Jurnal Pendidikan dan Pembelajaran Indonesia (JPPI)
p-ISSN: 2797-2879, e-ISSN: 2797-2860
Volume 6, nomor 4, 2026, hal. 3342-3359
Doi: https://doi.org/10.53299/jppi.v6i4.5682

Paired Learning Gains and Assessment Ceilings in Automotive Electrical Vocational Education

Muhamad Akbar Nur Murtadlo*, Erwin Komara Mindarta, Dani Irawan, Komarudin
Pendidikan Teknik Otomotif, Departemen Teknik Mesin dan Industri, Fakultas Teknik,
Universitas Negeri Malang, Malang, Indonesia

*Corresponding Author: murtadloakbar@gmail.com
Dikirim: 10-10-2026; Direvisi: 20-10-2026; Diterima: 22-10-2026

Abstract: Automotive electrical-system education requires an understanding of circuit principles and component functions, yet evaluations based on class averages can obscure individual differences and assessment limitations. This study examined changes in test performance before and after instruction recorded as problem-based learning in a headlamp and starter-system unit for Grade XI vocational automotive students. A one-group design was used to analyse 34 matched pre-test and post-test scores on a 0-to-100 scale. Descriptive statistics, paired *t* tests, Wilcoxon signed-rank tests, an exact sign test, standardized mean change, normalized gains, bootstrap confidence intervals, and exploratory analyses of baseline scores and gains were conducted. Mean scores increased from 71.91 to 95.88, representing a gain of 23.97 points. Thirty-two students improved, while two recorded lower post-test scores. All paired-change tests indicated statistically significant differences, with *p* values below 0.001, and the standardized within-student effect size was 1.09. An inverse association between baseline performance and arithmetic gain requires caution because the measures are mathematically linked. Seventeen students attained the maximum post-test score, indicating substantial ceiling concentration that may limit the assessment's discrimination at higher achievement levels. The results document a large, widespread improvement in recorded scores following the instructional unit. However, the absence of a comparison group prevents causal attribution of the improvement to problem-based learning. Student-level analysis and scrutiny of assessment properties offer a more informative basis for evaluating learning gains in vocational education.

Keywords: problem-based learning; automotive electrical systems; vocational education; pre-test–post-test; learning gains.

Abstrak: Pembelajaran kelistrikan otomotif menuntut penguasaan konsep rangkaian dan fungsi komponen, tetapi evaluasi berbasis nilai rata-rata sering mengabaikan variasi peningkatan antarsiswa dan keterbatasan pengukuran. Penelitian ini bertujuan menganalisis perubahan nilai sebelum dan sesudah pembelajaran yang tercatat menggunakan problem-based learning pada materi sistem lampu kepala dan sistem starter di kelas XI Teknik Kendaraan Ringan. Penelitian menggunakan desain satu kelompok dengan 34 pasangan nilai pre-test dan post-test pada skala 0 sampai 100. Analisis mencakup statistik deskriptif, uji *t* berpasangan, Wilcoxon, uji tanda, ukuran efek, peningkatan ternormalisasi, interval kepercayaan bootstrap, dan eksplorasi hubungan nilai awal dengan peningkatan. Rata-rata nilai meningkat dari 71,91 menjadi 95,88, atau bertambah 23,97 poin. Sebanyak 32 siswa mengalami peningkatan, sedangkan dua siswa mengalami penurunan. Seluruh uji perubahan berpasangan menunjukkan perbedaan yang signifikan dengan nilai *p* kurang dari 0,001; ukuran efek berpasangan sebesar 1,09. Hubungan negatif antara nilai awal dan peningkatan harus ditafsirkan secara hati-hati karena keterkaitan matematis kedua ukuran. Sebanyak 17 siswa mencapai nilai maksimum pada post-test, sehingga muncul indikasi efek plafon yang membatasi daya pembeda instrumen. Temuan menunjukkan peningkatan hasil tes yang besar dan meluas setelah pembelajaran, tetapi desain tanpa kelompok pembanding tidak

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

memungkinkan penentuan pengaruh kausal PBL. Penilaian individu dan evaluasi karakteristik instrumen perlu menjadi bagian dari kajian hasil belajar pendidikan vokasi.

Kata Kunci: pembelajaran berbasis masalah; kelistrikan otomotif; pendidikan vokasi; pre-test post-test; peningkatan hasil belajar.

INTRODUCTION

Rationale and vocational context

Automotive electrical maintenance requires learners to integrate circuit principles with component functions and diagnosis. A malfunctioning headlamp or starter demands recognition of current paths, switches, protective elements and component interactions. Yet classroom test scores provide evidence of assessed knowledge, not necessarily practical troubleshooting competence. Distinguishing these outcomes is essential when evaluating instruction in vocational education.

Headlamp and starter systems offer contrasting applications of electrical principles, including illumination circuits, relays, solenoids and mechanical cranking. An achievement test may assess component knowledge or circuit interpretation, but the actual construct depends on item design, scoring and validity evidence. Because item-level specifications were unavailable, the outcome in this study is described as recorded test performance rather than demonstrated workplace competency.

Problem-based learning (PBL) organizes inquiry around problems whose solutions require students to identify knowledge gaps, test explanations and revise their reasoning. Diagnosing a headlamp or starter fault could support such inquiry. However, the available teaching record identifies PBL without documenting task sequence, duration or implementation fidelity. The analysis therefore examines performance following instruction designated as PBL; it does not establish the effectiveness of any specific PBL activity.

Engineering PBL is not a single, uniform instructional package. Chen et al. (2021) describe multiple implementation forms and curricular challenges. Tan and Shen (2022) examine student experiences in electrical engineering, while Servant-Miklos and Kolmos (2022) identify different student conceptions of problem- and project-based learning. In automotive vocational education, Wagino et al. (2025) investigate PBL combined with e-learning through student perceptions and satisfaction, a different outcome from paired electrical-system test scores.

Evidence from adjacent technical settings is heterogeneous. Rafiq et al. (2023) examine inquiry–PBL integration among vocational students; Orji (2025) studies engagement and practical skills in electrical/electronic technology education; Awwad (2025) evaluates active-learning approaches in electronics; and Akbar et al. (2026) report a virtual-reality computer-systems intervention. Their distinct treatments, participants and outcomes preclude direct replication claims for the present class.

Research problem and analytical gap

Pre- and post-tests make a class-average shift easy to report, but a mean conceals differences among learners and may be distorted by score saturation. A useful paired evaluation should distinguish magnitude, precision, directional consistency and the capacity of an instrument to discriminate at high attainment. Scores near the maximum can compress the measured distribution without proving that competence has become uniform.

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

Baseline–gain associations also require caution. The gain is calculated by subtracting the baseline score from the final score, so a negative association can arise through mathematical coupling. A 0–100 scale further gives high-baseline learners less possible upward movement. These features, together with possible regression to the mean, prevent a baseline–gain correlation from being interpreted as proof that one subgroup benefits more from PBL.

A one-group pre-test–post-test design cannot separate instruction from other changes over time. Test familiarity, concurrent practice, maturation, differing test difficulty or administration may all contribute to the observed contrast. Paired tests establish evidence of differences between measurements; they do not estimate what would have happened under an alternative teaching method.

Broader syntheses motivate this classroom analysis but do not supply its missing counterfactual. Affandy et al. (2026) synthesize PBL in science education, Erdem (2024) examines cognitive outcomes across K–12 studies, Ahmad et al. (2023) compares chemistry pedagogies, and Uluçinar (2023) meta-analyses science achievement. General findings from these populations should not be transferred uncritically to automotive electrical instruction.

Context-specific evidence is equally important. Kanyesigye et al. (2023) investigate mechanical waves, Aristin et al. (2023) study geography, Arici and Yilmaz (2023) add augmented reality to PBL, and Alshaye et al. (2023) examine online programming tasks. These studies illustrate variations in settings, delivery and measured outcomes; none directly resolves the paired headlamp–starter assessment question.

Study objectives, questions and proposed contribution

This study evaluates 34 paired pre-test and post-test scores from Grade XI Light Vehicle Electrical Maintenance (XI TKR 2), covering headlamp and starter systems. The principal question is how much recorded performance changed within students and how consistently that change occurred. Secondary questions address sensitivity across statistical summaries, learner-level heterogeneity and concentration near the upper score boundary.

The contribution is a transparent account of score changes, individual differences and measurement limitations within a vocational class designated as using PBL. It is not an estimate of PBL superiority over alternative instruction. Precision about this evidential boundary strengthens the interpretation of otherwise striking numerical gains.

PBL learning analytics increasingly considers variation within and between students. Saleh et al. (2022) analyse collaborative inquiry, Saqr and López-Pernas (2023) investigate the timing of online interaction, and Mujumdar et al. (2024) use data mining to identify learning patterns. Because no process logs are available for the present classroom, only individual achievement trajectories—not group interactions or mechanisms—can be evaluated.

RESEARCH METHODS

Study design and source provenance

The study uses a one-group paired pre-test–post-test analysis of an existing class assessment record. Thirty-four students in XI TKR 2 completed the two

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

recorded assessments in Light Vehicle Electrical Maintenance, covering headlamp and starter systems. The register labels the instructional model as PBL. The available materials do not establish the school's name and location, academic year, session dates, instructional duration or student demographic profile. These contextual particulars require confirmation by the study investigators; they are not inferred from grades.

The analysis includes all 34 matched numeric pairs, with no missing assessment entries in the available register and no score imputation. Values were analysed on the recorded 0–100 scale. Repeated marks across learners were retained as legitimate observations. Only anonymized and aggregate outcomes are presented; the class register is not reproduced.

Instructional setting and exposure classification

The class record identifies PBL as the instructional model but does not specify its operational steps. Problem scenarios, group composition, teacher guidance, meeting duration, practical activities and adherence to the teaching plan were not documented in the materials available for this analysis. Consequently, the study characterizes temporal score change in a PBL-designated classroom rather than isolating the effects of a verifiable instructional treatment.

The headlamp and starter units could involve circuit tracing, fault diagnosis, teacher demonstration or independent exercises, but the record does not distinguish these activities. The timing of assessment and whether the same or equivalent test forms were used also remain undocumented. Claims concerning psychomotor skills or workplace readiness are avoided because only total achievement scores are available.

Implementation records are crucial when comparing PBL programmes. Devika et al. (2024) discuss transition challenges in engineering education; Das Neves et al. (2021) address teacher competencies for active learning; Bach et al. (2024) and Kassab et al. (2023) develop measures relevant to facilitation and engagement; and Tawfik et al. (2021) describe teachers' adaptations of PBL. These process measures were unavailable for this class.

Outcome definition and data quality procedures

For learner i, gain was defined as post-test score minus pre-test score. The primary quantity was the mean paired difference, supplemented by the median, dispersion and proportion with higher marks. A positive difference denotes increased recorded performance, not a causal treatment effect. Interpretation of all statistics is restricted to the observed grade-scale outcome.

All 34 observations were retained, including two negative differences. The assessments' number of items, item format, blueprint, scoring rubric, reliability and equivalence were not specified in the available documentation. Their absence matters because changes in total marks cannot, without supporting instrument evidence, be equated with changes in particular diagnostic competencies.

Assessment alignment therefore remains a priority. Ramos-Mejía and Padilla (2025) connect PBL to complex thinking in an electrochemistry course, whereas Suryanti and Nurhuda (2021) use an analytical rubric to examine critical thinking. These studies highlight the value of explicitly defined outcomes, unlike the present total-score-only record.

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

Statistical analysis

Pre-test, post-test and paired differences were summarized using means, SDs, medians, interquartile ranges and minima–maxima. A two-sided paired *t* test evaluated mean change with a 95% confidence interval. Shapiro–Wilk testing was used to assess the distribution of paired changes, followed by Wilcoxon signed-rank and exact sign tests as complementary checks.

The Wilcoxon signed-rank statistic evaluates signed ranks under its assumptions and should not automatically be interpreted as a mean test or a distribution-free median test. The sign test assesses the direction of nonzero changes. Concordance across these procedures was interpreted as robustness of the observed score contrast, not evidence of PBL causation.

Effect magnitude was summarized with within-person Cohen's *dz*, calculated as mean gain divided by the *SD* of paired differences. Class normalized gain and average student-level normalized gains were distinguished because they use different denominators. Matched-pairs rank-biserial correlation was treated as supplementary descriptive evidence.

A 20,000-resample paired percentile bootstrap provided a supplementary uncertainty interval for mean gain. A post hoc sensitivity analysis restricted the sample to learners with baseline scores above 40; these exploratory results are retained in Tabel 2 but not interpreted as a replacement for the full-sample analysis. Spearman's correlation described the baseline–gain association without assuming causal moderation.

Because gain contains the negative baseline score, its correlation with baseline is mathematically coupled. The scale also imposes the upper limit gain $\leq$ 100 minus pre-test score. Gambar 3 displays this limit. Ceiling concentration was summarized by counts at 100 and at least 95. Analyses emphasized effect sizes and descriptive bounds rather than the significance of numerous exploratory comparisons.

Ethical and reproducibility considerations

The analysis uses de-identified paired scores and does not reproduce the original student roster. Institutional permission, consent or applicable waiver documentation, and authorization for secondary use of minors' grades were not included among the materials available for verification. Consequently, this study does not assert an ethics approval or waiver that cannot be substantiated. The research team must supply the appropriate documentation before submission.

RESULTS AND DISCUSSION

Completeness and baseline–post-test summaries

All 34 learners contributed complete pairs. Pre-test marks ranged from 15 to 95 (mean 71.91, *SD* 20.67), whereas post-test marks ranged from 80 to 100 (mean 95.88, *SD* 5.70). The median increased from 75 to 97.5. Tabel 1 reports the full distributional comparison. These score summaries show increased attainment on the recorded tests, but the narrowing of marks should be considered alongside upper-scale saturation.

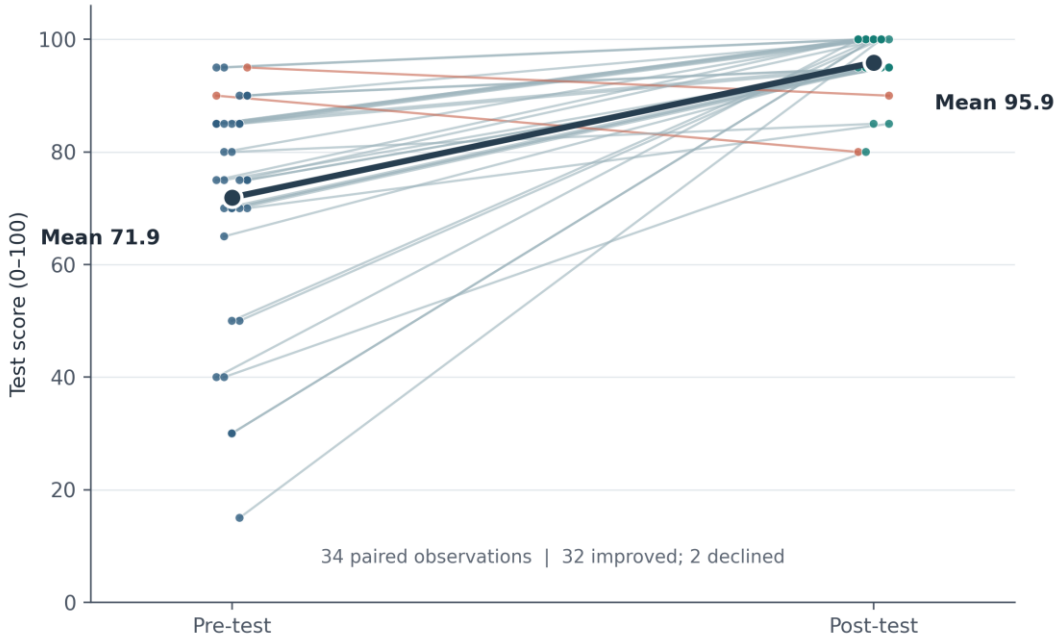
Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

Tabel 1. Descriptive statistics for paired automotive electrical systems test scores (n = 34).

Statistic	Pre-test	Post-test	Paired change
Valid observations	34	34	34
Mean (*SD*)	71.91 (20.67)	95.88 (5.70)	+23.97 (21.91)
Median (*IQR*)	75 (70–85)	97.5 (95–100)	+20 (10–25)
Range	15–95	80–100	−10 to +85
Scores ≥95	3/34 (8.82%)	29/34 (85.29%)	—
Scores =100	0/34 (0%)	17/34 (50.00%)	—

Catatan. Change = post-test − pre-test. *SD*, standard deviation; *IQR*, interquartile range. Percentages use 34 students as denominator.

Gambar 1 displays matched trajectories: most slopes are positive and many post-test observations cluster at 95 or 100. Two learners have declining scores. Paired visualization demonstrates that the change is broadly distributed rather than arising solely from different students contributing to each summary.



Gambar 1. Paired pre-test and post-test scores for 34 anonymized learners. Lines join scores from the same learner; the thicker line connects group means. Red trajectories indicate declines.

Magnitude and statistical evidence for change

Mean gain was 23.97 points (*SD* 21.91; 95% *t* confidence interval 16.33–31.61). The median gain was 20 points (*IQR* 10–25). The paired *t* test found $p < .001$, as did Wilcoxon signed-rank and exact sign tests. Shapiro–Wilk indicated nonnormal paired changes ($p = .0013$). The convergence of the paired procedures supports a strong difference between test occasions; none establishes a causal PBL effect.

The standardized paired difference was large (Cohen's *dz* = 1.094). Class normalized gain was 0.853 and mean individual normalized gain 0.745. A supplementary bootstrap confidence interval was 17.21–31.47 points; rank-biserial

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

correlation was 0.960. These quantities describe related but nonidentical aspects of the score contrast and must be interpreted against a bounded scale with many maximum marks.

Tabel 2. Primary paired comparison and complementary robustness statistics.

Measure	Estimate	Interpretive detail
Mean gain; 95% *t* CI	+23.97; [16.33, 31.61]	Points on 0–100 score scale
Bootstrap 95% CI	[17.21, 31.47]	20,000 paired percentile resamples
Paired *t* test	$t(33) = 6.38$; $p < .001$	Two-sided; mean change
Wilcoxon signed-rank	$W = 12$; $p < .001$	Two-sided; asymptotic
Exact sign test	32 positive, 2 negative; $p < .001$	Two-sided; nonzero pairs
Cohen's *dz*	1.094	Mean gain / *SD* of gains
Rank-biserial	0.960	Matched signed ranks
Class normalized gain	0.853	Ratio of class mean improvement to headroom
Mean individual gain	0.745	Mean of normalized individual gains
Shapiro–Wilk, gains	$W = 0.878$; $p = .00130$	Rejects normality of gain distribution
Baseline–gain Spearman	$\rho = -0.942$; $p < .001$	Mathematically coupled; exploratory
Sensitivity, baseline >40	$n = 29$; mean gain 16.90; $p < .001$	Post hoc restricted Wilcoxon

Catatan. CI, confidence interval; all inferential *p* values are two-sided. Complementary tests are not independent evidence of a treatment effect.

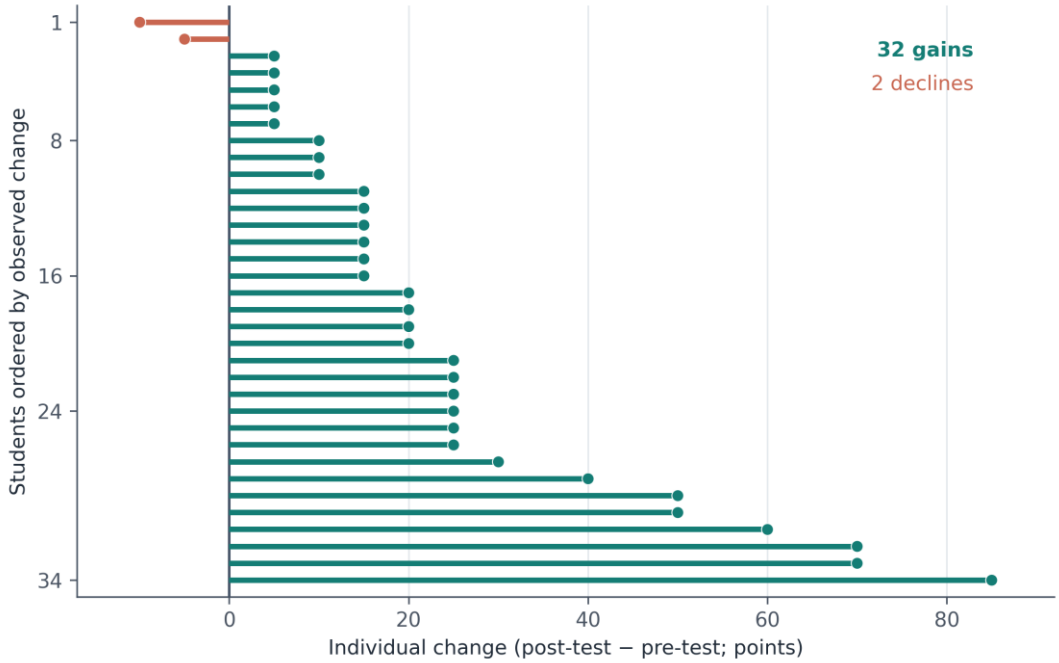
Tabel 2 assembles principal and supplementary results. Small *p* values quantify inconsistency with statistical null models under assumptions; they do not represent the probability that instruction caused improvement. Confidence intervals likewise should not be generalized automatically beyond one intact, nonrandomly selected class.

Student-level consistency and deviations

Thirty-two students (94.12%) improved and two (5.88%) declined; no score was unchanged. Gains ranged from minus 10 to plus 85 points. The two declines occurred among students starting at 95 and 90, whose final marks were 90 and 80. These high-baseline declines may reflect normal test variation, especially near the score ceiling, but cannot be diagnosed without item-level evidence.

Gambar 2 orders all 34 differences, making the uneven size of gains and the two declines visible. Most learners improved by comparatively modest amounts, while some recorded gains of 50 points or more. This heterogeneity supports examining student-level outcomes in addition to overall means.

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional



Gambar 2. Ordered changes in test scores at the student level. Learners are ordered by score change, not by identity; the vertical reference indicates no change.

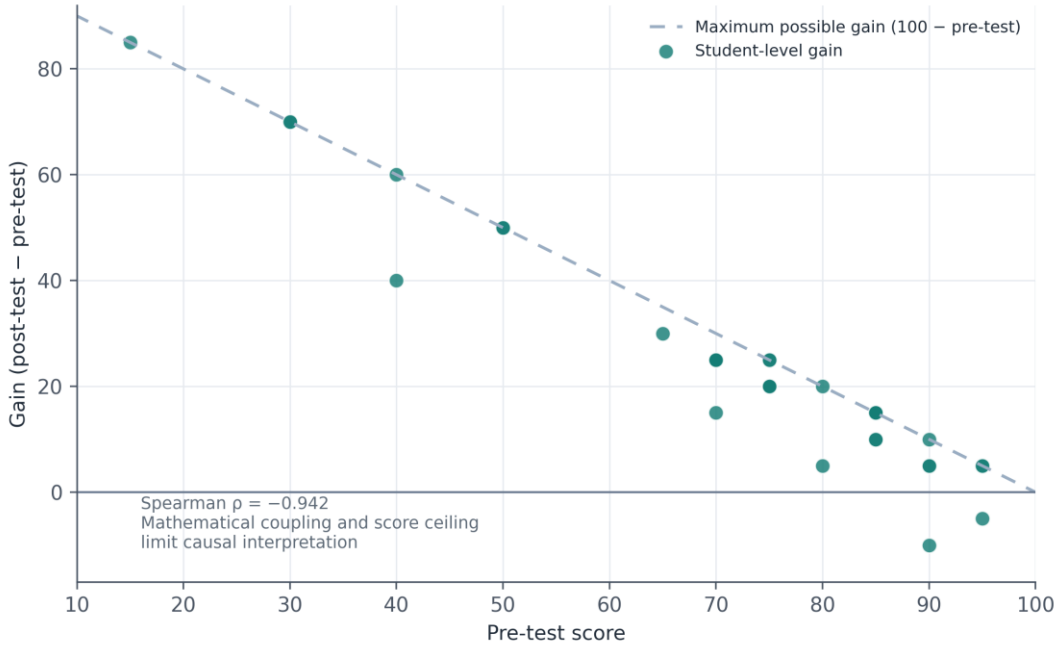
When five baseline scores of 40 or below were excluded in an exploratory check, the mean gain among 29 learners remained 16.90 points (Wilcoxon $p < .001$). The restriction confirms that large increases were not confined to very low starting scores, but it changes the study population and is not the primary estimate.

Baseline-score pattern and upper-bound constraints

Arithmetic gain was strongly inversely associated with baseline (Spearman's *rho* = -0.942, $p < .001$). The subgroup below 75 (13 students) gained 43.85 points on average, compared with 11.67 among 21 students starting at 75 or above. Such comparisons are mechanically constrained because learners closer to 100 have less room to gain; they do not demonstrate subgroup-specific PBL effects.

Gambar 3 shows that many gains lie near the maximum permitted by 100 minus baseline. Importantly, the raw pre-test–post-test correlation was weak (Pearson r = -0.085, p = .631), suggesting little stability in individual rank order across occasions. This could reflect instrument nonequivalence, low reliability, differential learning, or compressed post-test variance; the available scores cannot distinguish these explanations.

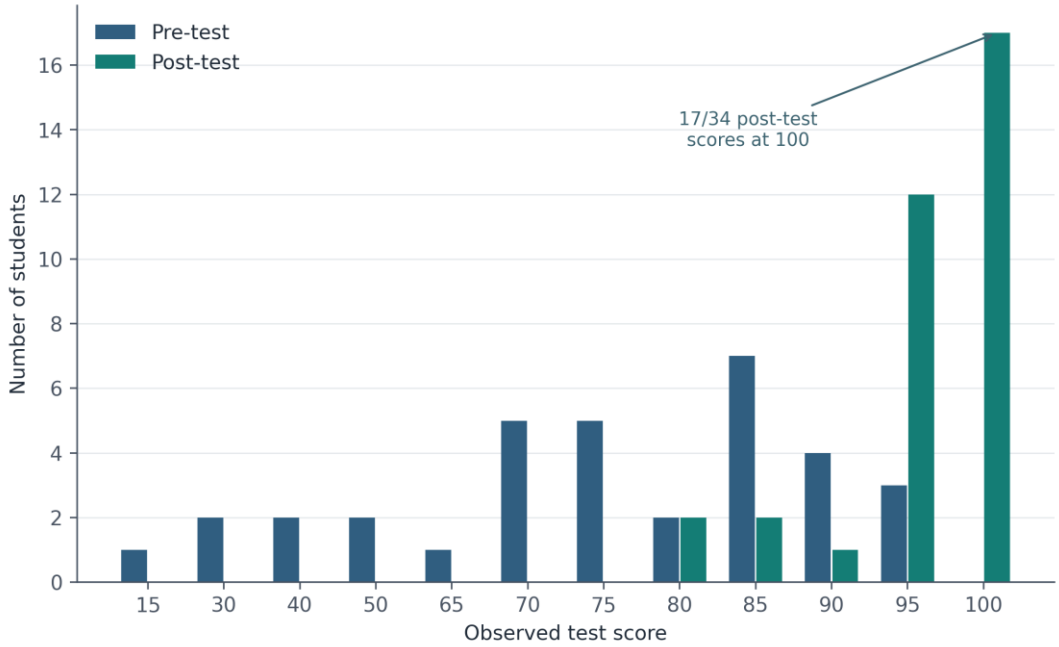
Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional



Gambar 3. Baseline scores, observed gains, and the score-ceiling boundary. The dashed line is the maximum possible arithmetic gain for a given baseline, not a fitted regression.

Score distribution and potential ceiling effect

Seventeen learners (50.00%) scored 100 at post-test and 29 (85.29%) scored at least 95. Before instruction, three (8.82%) scored at least 95 and none reached 100. Gambar 4 displays the pronounced shift toward the scale maximum, accompanying a reduction in observed *SD* from 20.67 to 5.70.



Gambar 4. Distribution of recorded marks at the two assessment occasions. Bars represent absolute frequencies; post-test scores accumulate at the maximum of 100.

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

When many students reach the maximum mark, the test cannot reliably discriminate their further achievement. The 17 perfect post-tests and 12 additional scores of 95 indicate substantial ceiling concentration; this does not prove the examination was invalid or that all learners mastered complex diagnostics. More demanding tasks might ask students to trace voltage loss through a headlamp circuit, distinguish an open relay coil from a failed ground, or justify starter-solenoid fault-isolation steps using measured voltages.

DISCUSSION

Interpretation and causal boundaries

The central result is a large increase in recorded achievement between the two testing occasions: 23.97 mean points, a 20-point median gain and improvement for 32 of 34 learners. Parametric and nonparametric analyses agree on the direction of change. The substantive contribution lies in reading that increase alongside the unequal student-level gains, weak cross-occasion ranking consistency and severe concentration at the top of the assessment scale.

The evidence establishes temporal change in a PBL-designated classroom, not the difference that would have arisen under another teaching method. Without a comparator, information about test equivalence or the elapsed instructional period, PBL effects are inseparable from prior exposure, practice, changes in test difficulty and ordinary learning over time. The results should therefore be reported as within-class score gains rather than experimental proof of comparative instructional effectiveness.

PBL could plausibly support conceptual learning through authentic troubleshooting, collaborative explanation and corrective feedback. In headlamp and starter systems, a meaningful problem might require students to identify components and evaluate competing causes of failure. However, no activity logs, fidelity checks or observations verify such processes in this class; any mechanism remains a hypothesis.

Other explanations warrant comparable scrutiny. Learners may have remembered the test format, studied independently or practised similar exercises between assessments. A post-test with lower difficulty could generate high marks without equivalent conceptual demand. Further, the near-ceiling post-test compresses individual variance and limits evidence about advanced understanding. A documented protocol and equivalent, calibrated assessments would be needed to isolate competing explanations.

Engineering applications make the implementation problem visible. Naik et al. (2024) evaluate PBL in freshman engineering; Islam and Chua (2025) incorporate an interactive platform into hybrid PBL; Das (2026) proposes a broader pedagogical framework; Reddy et al. (2024) investigate active-learning outcomes; and Ortiz Moreno et al. (2026) report a specified instructional redesign. These studies differ in exposure and measurement, reinforcing the need to document how PBL was implemented locally.

Heterogeneous gains and validity checks

The 94.12% improvement fraction indicates that the class-average shift did not depend on only a few high-gain learners. Nevertheless, differences spanning -10 to

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

+85 points reveal substantial heterogeneity, including two declines among initially high scorers. Their decreases could plausibly reflect assessment fluctuation near an upper score boundary, but the scores alone cannot determine why either student declined. Individual results should prompt formative investigation, not deficit labeling.

For educators, total-score trajectories can identify who may benefit from a review of circuit understanding; they do not identify a misconception. Item-level evidence would be needed to determine whether persistent problems involve relay control, grounding, voltage drop or starter actuation. Testing multiple questions per domain, recording scoring criteria and including practical demonstration would make subsequent individual feedback more informative.

PBL processes might also vary with the problems and tutors. Bendeliani and Torstensdotter (2026) discuss tutors' conceptions of the problem, Bai et al. (2023) examine varied case presentations, and de Pinho et al. (2021) study the sequence of classroom PBL activities. These sources justify documenting instructional scenarios rather than assuming particular classroom interactions occurred.

Group functioning is another potential source of variation. Li et al. (2022) review group function and outcomes, Saqr et al. (2023) examine the transfer of effective learning strategies, and Chang et al. (2022) compare collaborative arrangements paired with PBL. None of these group or strategy variables was available in the present record, preventing mediation analysis.

Delivery and scaffolding also differ across studies: Kristianto and Gandajaya (2023) examine offline and online PBL, Hursen (2021) combines PBL with Web 2.0 tools, and Abeysuriya et al. (2026) describe a structured immunology approach. Their mechanisms should not be transferred to this class without implementation evidence.

The inverse association between starting scores and arithmetic gains is especially easy to overread. Gain includes minus baseline by definition, and a student who initially scored 95 has only five available points of upward movement. The strong negative correlation consequently cannot support a claim that PBL specifically helped lower-achieving learners more than higher-achieving peers.

The raw pre-test–post-test correlation provides a separate measurement warning: r = -0.085, p = .631. For repeated scores intended to reflect related constructs, such weak association merits scrutiny, although it is not itself proof of invalidity. Near-ceiling compression, measurement error, nonequivalent items and real changes in ranking can all attenuate cross-occasion correlation. Without test forms or reliability estimates, their contributions cannot be apportioned.

Contextual evidence gives no basis for imposing an instructional subgroup mechanism on these observations. Lee (2025) examines first-time PBL participation, Avraam et al. (2025) studies learning approaches longitudinally, and Bustillo (2025) investigates PBL outcomes in a different setting. The present baseline strata were exploratory and are not adjusted comparisons of treatment response.

Score saturation and assessment implications

A more serious concern than the statistical significance of the mean gain is whether the assessment still distinguishes students near the top. Exactly half reached 100 and most others scored at least 95. These patterns justify describing a marked

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

observed score ceiling, but they do not establish why it occurred, how reliable the test was or whether practical automotive competence converged.

At the upper achievement range, advanced items could require diagnosing simultaneous faults, interpreting an incomplete wiring diagram, selecting safe multimeter measurements, distinguishing a weak battery from excessive starter-circuit resistance, and justifying a fault sequence based on evidence. Such tasks would extend discrimination beyond identifying component names or recalling a standard circuit. A validated analytic rubric and practical performance measure could complement the 0–100 classroom test.

Comparative PBL evidence likewise evaluates many distinct outcomes. Susanti et al. (2023) study problem-solving and critical thinking; Nantha et al. (2022) compare instructional methods for academic and problem-solving outcomes; and Singh et al. (2026) investigate engineering learning gains. These differences underscore why improvement in a single grade scale cannot be treated as proof of generalized competence.

Methodologically, the combination of paired trajectories, uncertainty estimates, sign consistency and ceiling diagnostics prevents a large mean from being interpreted in isolation. The *t* test describes average within-student change; Wilcoxon and sign analyses support its direction under different assumptions. The gain–baseline plot makes the score boundary explicit. Yet more statistical tests cannot repair missing comparator data or unverified instrument properties.

External validity is limited to one intact class of 34, with incompletely reported schooling and assessment context. The uncertainty intervals quantify the score contrast under conventional sampling assumptions; they do not automatically generalize to different schools, cohorts or instructional implementations. This is a focused class-level observation rather than a population estimate of PBL effectiveness.

Comparative research offers no justification for assuming PBL always outperforms alternatives. Borah et al. (2024) compare case-based and problem-based instruction, and Arrieta-Cohen et al. (2024) evaluate project-based intervention in a distinct setting. Ardalan and Molaeian (2025) discuss flexible engineering pedagogy, while Mora-Melia et al. (2024) integrate computational fluid dynamics into PBL. Their contexts and outcome definitions differ from this classroom's paired grade scale.

Other applied engineering studies emphasize structured activity and outcomes beyond aggregate marks (Awwad, 2025; Orji, 2025). The present data offer no direct evidence about practical performance, engagement or maintenance-task transfer. These distinctions should guide the design of future vocational assessments.

Implications and future research

For teachers, the most immediate implication is to look beyond a class mean. Paired changes can expose unexpectedly small gains and declines, while item-level analysis can identify questions requiring reteaching. A high final score should invite further diagnostic challenges rather than close the assessment process automatically.

For assessment designers, an upper-range instrument can combine conceptual questions with complex circuit-fault scenarios, reasoned selection of voltage and continuity tests, and observed safe troubleshooting. Recording the item blueprint,

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

scoring rubric, expert validation, reliability, equivalent-form evidence and a delayed assessment would substantially improve interpretability.

For researchers, the priority is a prospective comparison of a specified PBL sequence with a well-described alternative, preferably across more than one class. Researchers should document learning problems, session number, teacher support and implementation fidelity; administer appropriately matched assessments; and include practical diagnostic and retention outcomes. Such a design could test PBL mechanisms and comparative effectiveness directly rather than inferring them from one group's before–after contrast.

The important limitations are an uncontrolled single-class design, insufficient documentation of instructional exposure and assessment instruments, and absence of validated competence outcomes or retention data. Neither ethics authorization nor school context could be independently established from the supplied manuscript materials. The prominent post-test ceiling and mathematical coupling of baseline with gain further limit substantive interpretations.

Future research should collect institutional and consent documentation, instrument reliability and item-level responses before analysis. A comparator and delayed practical assessment would clarify whether apparent gains persist and transfer to unfamiliar electrical faults. These priorities follow directly from the observed score concentration and lack of process and measurement evidence, rather than from another significance test.

CONCLUSION

This analysis of 34 paired scores documents a substantial increase in recorded performance on headlamp and starter-system assessments. Mean marks rose from 71.91 to 95.88, with 32 learners improving and two declining. The mean paired gain was 23.97 points, and complementary tests consistently indicated a difference between occasions. Nevertheless, individual gains varied substantially and 17 learners attained the maximum post-test mark.

The findings support a descriptive within-class learning trajectory, not a causal estimate of PBL effectiveness. The nearly zero pre-test–post-test correlation and strong ceiling concentration heighten the importance of verifying instrument equivalence and reliability. Future controlled evaluation should document PBL delivery and measure challenging conceptual as well as practical diagnostic outcomes.

REFERENCE

Abeysuriya, V., Alwis, S. N., & Abeysuriya, V. (2026). Scaffolding problem-based learning in postgraduate immunology: A pyramidal mind-mapping approach to teaching complex immune mechanisms in COVID-19. *Advances in Physiology Education*, *50*(3), 855–862. https://doi.org/10.1152/advan.00021.2026

Affandy, H., Sunarno, W., Suryana, R., & Harjana. (2026). Unveiling the Global Instructional Impact of Problem-Based Learning in Science Education: A Comprehensive Systematic Review. *Interdisciplinary Journal of Problem-Based Learning*, *20*(1). https://doi.org/10.14434/ijpbl.v20i1.37007

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

Ahmad, Z., Ammar, M., Sellami, A., & Al-Thani, N. J. (2023). Effective Pedagogical Approaches Used in High School Chemistry Education: A Systematic Review and Meta-Analysis. *Journal of Chemical Education*, *100*(5), 1796–1810. https://doi.org/10.1021/acs.jchemed.2c00739

Akbar, A., Nugraha, E., Riza, L. S., Nazir, S., & Al Husaeni, D. N. (2026). IMPROVING STUDENTS COGNITIVE LEARNING OUTCOMES IN COMPUTER SYSTEMS LEARNING THROUGH UNITY-BASED VIRTUAL REALITY MEDIA. *Journal of Technology and Science Education*, *16*(2), 470–485. https://doi.org/10.3926/jotse.3677

Alshaye, I. A., Tasir, Z., & Jumaat, N. F. (2023). The Effectiveness of Online Problem-Based Learning Tasks on Riyadh's Secondary School Students' Problem-Solving Ability and Programming Skills. *Open Education Studies*, *5*(1). https://doi.org/10.1515/edu-2022-0208

Ardalan, A. A., & Molaeian, S. (2025). Flexible Education-A Paradigm Shift in Engineering Pedagogy. *Journal of Engineering Education Transformations*, *39*(1), 109–125. https://doi.org/10.16920/jeet/2025/v39i1/25118

Arici, F., & Yilmaz, M. (2023). An examination of the effectiveness of problem-based learning method supported by augmented reality in science education. *Journal of Computer Assisted Learning*, *39*(2), 446–476. https://doi.org/10.1111/jcal.12752

Aristin, N. F., Hastuti, K. P., Arisanty, D., Adyatma, S., & Donna, C. (2023). Effectiveness of problem-based learning models to improve learning outcomes of geography in the new normal learning era. *Journal of Education and Learning*, *17*(4), 623–632. https://doi.org/10.11591/edulearn.v17i4.20834

Arrieta-Cohen, M. C., Torres-Arizal, L. A., & Gómez-Yepes, R. L. (2024). Evaluating the Impact of an Educational Intervention Using Project-Based Learning on Postpandemic Recovery in Rural Colombia. *Education Sciences*, *14*(12). https://doi.org/10.3390/educsci14121341

Avraam, D., Televantou, I., Albert, A. P., Hitchings, A. W., Nicolaou, S. A., Papageorgiou, A., McCrorie, P., & Nicolaou, P. (2025). Exploring the relationship between learning approaches and problem-based learning: Insights from a longitudinal study in medical students. *BMC Medical Education*, *25*(1). https://doi.org/10.1186/s12909-025-07171-1

Awwad, F. (2025). Enhancing Electronics Courses Education: Active Learning Strategies for Undergraduate Engineering Students. *International Journal of Engineering Pedagogy*, *15*(2), 42–73. https://doi.org/10.3991/ijep.v15i2.51739

Bach, A., Blüthmann, I., Wulff, I., Thierfelder, I., Krebber, M., Watermann, R., Thiel, F., & Peters, H. (2024). Development and validation of a questionnaire on problem-based learning to evaluate facilitation by the tutor and the student group. *Medical Teacher*, *46*(5), 648–656. https://doi.org/10.1080/0142159X.2023.2271156

Bai, S., Zhang, L., Ye, Z., Yang, D., Wang, T., & Zhang, Y. (2023). The benefits of using atypical presentations and rare diseases in problem-based learning in

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

undergraduate medical education. *BMC Medical Education*, *23*(1). https://doi.org/10.1186/s12909-023-04079-6

Bendeliani, N., & Torstensdotter, G. L. (2026). The role of the "problem" in problem-based learning: PBL tutors' conceptions. *Scandinavian Journal of Educational Research*. https://doi.org/10.1080/00313831.2026.2625211

Borah, N., Paudel, S., & Stivers, A. (2024). Comparing the effectiveness of case-based learning and problem-based learning in a core finance class. *Journal of Education for Business*, *99*(2), 113–124. https://doi.org/10.1080/08832323.2023.2277272

Bustillo, R. (2025). Too Soon for Problem-Based Learning? Learning Outcomes. *International Journal of Pedagogy and Curriculum*, *32*(2), 59–76. https://doi.org/10.18848/2327-7963/CGP/v32i02/59-76

Chang, Y.-H., Yan, Y.-C., & Lu, Y.-T. (2022). Effects of Combining Different Collaborative Learning Strategies with Problem-Based Learning in a Flipped Classroom on Program Language Learning. *Sustainability (Switzerland)*, *14*(9). https://doi.org/10.3390/su14095282

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

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

Das Neves, R. M., Lima, R. M., & Mesquita, D. (2021). Teacher competences for active learning in engineering education. *Sustainability (Switzerland)*, *13*(16). https://doi.org/10.3390/su13169231

de Pinho, L. A., Alves, L. A., Comarú, M. W., da Luz, M. R. M. P., & Lopes, R. M. (2021). A processual view on the use of problem-based learning in high school physiology teaching. *Advances in Physiology Education*, *45*(4), 750–757. https://doi.org/10.1152/ADVAN.00056.2020

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

Erdem, C. (2024). A comparative meta-analysis of the effects of problem-based learning model on K-12 students' cognitive outputs. *Educational Studies*, *50*(6), 1498–1519. https://doi.org/10.1080/03055698.2022.2103650

Hursen, C. (2021). The Effect of Problem-Based Learning Method Supported by Web 2.0 Tools on Academic Achievement and Critical Thinking Skills in Teacher Education. *Technology, Knowledge and Learning*, *26*(3), 515–533. https://doi.org/10.1007/s10758-020-09458-2

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

Islam, M. R., & Chua, K. J. (2025). An Innovative Interactive Platform Assisted Hybrid Problem-Based Learning Approach for Concept-Heavy Engineering Modules. *European Journal of Education*, *60*(2). https://doi.org/10.1111/ejed.70066

Kanyesigye, S. T., Uwamahoro, J., & Kemeza, I. (2023). The Impact of Problem-Based Learning on Students' Achievement in Mechanical Waves in Secondary Schools. *Research in Science Education*, *53*(6), 1013–1033. https://doi.org/10.1007/s11165-023-10119-4

Kassab, S. E., El-Baz, A., Hassan, N., Hamdy, H., Mamede, S., & Schmidt, H. G. (2023). Construct validity of a questionnaire for measuring student engagement in problem-based learning tutorials. *BMC Medical Education*, *23*(1). https://doi.org/10.1186/s12909-023-04820-1

Kristianto, H., & Gandajaya, L. (2023). Offline vs online problem-based learning: A case study of student engagement and learning outcomes. *Interactive Technology and Smart Education*, *20*(1), 106–121. https://doi.org/10.1108/ITSE-09-2021-0166

Lee, Y.-C. (2025). Changes in Learning Outcomes of Students Participating in Problem-Based Learning for the First Time: A Case Study of a Financial Management Course. *Asia-Pacific Education Researcher*, *34*(1), 511–530. https://doi.org/10.1007/s40299-024-00873-y

Li, A., Bilgic, E., Keuhl, A., & Sibbald, M. (2022). Does your group matter? How group function impacts educational outcomes in problem-based learning: A scoping review. *BMC Medical Education*, *22*(1). https://doi.org/10.1186/s12909-022-03966-8

Mora-Melia, D., Gutiérrez-Bahamondes, J. H., Iglesias-Rey, P. L., & Martinez-Solano, F. J. (2024). Exploring the synergy of problem-based learning and computational fluid dynamics in university fluid mechanics instruction. *Computer Applications in Engineering Education*, *32*(6). https://doi.org/10.1002/cae.22782

Mujumdar, S. B., Acharya, H., Shirwaikar, S., & Bafna, P. B. (2024). Assessing impact of problem-based learning using data mining to extract learning patterns. *Journal of Applied Research in Higher Education*, *16*(2), 610–628. https://doi.org/10.1108/JARHE-05-2022-0165

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

Nantha, C., Pimdee, P., & Sitthiworachart, J. (2022). A Quasi-Experimental Evaluation of Classes Using Traditional Methods, Problem-Based Learning, and Flipped Learning to Enhance Thai Student-Teacher Problem-Solving Skills and Academic Achievement. *International Journal of Emerging Technologies in Learning*, *17*(14), 20–38. https://doi.org/10.3991/ijet.v17i14.30903

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

Orji, C. T. (2025). Efficacy of problem-based intervention on engagement and practical skills acquisition among electrical/electronic technology education students. *Discover Education*, *4*(1). https://doi.org/10.1007/s44217-025-00736-8

Ortiz Moreno, H., Godoy-Rangel, C., & Ahumada-Lazo, R. (2026). From lectures to active learning: A case study on problem-based learning in the material balances course. *International Journal of Mechanical Engineering Education*, *54*(4), 1101–1122. https://doi.org/10.1177/03064190251356618

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

Ramos-Mejía, A., & Padilla, K. (2025). A Problem-Based Learning Electrochemistry Course for Undergraduate Students to Develop Complex Thinking. *Education Sciences*, *15*(3). https://doi.org/10.3390/educsci15030320

Reddy, S. N., Pathlavath, M., Narsareddygari, S., & Naik, S. M. (2024). Investigating The Transformative Effects of Active Learning Methodologies in The Field of Engineering Education to Improve Learning Outcomes in Students by Unleashing Their Potential. *Journal of Engineering Education Transformations*, *37*(Special Issue 2), 562–567. https://doi.org/10.16920/jeet/2024/v37is2/24088

Saleh, A., Phillips, T. M., Hmelo-Silver, C. E., Glazewski, K. D., Mott, B. W., & Lester, J. C. (2022). A learning analytics approach towards understanding collaborative inquiry in a problem-based learning environment. *British Journal of Educational Technology*, *53*(5), 1321–1342. https://doi.org/10.1111/bjet.13198

Saqr, M., & López-Pernas, S. (2023). The temporal dynamics of online problem-based learning: Why and when sequence matters. *International Journal of Computer-Supported Collaborative Learning*, *18*(1), 11–37. https://doi.org/10.1007/s11412-023-09385-1

Saqr, M., Matcha, W., Uzir, N. A., Jovanović, J., Gašević, D., & López-Pernas, S. (2023). Transferring effective learning strategies across learning contexts matters: A study in problem-based learning. *Australasian Journal of Educational Technology*, *39*(3), 35–57. https://doi.org/10.14742/ajet.8303

Servant-Miklos, V. F. C., & Kolmos, A. (2022). Student conceptions of problem and project based learning in engineering education: A phenomenographic investigation. *Journal of Engineering Education*, *111*(4), 792–812. https://doi.org/10.1002/jee.20478

Singh, S., Singh Gautam, B., Arjunan, T. V., Das, B., & Tripathi, A. K. (2026). A pedagogical teaching-learning framework for engineering undergraduates: Learning gains in a material science and metallurgy course. *International Journal of Mechanical Engineering Education*. https://doi.org/10.1177/03064190261427723

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional

Suryanti, N., & Nurhuda, N. (2021). The Effect of Problem-Based Learning with an Analytical Rubric on the Development of Students' Critical Thinking Skills. *International Journal of Instruction*, *14*(2), 665–684. https://doi.org/10.29333/iji.2021.14237a

Susanti, M., Suyanto, S., Jailani, J., & Retnawati, H. (2023). Problem-based learning for improving problem-solving and critical thinking skills: A case on probability theory course. *Journal of Education and Learning*, *17*(4), 507–525. https://doi.org/10.11591/edulearn.v17i4.20866

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

Tawfik, A. A., Gish-Lieberman, J. J., Gatewood, J., & Arrington, T. L. (2021). How K-12 Teachers Adapt Problem-Based Learning. *Interdisciplinary Journal of Problem-Based Learning*, *15*(1). https://doi.org/10.14434/ijpbl.v15i1.29662

Uluçinar, U. (2023). The Effect of Problem-Based Learning in Science Education on Academic Achievement: A Meta-Analytical Study. *Science Education International*, *34*(2), 72–85. https://doi.org/10.33828/sei.v34.i2.1

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 ELEARNING INTEGRATION IN VOCATIONAL EDUCATION. *Journal of Educators Online*, *22*(4). https://doi.org/10.9743/JEO.2025.22.4.15

Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional