sustainability competencies (Bin Mamat, 2026; Zhao et al., 2026). In mechanical education, PBL combined with example-based learning supported technical and non-technical skills, and a brake-system course used real problems to connect component function with practical interpretation (Morales et al., 2026; Sanchez-Gomez, 2022). Industrial-engineering studies report benefits from PBL combined with project work, flipped instruction, or Industry 4.0 activities, although these are multi-component interventions (Farooq & Jang, 2026; Gómez-Coma et al., 2023; Jeyanathan et al., 2026; Rodriguez-Sanchez et al., 2024). Continuing engineering education in manufacturing likewise positions PBL as a route from evolving technical knowledge to industrial practice (Andersen & Rösiö, 2023).

Despite the breadth of PBL research across engineering and vocational education, direct evidence remains sparse for component-level heavy-equipment learning in which students must reason about an integrated thermal-mechanical system. The literature spans civil, electrical, chemical, biomedical, manufacturing, and broader TVET contexts, but it rarely distinguishes a larger pretest-to-posttest gain from a higher final level of attainment after baseline performance is considered. That distinction is consequential in intact-class studies because both experimental and comparison groups can improve substantially during ordinary instruction. The empirical gap addressed here is therefore not the generic question of whether PBL can raise test scores, but whether the PBL condition is associated with superior final cooling-system attainment once general learning gains and baseline differences are separated analytically.

This study therefore treats instructional effectiveness as a two-dimensional outcome: displacement from baseline and destination at the end of instruction. The distinction is not semantic. In a bounded competency scale, two groups can display comparable numerical movement while differing in the level of performance ultimately reached. Framing the analysis around both estimands makes the contribution more specific to vocational education, where the educational question is not only whether students improve, but whether they finish close to the level of technical understanding expected for reliable diagnostic work.

The present study was guided by the hypothesis that students in the PBL group would demonstrate higher final cooling-system competency assessment scores than students in the control group after baseline performance was taken into account. A secondary analytical objective was to determine whether any post-test advantage was accompanied by a larger raw pre-to-post gain. These objectives intentionally distinguish final attainment from change because the two quantities answer different educational questions. Gain asks how far a student moved from the starting point. Final attainment asks where the student ended after instruction. In vocational education, the second question may be particularly consequential because occupational readiness depends on reaching an adequate level of technical understanding, not merely on improving relative to a low baseline.

The study therefore contributes an argument that is both pedagogical and methodological. Pedagogically, it examines PBL in the focused context of cooling-system learning within heavy-duty diesel engine education. Methodologically, it demonstrates why evidence for an instructional advantage should not be inferred from within-group significance alone. The analysis combines descriptive statistics, within-group change tests, between-group comparisons, effect sizes, normalized gain, and analysis of covariance (ANCOVA). Because the post-test distributions are bounded and show unequal group variances, robust and non-parametric sensitivity analyses are also considered. This layered approach is intended to produce an interpretation that follows the actual structure of the data rather than forcing the results into a predetermined narrative.

Accordingly, the article addresses three questions. First, to what extent did cooling-system assessment scores increase from pre-test to post-test within the PBL and control groups? Second, did the groups differ in raw gain and final post-test performance? Third, did the difference in final performance remain after adjustment for baseline score? The central expectation was that PBL would be associated with a higher final level of attainment, while the magnitude of raw gain was treated as an empirical question rather than assumed to be larger in advance.

The resulting argument is deliberately conservative. A quasi-experimental classroom dataset cannot, by itself, establish that all observed differences were caused exclusively by PBL. Intact classes may differ in characteristics that are not represented in the score file, and the instructional record available for analysis does not provide the detail needed to isolate every component of the teaching process. The study therefore focuses on what can be established directly from the observed scores: the size and distribution of learning changes, the contrast between groups, and the persistence of the post-test difference after statistical adjustment for baseline performance. This approach allows the manuscript to make a defensible contribution while retaining an appropriate boundary between empirical evidence and pedagogical interpretation.

METHOD

Study design and educational context

The study used a two-group pretest-posttest comparison based on classroom assessment records from Grade XI Heavy Equipment Engineering students. The learning context was the Diesel Engine Alat Berat subject, with the instructional content focused on the cooling system. The dataset identifies two intact classes. XI TAB 1 was coded as the experimental group and associated with Problem-Based Learning, whereas XI TAB 2 was coded as the control group. Because the available records describe pre-existing classes rather than individual random assignment, the analysis was treated as quasi-experimental.

The use of a pretest-posttest structure was important for two reasons. First, it allowed the amount of change within each class to be quantified. Second, the pre-test provided a baseline covariate that could be used to determine whether a difference in post-test performance persisted after accounting for initial score variation. The design therefore supported complementary estimands rather than relying on a single comparison.

Participants and analytic sample

The score file contained 61 students with complete pre-test and post-test observations. The experimental/PBL class included 32 students and the control class included 29 students. No student-level observations were missing within the analyzed score ranges. Student names were not used in the statistical analysis or in the reporting of results; each record was treated only as a paired set of pre-test and post-test scores with a group label.

The experimental group's pre-test scores ranged from 35 to 65, and its post-test scores ranged from 80 to 100. The control group's pre-test scores ranged from 35 to 60, and its post-test scores ranged from 75 to 95. Scores were recorded on a 0-100 scale in five-point increments in the supplied dataset. Because the available file contained total scores rather than item-level responses, the analysis evaluates observed assessment attainment and does not estimate internal-consistency reliability, item difficulty, discrimination, or latent proficiency.

Assessment instrument and scoring

Pre-test and post-test performance was recorded as a total cooling-system assessment score on a 0–100 scale. Every observed total occurred in five-point increments, yielding 20 equal score intervals across the scale. The available research records contain total scores but not the original test blueprint, item-level responses, or scoring rubric; therefore, the item format and item count cannot be verified from the supplied documentation. The five-point spacing is consequently treated as a property of the recorded scoring scale rather than assumed to represent a specific 20-item multiple-choice or rubric-based instrument.

Instructional condition

The instructional record identifies Problem-Based Learning as the learning model for the experimental condition. Within the analytical framework of this study, the group variable therefore distinguishes students taught in the class identified as the PBL experimental group from students in the class identified as the control group. The score file does not encode session-by-session implementation features, fidelity measures, teacher behaviors, problem scenarios, or the specific instructional sequence used in the control class. Consequently, the statistical analysis estimates the difference between the observed classroom conditions as implemented, rather than the isolated effect of any single PBL component.

This distinction is important for interpretation. The experimental contrast is an instructional-condition contrast at the class level. It is not a laboratory manipulation in which every exposure other than PBL was held constant. The conclusions are therefore framed in terms of association with the PBL condition and baseline-adjusted differences in final attainment.

Outcome variable and derived measures

The primary observed outcome was the post-test score for cooling-system learning. The corresponding pre-test score was used as an indicator of baseline performance. Three derived measures were calculated.

First, raw gain was defined for each student as:

Gain = Post-test score - Pre-test score.

Second, individual normalized gain was calculated as:

Normalized gain = (Post-test score - Pre-test score) / (100 - Pre-test score).

This metric expresses the observed improvement as a proportion of the maximum numerical improvement still available to the student at baseline. Because no student had a pre-test score of 100, the denominator was non-zero for all records.

Third, baseline-adjusted post-test means were estimated from an ANCOVA model with post-test score as the dependent variable, group as the categorical predictor, and pre-test score as the covariate. Adjusted means were evaluated at the grand mean pre-test score across the complete sample.

Statistical analysis

The analysis was conducted in sequential layers. Descriptive statistics were first calculated for pre-test, post-test, raw gain, and normalized gain. Means and standard deviations were used to summarize central tendency and dispersion, while medians, interquartile ranges, and ranges were examined to characterize the bounded and discrete score distributions.

Within-group change was tested using paired-samples t tests. The standardized magnitude of paired change was summarized using Cohen's dz, calculated as the mean paired difference divided by the standard deviation of the paired differences. Because several score distributions departed from normality, Wilcoxon signed-rank tests were also calculated as sensitivity analyses for the within-group comparisons.

Between-group comparisons of pre-test score, post-test score, raw gain, and individual normalized gain were performed using Welch's independent-samples t test. Welch's procedure was selected because it does not assume equal variances and is therefore preferable when group dispersions differ. For each comparison, the mean difference was defined as PBL minus Control, and a 95% confidence interval was calculated using the Welch-Satterthwaite degrees of freedom. Between-group standardized differences were summarized using Hedges' g, which applies a small-sample correction to the pooled standardized mean difference. Mann-Whitney U tests were used as non-parametric sensitivity analyses.

To evaluate the primary hypothesis while accounting for baseline differences, an ANCOVA model was fitted:

Post-test = intercept + group + pre-test + error.

The assumption of homogeneous regression slopes was examined by fitting an interaction model including pre-test × group. A non-significant interaction was interpreted as evidence that a common baseline-post-test slope was a reasonable working approximation. Because the post-test variance differed between groups, heteroscedasticity-consistent HC3 standard errors were used for the primary coefficient-level inference from the ANCOVA model. Conventional type-II ANCOVA results and partial eta squared were also reported to describe the proportion of residual outcome variation attributable to group after adjustment for baseline.

Distributional assumptions were inspected using Shapiro-Wilk tests and variance homogeneity using a median-centered Levene test. These diagnostics were not used mechanically to accept or reject the analysis; rather, they informed the choice to present Welch, robust, and non-parametric results together. Statistical tests were two-sided with alpha set at 0.05. Given the modest sample size and the bounded five-point score increments, emphasis was placed on effect estimates, confidence intervals, and consistency across analytical approaches rather than on p values alone.

Because instructional condition was assigned at the intact-class level and only one class represented each condition, the design contains no independent class-level replication from which to estimate between-class variance or a cluster-robust treatment effect. The reported student-level inferential statistics are therefore interpreted as contrasts between the two observed classroom conditions, not as class-replicated causal estimates of PBL.

Analytical logic

The inferential sequence was designed to prevent three common interpretive errors. The first is equating a significant pre-to-post increase in the experimental group with superiority over the control group. The second is interpreting a higher post-test mean without checking whether it can be explained by baseline imbalance. The third is assuming that raw gain and final attainment are interchangeable indicators of instructional effectiveness.

The analysis therefore asked progressively narrower questions. Paired tests established whether each group changed. Welch tests then established whether the groups differed in their observed endpoints and changes. Finally, ANCOVA assessed whether the post-test difference remained after accounting for baseline score. The normalized-gain analysis was supplementary and was used to examine improvement relative to the numerical room available on the 0-100 scale. This hierarchy of analyses supports an interpretation in which each statistical result answers a specific educational question.

RESULTS AND DISCUSSION

Descriptive pattern of pre-test and post-test scores

The full dataset consisted of 61 complete paired observations, with 32 students in the PBL group and 29 in the control group. Descriptive statistics are presented in Table 1. The PBL group entered the study with a mean pre-test score of 50.00 (SD = 9.84), compared with 45.69 (SD = 8.53) in the control group. The corresponding medians were 50 and 45. The experimental group therefore had a 4.31-point higher baseline mean.

Table 1. Descriptive statistics for cooling-system assessment scores by instructional group.

Group	n	Pre-test Mean ± SD	Pre-test Median [IQR]	Post-test Mean ± SD	Post-test Median [IQR]	Raw gain Mean ± SD	Group-level normalized gain
Control	29	45.69 ± 8.53	45 [40–50]	84.83 ± 8.40	85 [75–95]	39.14 ± 11.81	0.721
PBL	32	50.00 ± 9.84	50 [40–60]	89.06 ± 6.02	90 [85–90]	39.06 ± 10.81	0.781

Note. PBL = Problem-Based Learning; IQR = interquartile range. Group-level normalized gain = (post-test mean − pre-test mean)/(100 − pre-test mean).

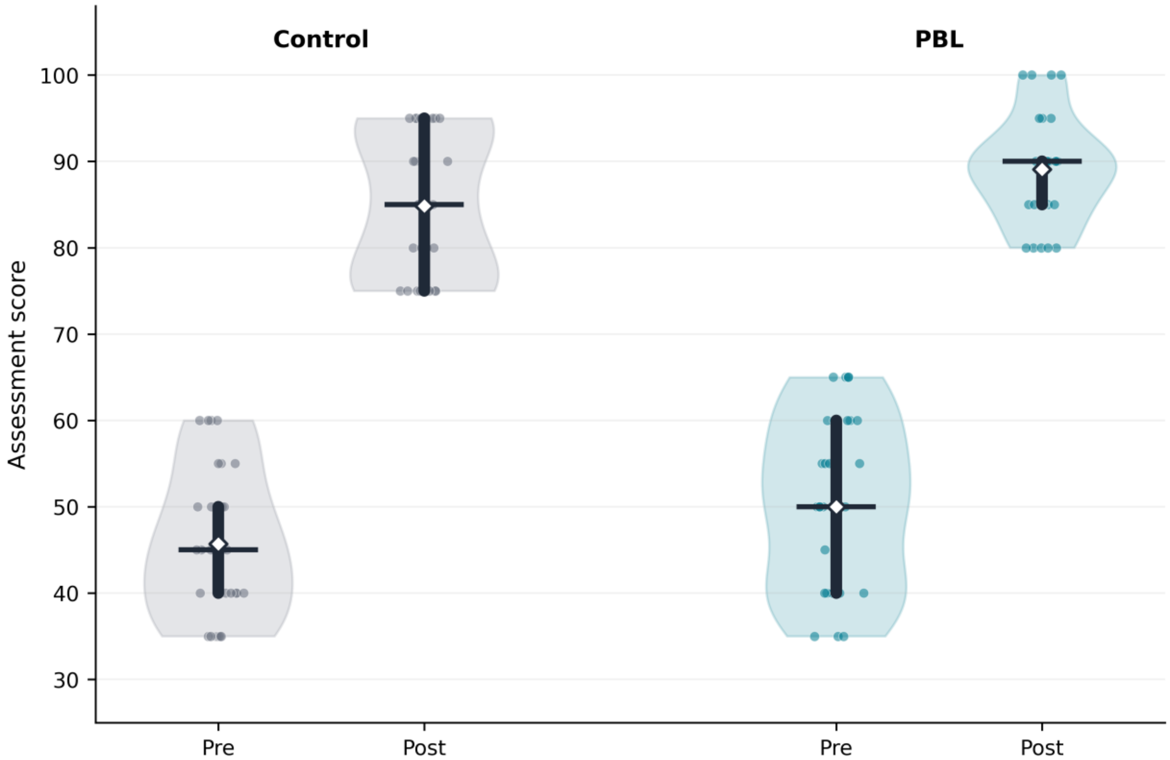

Fig. 1. Distributions of pre-test and post-test cooling-system assessment scores. Points represent students; diamonds indicate group means, thick vertical segments indicate interquartile ranges, and horizontal bars indicate medians.

At post-test, both groups shifted to substantially higher score ranges. The PBL group achieved a mean of 89.06 (SD = 6.02; median = 90; range = 80-100), whereas the control group achieved a mean of 84.83 (SD = 8.40; median = 85; range = 75-95). The raw post-test difference was therefore 4.23 points in favor of the PBL group. Fig. 1 shows the score distributions and makes two features visible: the large upward shift from pre-test to post-test in both groups and the tighter post-test distribution in the PBL group.

The difference in dispersion at post-test was notable. The PBL standard deviation decreased from 9.84 at baseline to 6.02 at post-test, whereas the control-group standard deviation remained approximately unchanged from 8.53 to 8.40. The post-test Levene test was significant ($p = 0.0039$), indicating unequal variances between groups. This result justified the use of Welch's between-group tests and heteroscedasticity-robust standard errors in the baseline-adjusted analysis.

Within-group learning changes

Both groups exhibited strong pre-to-post improvement. The control group increased by an average of 39.14 points (SD of the paired gain = 11.81; 95% CI 34.65 to 43.63). The paired-samples test yielded $t(28) = 17.85$, $p < 0.001$. The standardized paired effect was $d_z = 3.31$. The Wilcoxon signed-rank sensitivity test also showed a clear upward shift ($p < 0.001$).

The PBL group increased by an average of 39.06 points (SD of the paired gain = 10.81; 95% CI 35.17 to 42.96). The paired-samples test yielded $t(31) = 20.44$, $p < 0.001$, with a standardized paired effect of $d_z = 3.61$. The Wilcoxon signed-rank result was also significant ($p < 0.001$). Fig. 2 plots individual trajectories. Every trajectory moves upward from pre-test to post-test, confirming that the improvement was not produced by a small subset of extreme observations.

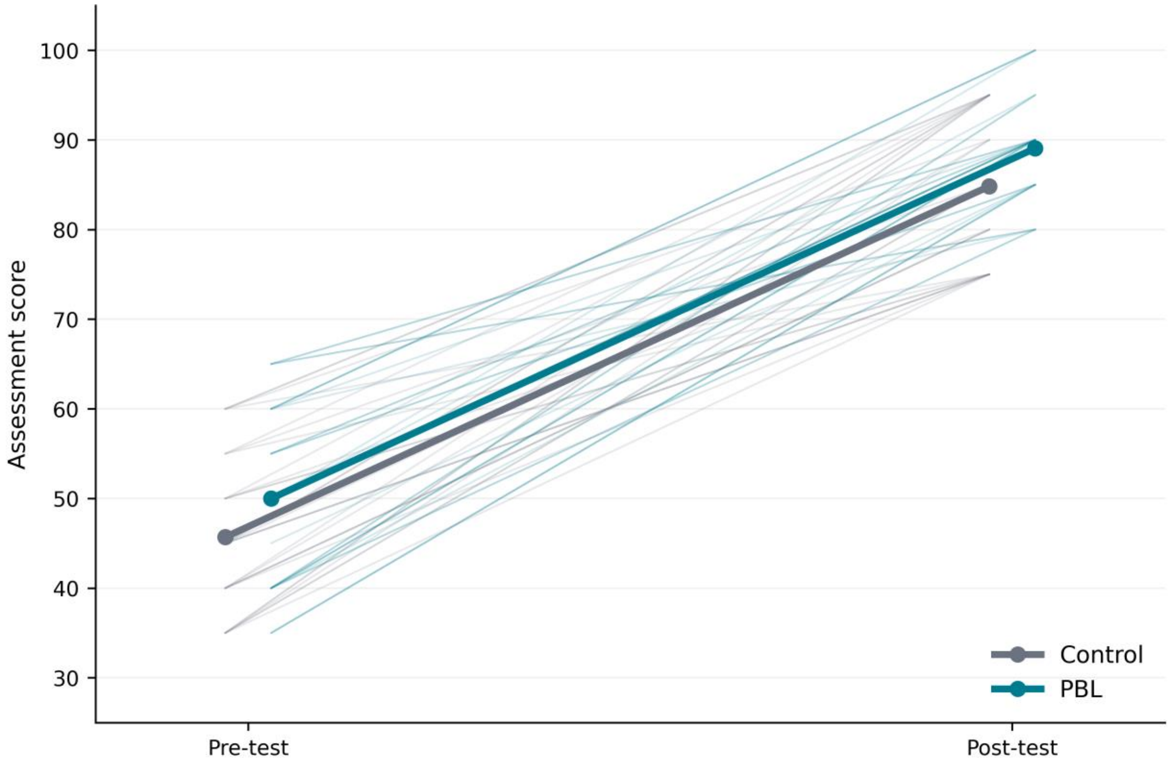

Fig. 2. Individual pre-test to post-test trajectories by instructional group. Thin lines represent individual students; thick lines connect group means.

These within-group results establish that substantial learning occurred under both observed instructional conditions. They do not, however, establish that the PBL group improved more than the control group. That question requires a direct comparison of the change scores.

Baseline comparison and raw post-test difference

The baseline difference between groups was 4.31 points in favor of the PBL group. Welch's t test yielded $t = 1.83$ with approximately 58.89 degrees of freedom, $p = 0.072$, and a 95% confidence interval from -0.40 to 9.02 points. The corresponding Hedges' g was 0.46. The Mann-Whitney sensitivity test produced $p = 0.088$. These results do not provide conventional statistical evidence of a baseline difference at alpha = 0.05, but the effect estimate is large enough that baseline should not be ignored in a quasi-experimental comparison.

For the post-test, the PBL group exceeded the control group by 4.23 points. Welch's t test yielded $t = 2.24$ with approximately 50.28 degrees of freedom, $p = 0.029$, with a 95% confidence interval from 0.44 to 8.03 points. The standardized between-group effect was Hedges' $g = 0.58$. The non-parametric Mann-Whitney test was less decisive ($p = 0.065$), reflecting the coarse five-point increments, bounded scale, and clustering of post-test scores at high values. The raw mean comparison therefore favors the PBL group, but the inferential strength is sensitive to analytical form.

Raw gain and normalized gain

The most striking result was the near identity of the average raw gain. The PBL group improved by 39.06 points and the control group by 39.14 points, a difference of -0.08 points when calculated as

PBL minus Control. Welch's test produced t = -0.03, p = 0.979, with a 95% confidence interval from -5.90 to 5.75 points. Hedges' g was effectively zero (-0.007). The Mann-Whitney test similarly indicated no meaningful difference in the distribution of raw gain (p = 0.854).

Thus, the data do not support an interpretation in which PBL generated a larger absolute increase in test score. Both groups gained almost exactly the same number of points on average.

Normalized gain provided a complementary view. At the group-mean level, normalized gain was 0.781 for the PBL group and 0.721 for the control group. When normalized gain was calculated individually and compared between groups, the PBL mean exceeded the control mean by approximately 0.059. Welch's test yielded p = 0.145, with a 95% confidence interval from -0.021 to 0.139 and Hedges' g = 0.38. Therefore, although the normalized-gain point estimate favored PBL, the sample did not provide sufficiently precise evidence to conclude that normalized gain differed between the groups.

Baseline-adjusted final attainment

Because the PBL group began approximately 4.31 points higher, the primary analysis adjusted post-test scores for baseline performance. In the ANCOVA model, the estimated group coefficient was 3.98 points in favor of PBL after pre-test score was included as a covariate. Conventional ANCOVA yielded F(1, 58) = 4.29, p = 0.043, with partial eta squared = 0.069. When HC3 heteroscedasticity-robust standard errors were used, the PBL coefficient remained positive and statistically detectable: 3.98 points, 95% CI 0.03 to 7.92, p = 0.048.

At the sample grand mean pre-test score of 47.95, the estimated adjusted post-test mean was 88.94 for PBL and 84.96 for Control. The adjusted difference was therefore essentially four points. The pre-test × group interaction was not significant, F = 0.073, p = 0.788, providing no evidence that the baseline-post-test association required different regression slopes for the two groups. Fig. 3 displays the observed pre-test and post-test relationship by group.

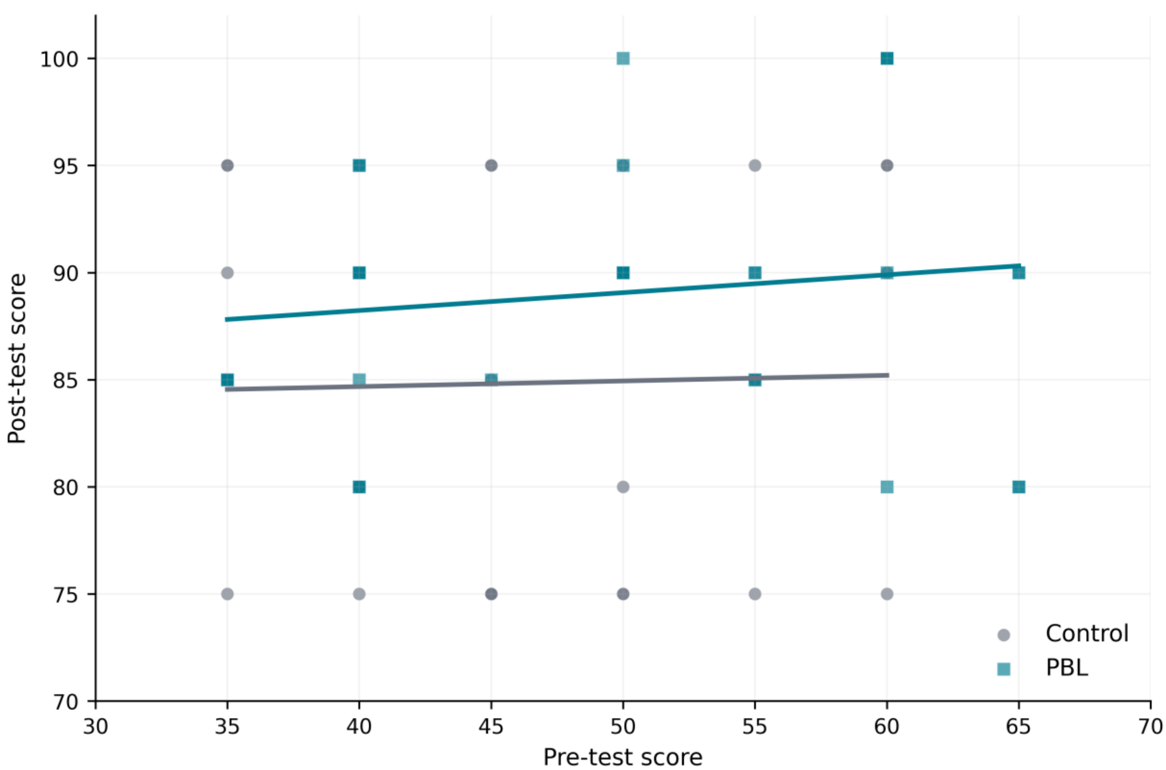

Fig. 3. Relationship between pre-test and post-test scores by group. Lines show group-specific ordinary least-squares trends for visualization; the inferential group comparison is based on ANCOVA with baseline adjustment.

The baseline covariate itself was weak in the fitted model (coefficient approximately 0.06, $p > 0.50$), indicating that within the restricted pre-test range represented in this dataset, higher baseline scores were not strongly associated with higher post-test scores after group membership was considered. This weak relationship is consistent with the substantial upward convergence of students toward high post-test values.

Integrated interpretation of the statistical evidence

Fig. 4 summarizes the central effect estimates on the original score-point scale. At baseline, the PBL group was 4.31 points higher, although the confidence interval included zero. The raw post-test difference was 4.23 points and its Welch confidence interval excluded zero. After baseline adjustment, the estimated difference remained 3.98 points; the robust confidence interval barely excluded zero. By contrast, the raw gain difference was essentially zero.

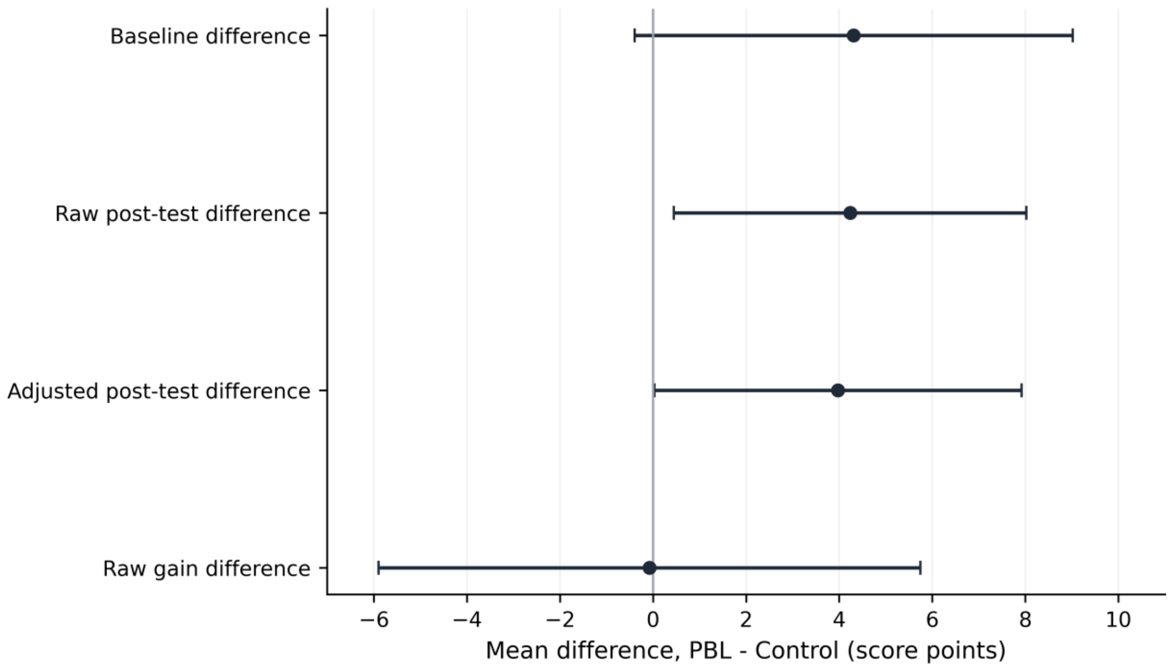

Fig. 4. Group effect estimates on the original 0–100 score scale. Positive values favor PBL. Error bars show 95% confidence intervals; the adjusted post-test estimate uses HC3 robust standard errors.

Table 2. Inferential comparison of pre-test, post-test, change, and baseline-adjusted outcomes.

Comparison	**Estimate**	**95% CI**	**Test statistic**	**p**	**Effect size**
Baseline: PBL − Control	+4.31	−0.40 to 9.02	Welch t(58.89)=1.83	.072	Hedges' g=0.46
Control: Post − Pre	+39.14	34.65 to 43.63	Paired t(28)=17.85	<.001	dz=3.31
PBL: Post − Pre	+39.06	35.17 to 42.96	Paired t(31)=20.44	<.001	dz=3.61
Post-test: PBL − Control	+4.23	0.44 to 8.03	Welch t(50.28)=2.24	.029	Hedges' g=0.58

Raw gain: PBL − Control	−0.08	−5.90 to 5.75	Welch t(56.99)=−0.03	.979	Hedges' g=−0.01
Individual n-gain: PBL − Control	+0.059	−0.021 to 0.139	Welch t(55.79)=1.48	.145	Hedges' g=0.38
Adjusted post-test: PBL − Control	+3.98	0.03 to 7.92	HC3 t(58)=2.02	.048	ηp²=0.069*

Note. All between-group estimates are PBL minus Control. HC3 = heteroscedasticity-consistent standard error. *Partial eta squared is from the conventional type-II ANCOVA (F(1,58)=4.29, p=.043).

Taken together, these results establish a pattern that cannot be captured by a single "improvement" statistic. Both groups improved greatly. The PBL group did not improve by more raw points. Nevertheless, the PBL group finished at a higher average level, and an approximately four-point advantage remained after baseline adjustment. The evidence for this endpoint advantage is positive but should be characterized as modest in precision because the sample is small, the post-test distribution is bounded and heteroscedastic, and the non-parametric post-test comparison did not cross the conventional 0.05 threshold.

Accordingly, the most defensible interpretation is not that PBL caused a larger learning gain. Rather, the observed PBL condition was associated with higher final cooling-system assessment attainment. The distinction is important because it directs the substantive discussion toward endpoint competency, scale constraints, and the educational meaning of reaching a higher final level rather than toward an unsupported claim of greater raw improvement.

Principal finding

The study produced a clear but nuanced result. Students in both the PBL and control groups learned substantially during the instructional period, as shown by average pre-to-post increases of approximately 39 points in each group. There was essentially no between-group difference in raw gain. At the same time, the PBL group achieved a higher final score, and the estimated difference remained close to four points after pre-test performance was controlled. The main contribution of the study therefore lies in distinguishing two questions that are often collapsed into one: how much students improved and what level of competency they ultimately attained.

The baseline-adjusted endpoint difference was modest (3.98 points) and should not be interpreted as a dramatic pedagogical effect, particularly because its robust confidence interval was close to the null boundary. Nevertheless, because this difference remained after baseline adjustment while the raw-gain contrast was essentially zero, the findings support a narrower claim of higher final attainment associated with PBL rather than a larger learning gain.

This distinction matters because a treatment can be associated with a higher final level without producing a larger raw change. In the present data, the PBL group began somewhat higher and ended somewhat higher, while both groups moved upward by nearly the same absolute amount. A simplistic interpretation of the paired tests would conclude that each condition was highly effective. A simplistic interpretation of raw gain would conclude that PBL offered no advantage. A simplistic interpretation of post-test means alone would conclude that PBL was superior. None of these isolated statements adequately represents the data. The combined evidence indicates strong general learning in both classes and a smaller, baseline-adjusted endpoint advantage associated with the PBL class.

Learning gain versus competency attainment

The difference between gain and attainment is particularly relevant to bounded educational scores. Raw gain is easy to interpret because it is expressed in the original score units, but it is affected by the student's starting position and by the upper limit of the scale. A student who begins at 60 can gain at most 40 points on a 100-point test, whereas a student who begins at 35 can gain 65 points. When groups have different baseline means, identical raw gains do not necessarily imply identical learning efficiency relative to the available score range.

The group-level normalized gain was 0.781 for PBL and 0.721 for Control, suggesting that the PBL group converted a somewhat larger share of its available improvement into observed post-test performance. However, the individual normalized-gain comparison was not statistically decisive. This finding should prevent the normalized-gain metric from being used as a substitute for the primary result. Instead, it helps explain why identical raw gains can coexist with different final scores.

The endpoint perspective may be educationally meaningful in vocational instruction. Occupational learning is often threshold-oriented in practice: students eventually need to demonstrate a sufficiently developed understanding to interpret a system, diagnose a fault, or perform a procedure correctly. A learner who improves substantially but remains below the level required for reliable performance may still require additional instruction. Conversely, reaching a high final level may be important even if the numerical gain is similar to that of another learner who started lower. The current dataset does not contain a validated mastery threshold, so no pass/fail inference is made. Nevertheless, the distinction between change and endpoint level remains conceptually relevant.

Why PBL could plausibly support final attainment

The observed pattern is compatible with the instructional logic of PBL, although the score data alone cannot establish a mechanism. Cooling-system problems are inherently relational. Overheating, insufficient heat rejection, abnormal coolant flow, pressure loss, thermostat malfunction, fan problems, or circulation failure cannot be understood reliably through isolated component recall. Diagnosis requires the learner to integrate component function, system sequence, operating conditions, symptoms, and evidence.

A problem-centered learning environment can potentially organize this knowledge around causal structures. Students confronted with a realistic technical problem must determine which observations matter, identify possible explanations, compare alternatives, and justify a course of action. Such activity may support the formation of integrated mental models in which a component is understood through its contribution to system behavior. If this occurs, the benefit may become visible in final task performance even when raw score gain is not dramatically larger.

Several mechanisms make the endpoint advantage plausible without implying that the present dataset identifies them. Engineering students in systemic PBL describe learner agency through team dynamism, peer support, self-directed learning, project management, and career effort (Lyngdorf et al., 2023), while longitudinal evidence shows increasing use of metacognitive and self-directed strategies in a PBL curriculum (Marra et al., 2022). PBL teamwork also involves complementary competence, team culture, conflict management, and shared situational awareness (Boelt et al., 2023). Real-life problems can support engineering competence beliefs and professional identity (Chen et al., 2022), while facilitation and cooperative interaction support engagement and group performance (Čubela et al., 2023; Havenga & Swart, 2022). Vocationally oriented PBL has additionally been associated with deeper understanding of problem solving, professional knowledge, and employability potential (Miliou et al., 2022).

However, this explanation remains an interpretation rather than a directly observed causal pathway in the present study. The dataset contains scores and group labels but does not include process measures such as student dialogue, diagnostic reasoning sequences, engagement, time on task, quality of problem representation, or fidelity of PBL implementation. The study therefore cannot determine whether the final-score difference arose from collaborative reasoning, increased engagement, contextualized practice, teacher facilitation, peer explanation, or another feature of the classroom condition. Future work should measure these processes explicitly if the mechanism of PBL is a central research objective.

The importance of the control group

One of the strongest lessons from the data is the importance of retaining the control group in the interpretation. The PBL group's pre-to-post improvement is statistically very large. If the analysis had included only that class, the result could easily have been presented as compelling evidence that PBL dramatically improved learning. The control group reveals that such a conclusion would be incomplete because its students improved by almost exactly the same raw number of points.

This observation does not mean that the control condition was equivalent to PBL in every educational sense. It means only that the available score metric does not show a difference in absolute gain. The two instructional conditions may differ in final attainment, dispersion, learning processes, retention, transfer, practical performance, or student experience. Some of these dimensions are partially visible in the post-test distribution, while others are not represented in the current dataset. The control group therefore serves not merely as a statistical comparator but as a safeguard against attributing all learning that occurs during a course to the experimental pedagogy.

The result also illustrates why "significant in one group" versus "not significant in another group" is not a valid test of an intervention difference. In this dataset, both groups are significant within themselves, but even if only one paired test had reached significance, the correct comparison would still be the difference in gains or an appropriately adjusted model. The study's analytical contribution is therefore strengthened by reporting the direct group contrast rather than relying on within-group p values.

Baseline adjustment and the quasi-experimental interpretation

The PBL group started 4.31 points higher on average. Although the baseline Welch test yielded $p = 0.072$, the absence of conventional statistical significance should not be interpreted as proof that the groups were equivalent. Baseline significance testing is especially limited in modest samples because a practically relevant difference may fail to cross an arbitrary p-value threshold. In a quasi-experimental design, the baseline score is available precisely to help account for this imbalance.

ANCOVA indicated that the estimated PBL advantage at post-test was 3.98 points after adjustment for pre-test score. This is close to the unadjusted post-test difference of 4.23 points, which shows that baseline adjustment changed the estimate only modestly. The interaction test provided no evidence that the relationship between pre-test and post-test differed substantially by group. These findings support the use of a common-slope model.

At the same time, the strength of the inference should not be overstated. The robust confidence interval for the adjusted group effect extends from approximately 0.03 to 7.92 points and therefore lies close to the null boundary. The post-test Mann-Whitney sensitivity test produced $p = 0.065$. These results do not erase the positive estimate, but they show that the exact inferential conclusion is sensitive to distributional assumptions and analytical method. The appropriate scientific statement is that the data provide evidence consistent with a modest PBL-associated endpoint advantage, not that they establish an unequivocal or large causal effect.

The result should therefore be interpreted against a literature base that is positive but not uniformly comparative. Some studies evaluate combined pedagogies or perceptions rather than isolated PBL effects. A structural-engineering comparison, for example, found the highest achievement under project- and technology-based assignments rather than PBL alone (Ghoniem & Ghoniem, 2022), while material-balances research embeds PBL within a broader active-learning design (Ortiz Moreno et al., 2025). A review of PBL in engineering education for sustainability similarly observed that assessment often remains content- and score-based despite broader competence claims (D'Escoffier et al., 2024). These studies reinforce the need to state precisely what the present data show: strong learning in both groups, no raw-gain advantage, and a modest PBL-associated advantage in final assessed attainment.

Score distribution, variance, and possible ceiling compression

The post-test distributions differ not only in their means but also in their shapes and spreads. All PBL students scored at least 80, whereas the control post-test range extended down to 75. The PBL distribution reached 100, while the control maximum was 95. The PBL post-test standard deviation was 6.02 compared with 8.40 in Control, and Levene's test detected unequal post-test variances.

One possible interpretation is that PBL was associated with more consistent final attainment. A lower standard deviation at a higher mean can indicate that students converged toward a common level of performance. This pattern may be attractive in vocational education because instructional success concerns not only raising the average but also limiting the number of learners who remain substantially behind. However, the evidence should be treated cautiously because the score scale is discrete and bounded. As more students approach the maximum, variance can contract mechanically.

Ceiling compression is particularly relevant here. Four PBL students reached a score of 100, leaving no room for further observed improvement on the scale. Students beginning at 60 or 65 also had less potential numerical gain than students beginning at 35 or 40. These features can attenuate raw gain differences and reduce the usefulness of normal-theory assumptions. The use of robust standard errors and non-parametric sensitivity tests was therefore appropriate, but a future study would benefit from an assessment instrument with enough difficulty and score resolution to discriminate performance at the upper end.

Educational implications

For teaching practice, the findings suggest that PBL can be considered a promising approach for cooling-system instruction when the goal is to support high final attainment. The result should not be translated into the claim that simply labeling a lesson "PBL" will produce a four-point increase. The classroom condition is a package that may include problem design, facilitation, group interaction, time allocation, feedback, prior preparation, and assessment alignment. Effective implementation therefore depends on how the problem is structured and how the learning process is guided.

The literature also warns against assuming that every PBL implementation produces the same learning profile. Technical-vocational studies emphasize authentic, repeated, feedback-rich practice but identify resource constraints, uneven readiness, and monitoring problems (Bascuguin & Generoso, 2026). Technical-education entrants and hybrid PBL cases suggest that learners need scaffolding as responsibility shifts away from teacher-centered routines (Dabir et al., 2022; Lou et al., 2022). TVET research highlights facilities and lecturer competence, while competency-based models stress alignment among learning content, delivery, media, and assessment (Liangyu et al., 2024; Rahman et al., 2022). Engineering instructors likewise view PBL as promising but implementation remains an instructional-design problem (Herrera-Perez et al., 2026). Multidisciplinary cases reinforce the need to align tasks, disciplinary integration, and intended outcomes (Syed et al., 2024).

A technically useful PBL task for cooling-system learning would ideally require students to move from symptoms to hypotheses and from hypotheses to evidence. For example, a scenario involving engine overheating could require learners to differentiate restricted coolant circulation, insufficient airflow, thermostat malfunction, pressure loss, or abnormal heat load. The educational value lies in requiring a reasoned diagnostic pathway rather than allowing students to select an answer through recall alone. Although the present dataset does not document the exact problem scenarios used, this type of structure illustrates the mechanism that future implementations can deliberately test.

The findings also imply that evaluation should be aligned with the intended competency. If the instructional goal is diagnostic reasoning, an assessment dominated by factual recall may underestimate the effect of problem-centered learning. Conversely, if the test measures only final answers without capturing reasoning quality, it may not show how students reached those answers. Future assessments could combine objective knowledge items with scenario-based diagnosis, explanation of causal chains, selection of inspection steps, and practical performance. Such multidimensional assessment would make the term “competency” more defensible than reliance on a single total score.

Implications for research design

The current analysis provides several design lessons for subsequent studies. First, intact-class comparisons should collect more baseline information than a single pre-test score. Relevant variables may include prior diesel-engine coursework, general academic achievement, attendance, practical workshop experience, or other indicators related to technical learning. These variables can help determine whether the groups differ before the intervention in ways that could influence the outcome.

Second, future designs should measure implementation fidelity. PBL is not a single exposure with a fixed dose. Its effects can depend on the authenticity of the problem, the extent of student responsibility, the quality of facilitation, opportunities for reflection, and the way evidence is used during problem solving. Without implementation data, a positive or null result cannot reveal which features of the pedagogy were actually responsible.

Third, repeated or delayed assessments would extend the present endpoint analysis. A post-test immediately after instruction measures short-term attainment, whereas a delayed test can evaluate retention. Practical transfer tasks can assess whether students use the acquired knowledge in unfamiliar diagnostic situations. These outcomes may be particularly important for vocational education, where the ultimate objective is not only test performance but reliable performance in realistic technical contexts.

Fourth, larger multi-class or multi-school samples would improve precision and allow the class-level nature of the intervention to be modeled explicitly. In the present dataset, treatment condition is confounded with class: XI TAB 1 is the PBL class and XI TAB 2 is the control class. Any unobserved difference between the classes could therefore contribute to the estimated treatment contrast. Replicating the design across several classes, ideally with random assignment of classes or instructors to conditions where feasible, would substantially strengthen causal inference.

Strengths and limitations

A strength of the study is the availability of complete paired observations for both groups. The dataset permits direct examination of baseline, endpoint, and within-student change rather than relying on post-test scores alone. A second strength is the analytical separation of within-group improvement from between-group effect. This prevents the very large paired effect sizes from being misinterpreted as evidence of treatment superiority. A third strength is the use of baseline adjustment and robust inference, which is particularly relevant given unequal post-test variances and the modest baseline imbalance.

The study also has important limitations. The first is the quasi-experimental structure based on intact classes. Because students were not represented as individually randomized in the available records, unmeasured class differences may confound the group effect. The second is that group and class are perfectly aligned, making it impossible to separate a PBL effect from other class-specific influences with this dataset alone. With one class per condition, treatment and class are inseparable at the design level; adjustment for individual pre-test score cannot resolve this class-level confounding.

Third, the outcome is a total assessment score without item-level information. The available records do not establish which dimensions of cooling-system competence were measured, how the items were developed, or the reliability and validity of the instrument. The term “competency” should therefore be interpreted as performance on the observed cooling-system assessment unless additional instrument documentation demonstrates broader practical competence. The observed five-point score spacing does not, by itself, establish whether the original assessment used 20 multiple-choice items, a rubric, or another scoring rule.

Fourth, the control condition is labeled as Control but is not pedagogically characterized in the score file. The estimated difference is therefore a comparison between the recorded PBL class and the recorded control class, not a precisely isolated contrast between PBL and a fully documented alternative teaching method. Fifth, the sample size is modest, which contributes to relatively wide confidence intervals and borderline inferential results for the adjusted group effect.

Sixth, the post-test distribution is bounded and concentrated near the upper end, particularly in the PBL group. This creates potential ceiling compression, reduces score resolution, and helps explain why parametric and rank-based tests do not yield identical significance decisions. Seventh, there is no delayed retention test, no practical transfer task, and no measure of problem-solving process, so the study cannot determine whether the observed endpoint advantage persists over time or generalizes to hands-on diagnosis.

These limitations define the appropriate boundary of the conclusion. The data support a modest baseline-adjusted difference in final assessment attainment associated with the PBL class. They do not support the stronger claim that PBL produced a larger absolute gain, nor do they prove that PBL alone caused the difference.

Core contribution

The central contribution of this study is therefore not a generic confirmation that PBL “works.” The more informative contribution is the empirical separation of two forms of instructional success. Both classes displayed large learning gains, showing that structured instruction was effective in moving students far above their starting scores. The PBL class, however, finished at a higher level, and that endpoint difference remained after baseline adjustment. The distinction reveals that the educational value of PBL in this dataset is better described as an advantage in final attainment than as an advantage in the number of raw points gained.

This framing extends prior PBL research by distinguishing score change from endpoint performance in an intact-class vocational setting. Methodologically, it shows that conclusions about instructional effectiveness may differ depending on whether researchers examine within-group change, between-group gain, or baseline-adjusted final attainment. Future studies should test whether this distinction persists in practical diagnostic performance, delayed retention, and multi-class designs. Accordingly, the adjusted endpoint contrast should be read as an effect estimate for these two observed classes rather than as a cluster-replicated causal effect of PBL.

CONCLUSION

This quasi-experimental pretest-posttest analysis examined cooling-system assessment outcomes among 61 Grade XI Heavy Equipment Engineering students, including 32 students in the class identified as receiving Problem-Based Learning and 29 students in the control class. Both groups showed substantial improvement from pre-test to post-test, with average gains of approximately 39 points. The raw gain difference between groups was essentially zero, demonstrating that the available data do not support the claim that PBL generated a larger absolute increase in score.

The more informative result concerned final attainment. The PBL group achieved a post-test mean of 89.06 compared with 84.83 in the control group. The unadjusted difference was 4.23 points, and the baseline-adjusted difference remained approximately 3.98 points. Robust inference indicated that this adjusted advantage was positive but estimated with limited precision. The overall evidence is therefore consistent with a modest PBL-associated advantage in final cooling-system attainment rather than a greater raw learning gain.

This distinction should shape both the interpretation and future evaluation of vocational pedagogy. Instructional effectiveness is not fully represented by whether students improve, because improvement may occur across multiple teaching conditions. Evaluations should also ask what level of technical understanding students reach, whether that level persists after baseline adjustment, and whether the observed attainment transfers to diagnostic and practical performance. Within the limits of the present intact-class dataset, PBL emerges as a promising instructional condition for supporting higher final cooling-system assessment performance, while stronger causal conclusions require replication with better-controlled assignment, richer competency measures, documented instructional fidelity, and delayed or practical outcomes.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the Teacher Professional Education Program, Graduate School, State University of Malang, and SMK Negeri 6 Malang for their support in facilitating the learning activities and data collection. The authors also thank the participating students and teaching staff for their cooperation throughout the study.

REFERENCES

Alaloul, W. S., & Qureshi, A. H. (2022). Cooperative problem-based learning experience and coaching strategies of engineering course. *International Journal of Evaluation and Research in Education*, *11*(2), 848–861. https://doi.org/10.11591/ijere.v11i2.22243

Andersen, A.-L., & Rösiö, C. (2023). Continuing Engineering Education in Changeable and Reconfigurable Manufacturing – Implications of Problem-Based Learning in Industrial Practice*. *International Journal of Engineering Education*, *39*(5), 1118–1130. https://www.ijee.ie/contents/c390523.html

Ariza, J. Á. (2022). Can In-Home Laboratories Foster Learning, Self-Efficacy, and Motivation during the COVID-19 Pandemic? - A Case Study in Two Engineering Programs. *International Journal of Engineering Education*, *38*(2), 310–321. https://www.ijee.ie/contents/c380222.html

Bascuguin, C. B., & Generoso, E. T. (2026). Pedagogical flexibility and adaptive teaching strategies for practical skills mastery in computer systems servicing: A mixed-methods brief report from Batangas, Philippines. *Frontiers in Education*, *11*. https://doi.org/10.3389/feduc.2026.1813666

Bilbao, J., Bravo, E., García, O., & Rebollar, C. (2024). Recovering autonomous work after the pandemic: Analysis in Calculus for incoming Students in Technical Education degrees. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-04263-z

Bin Mamat, A. M. I. (2026). Transforming turbomachinery education: A constructivist approach with an interactive PBL dashboard. *International Journal of Mechanical Engineering Education*. https://doi.org/10.1177/03064190251414098

Boelt, A. M., Holgaard, J. E., & Kolmos, A. (2023). A Thematic Analysis of Engineering Students' Experiences of Teamwork in Problem-Based Learning. *International Journal of Engineering Education*, *39*(3), 627–642. https://www.ijee.ie/1atestissues/Vol39-3/13_ijee4338.pdf

Chen, J., Du, X., & Kolmos, A. (2022). Students' Views on Sources of Engineering Identity Development in a Collaborative PBL Environment. *International Journal of Engineering Education*, *38*(2), 525–542. https://www.ijee.ie/1atestissues/Vol38-2/21_ijee4181.pdf

Chen, J., Kolmos, A., & Du, X. (2021). Forms of implementation and challenges of PBL in engineering education: A review of literature. *European Journal of Engineering Education*, *46*(1), 90–115. https://doi.org/10.1080/03043797.2020.1718615

Čubela, D., Rossner, A., & Neis, P. (2023). Using Problem-Based Learning and Gamification as a Catalyst for Student Engagement in Data-Driven Engineering Education: A Report. *Education Sciences*, *13*(12). https://doi.org/10.3390/educsci13121223

Dabir, V., Kulkarni, A., & Khare, K. (2022). Perceiving the Suitability of Adopting Problem Based Learning Pedagogy for Technical Education Entrants. *Journal of Engineering Education Transformations*, *36*(1), 33–39. https://doi.org/10.16920/jeet/2022/v36i1/22134

Das, D. K. (2026). Enhancing student performance and competency in undergraduate engineering education: A holistic pedagogical framework development. *Cogent Education*, *13*(1). https://doi.org/10.1080/2331186X.2026.2694109

Devika, S. V., Siddapuram, A., Kaur, R., & Bollampally, A. (2024). From Lecture-Based Learning to Problem-Based Learning: A Review on Navigating the Transformation in Engineering Education. *Journal of Engineering Education Transformations*, *38*(Special Issue 1), 179–183. https://doi.org/10.16920/jeet/2024/v38is1/24229

Devika, S. V., Siddapuram, A., Kaur, R., & Kondala Rao, P. (2024). Enhancing Program Outcomes and Academic Performance through Problem-Based Learning in an Antennas and Wave Propagation Course. *Journal of Engineering Education Transformations*, *38*(Special Issue 1), 140–145. https://doi.org/10.16920/jeet/2024/v38is1/24223

D'Escoffier, L. N., Guerra, A., & Braga, M. (2024). Problem-Based Learning and Engineering Education for Sustainability Where we are and where could we go? Journal of Problem Based Learning in Higher Education, 12(1), 18–45. *https://doi.org/10.54337/ojs.jpblhe.v12i1.7799*.

Farooq, M. U., & Jang, Y. J. (2026). Advancing Manufacturing Process Education for Industry 4.0: A Multi-Case Study. *Computer Applications in Engineering Education*, *34*(1). https://doi.org/10.1002/cae.70136

Ghoniem, A., & Ghoniem, E. (2022). Inducing competence-based assignment in traditional structural engineering education: A case study. *Computer Applications in Engineering Education*, *30*(3), 907–916. https://doi.org/10.1002/cae.22493

Gómez-Coma, L., Díaz-Sainz, G., Fallanza, M., Ortiz, A., & Ortiz, I. (2023). Integration of chemical engineering skills in the curriculum of a master course in industrial engineering. *Education for Chemical Engineers*, *45*, 68–79. https://doi.org/10.1016/j.ece.2023.08.002

Havenga, M., & Swart, A. J. (2022). Preparing first-year engineering students for cooperation in real-world projects. *European Journal of Engineering Education*, *47*(4), 558–576. https://doi.org/10.1080/03043797.2022.2037521

Herrera-Perez, J., Falcón-Mendoza, A., Granizo-Pumagualli, C., & Herrera-Perez, V. (2026). Teachers' perspectives on problem based learning in engineering leveling courses teaching vector physics and statics. *Discover Education*, *5*(1). https://doi.org/10.1007/s44217-026-01867-2

Jeyanathan, J. S., Nagaraj, P., & Anitha, P. (2026). Bridging Theory and Practice: Assessing Project and Problem-Based Learning (P2BL) in Undergraduate IoT Course. *Journal of Engineering Education Transformations*, *39*(Special Issue 2), 554–561. https://doi.org/10.16920/jeet/2026/v39is2/26066

Katageri, B. G., & Raikar, R. V. (2022). Problem-Based Learning Experience in Civil Engineering. *Journal of Engineering Education Transformations*, *35*(3), 143–151. https://doi.org/10.16920/jeet/2022/v35i3/22097

Liangyu, Y., Rahman, K. A. A., & Jurime, N. H. (2024). Project-Oriented Problem-Based Learning and Innovation within TVET Ruminant Education. *Journal of Technical Education and Training*, *16*(3), 101–113. https://doi.org/10.30880/jtet.2024.16.03.008

Long, J., Dragich, E., & Saterbak, A. (2022). Problem-Based Learning Impacts Students' Reported Learning and Confidence in an Undergraduate Biomedical Engineering Course. *Biomedical Engineering Education*, *2*(2), 209–232. https://doi.org/10.1007/s43683-022-00067-2

Lou, S.-J., Huang, C.-Y., Cheng, Y.-M., & Chung, C.-C. (2022). Hybrid PBL Teaching Practice under COVID-19 Impact—A Case Study. *International Journal of Engineering Education*, *38*(2), 437–451. https://www.ijee.ie/contents/c380222.html

Lyngdorf, N. E. R., Du, X., & Lundberg, A. (2023). First-year engineering students' learner agency sources in a systemic PBL environment: A Q study. *European Journal of Engineering Education*, *48*(6), 1130–1147. https://doi.org/10.1080/03043797.2023.2233427

Maksum, H., Rizal, F., & Ulya, R. H. (2023). Problem-Based Blended Learning Models in Vocational Education: A Developmental Research. *International Journal of Information and Education Technology*, *13*(12), 1968–1976. https://doi.org/10.18178/ijiet.2023.13.12.2011

Maksum, H., Wagino, W., Purwanto, W., Muslim, M., Putra, R., & Koto, R. D. (2025). Quantifying the Impact of Problem-Based Learning Enhanced by E-Learning: An Advanced Statistical and Effect Size Approach in Vocational Studies. *TEM Journal*, *14*(3), 2744–2754. https://doi.org/10.18421/TEM143-75

Mann, L., Chang, R., Chandrasekaran, S., Coddington, A., Daniel, S., Cook, E., Crossin, E., Cosson, B., Turner, J., Mazzurco, A., Dohaney, J., O'Hanlon, T., Pickering, J., Walker, S., Maclean, F., & Smith, T. D. (2021). From problem-based learning to practice-based education: A framework for shaping future engineers. *European Journal of Engineering Education*, *46*(1), 27–47. https://doi.org/10.1080/03043797.2019.1708867

Marra, R. M., Hacker, D. J., & Plumb, C. (2022). Metacognition and the development of self-directed learning in a problem-based engineering curriculum. *Journal of Engineering Education*, *111*(1), 137–161. https://doi.org/10.1002/jee.20437

Miliou, O., Ioannou, A., Georgiou, Y., Vyrides, I., Xekoukoulotakis, N., Willert, S., Andreou, A., Andreou, P., Komnitsas, K., Zaphiris, P., & Yiatros, S. (2022). The Design of a Postgraduate Vocational Training Programme to Enhance Engineering Graduates' Problem-Solving Skills Through PBL. *International Journal of Engineering Education*, *38*(5 A), 1257–1273. https://doi.org/10.5281/zenodo.7589566

Morales, F. J., Aranda, M. T., & Romero, L. M. (2026). Empowering students through practical learning: A hands-on approach to brake systems. *International Journal of Mechanical Engineering Education*. https://doi.org/10.1177/03064190261418726

Naik, S. M., Bandi, S., & Reddy, C. L. (2024). Enhancing Engineering Education: Exploring the Impact of Problem-Based Learning on Freshman Students' Skills and Engagement – A Case Study. *Journal of Engineering Education Transformations*, *38*(Special Issue 1), 133–139. https://doi.org/10.16920/jeet/2024/v38is1/24222

Naveh, G., Bakun-Mazor, D., Tavor, D., & Shelef, A. (2022). Problem-based Learning in A Theoretical Course in Civil Engineering: Students' Perspectives. *Advances in Engineering Education*, *10*(3), 46–67. https://doi.org/10.18260/3-1-1153-36033

Okolie, U. C., Elom, E. N., Igwe, P. A., Binuomote, M. O., Nwajiuba, C. A., & Igu, N. C. N. (2021). Improving graduate outcomes: Implementation of problem-based learning in TVET systems of Nigerian higher education. *Higher Education, Skills and Work-Based Learning*, *11*(1), 92–110. https://doi.org/10.1108/HESWBL-12-2018-0140

Okolie, U. C., Ogwu, E. N., Osuji, C. U., Ogba, F. N., Igwe, P. A., & Obih, S. O. (2023). A critical perspective on TVET teachers' pedagogical practices: Insights into the guiding pedagogical principles in practice. *Journal of Vocational Education and Training*, *75*(3), 439–458. https://doi.org/10.1080/13636820.2021.1894221

Orji, C. T. (2025b). Problem-based Intervention via Task-based Scenarios Analysis of Practical Skills Acquisition Relations in Vocational Electrical Technology. *Journal of Problem Based Learning in Higher Education*, *13*(1), 75–94. https://doi.org/10.54337/ojs.jpblhe.v13i1.9979

Ortiz Moreno, H., Godoy-Rangel, C., & Ahumada-Lazo, R. (2025). From lectures to active learning: A case study on problem-based learning in the material balances course. *International Journal of Mechanical Engineering Education*. https://doi.org/10.1177/03064190251356618

Patnawar, S. T. (2023). A Comprehensive Review on PBL and Digital PBL in Engineering Education—Status, Challenges and Future Prospects. *Journal of Engineering Education Transformations*, *37*(2), 142–157. https://doi.org/10.16920/jeet/2023/v37i2/23157

Rafiq, A. A., Triyono, M. B., & Djatmiko, I. W. (2023). The Integration of Inquiry and Problem-Based Learning and Its Impact on Increasing the Vocational Student Involvement. *International Journal of Instruction*, *16*(1), 659–684. https://doi.org/10.29333/iji.2023.16137a

Rahman, K. A. A., Rozali, M. Z., Samah, N. A., Bakar, M. A., Ahmad, N. A., Gerijih, D. D., & Zakariah, S. H. (2022). Conceptual Model of Video Learning based on Project-Oriented Problem-Based Learning and Competency-Based Education for Technical and Vocational Education. *Journal of Technical Education and Training*, *14*(1), 38–53. https://doi.org/10.30880/jtet.2022.14.01.004

Rodriguez-Sanchez, C., Orellana, R., Fernandez Barbosa, P. R., Borromeo, S., & Vaquero, J. (2024). Insights 4.0: Transformative learning in industrial engineering through problem-based learning and project-based learning. *Computer Applications in Engineering Education*, *32*(4). https://doi.org/10.1002/cae.22736

Sanchez-Gomez, C. A. (2022). Implementing a joint learning method (PBL and EBL) to innovate the development of mechanical engineering technical and non-technical skills. *International Journal of Mechanical Engineering Education*, *50*(1), 176–196. https://doi.org/10.1177/0306419020950751

She-Ping, T., Li-Juan, W., Zhang, F., & Tao, H. A. N. (2023). Application of Problem-Based Learning to Promote Critical Thinking Disposition Among Engineering Undergraduates*. *International Journal of Engineering Education*, *39*(4), 877–885. https://www.ijee.ie/contents/c390423.html

Syed, K., Vijayalakshmi, M., Kandakatla, R., & Narayanan, S. (2024). Inculcating Multidisciplinary Learning in Electrical Engineering Undergraduate Curriculum through Problem-Based Learning. *Journal of Engineering Education Transformations*, *38*(Special Issue 1), 210–215. https://doi.org/10.16920/jeet/2024/v38is1/24234

Tan, S., & Shen, Z. (2022). Student Experiences with Problem-Based Learning: Findings from an Electrical Engineering Course. *International Journal of Engineering Education*, *38*(4), 1044–1055. https://www.ijee.ie/contents/c380422.html

Wagino, W., Maksum, H., Purwanto, W., Ambiyar, A., Samala, A. D., Criollo-C, S., & Koto, R. D. (2025). Assessing student expectations, perceived quality, value, and satisfaction through PBL and e-learning integration in vocational education. *Journal of Educators Online*, *22*(4). https://doi.org/10.9743/JEO.2025.22.4.15

Wang, J., & Somasundram, P. (2025). Effect of STEM-PBL advanced mathematics course on engineering students' problem-solving ability in higher vocational college. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1628482

Wulansari, R. E., Fortuna, A., Marta, R., Primawati, P., Masek, A., Kaya, D., Prasetya, F., Sakti, R. H., Luthfi, A., Rizki, I. A., Hasna, F. H., & Eshaghi, S. (2024). Revolutionizing Learning: Unleashing the Power of Technology Gamification-Augmented Reality in Vocational Education. *TEM Journal*, *13*(3), 2384–2397. https://doi.org/10.18421/TEM133-65

Zhao, Y., Dong, H., & Zhang, X. (2026). Fostering Technical and Sustainability Competencies Through an Integrated PBL Approach in an Undergraduate Mechanical Vibration Course. *Sustainability (Switzerland)*, *18*(5). https://doi.org/10.3390/su18052660