Taushiah: Jurnal Hukum, Pendidikan dan Kemasyarakatan **ISSN: 2599-1353**
Vol. 15 No.2 (2025)

THE EFFECT OF DEEP LEARNING ON ISLAMIC RELIGIOUS EDUCATION LEARNING OUTCOMES

Fatimah Hasibuan, Mustapa, Rini Dewi Andriani
Universitas Islam Sumatera Utara
Email: fatimahhasibuan245@gmail.com, mustapa@fai.uisu.ac.id, parianto@fai.uisu.ac.id

Abstrak

Penelitian ini bertujuan menganalisis kondisi awal hasil belajar, peningkatan capaian, dan perbedaan hasil belajar Pendidikan Agama Islam melalui pendekatan pembelajaran mendalam berbantuan strategi pembelajaran berbasis proyek pada materi Fikih Muamalah. Penelitian menggunakan pendekatan kuantitatif dengan desain kuasi-eksperimen pretest–posttest kelompok kontrol nonekuivalen. Sampel dipilih secara purposif sebanyak 45 siswa kelas VIII, terdiri atas 22 siswa kelas kontrol dan 23 siswa kelas eksperimen. Pengumpulan data menggunakan tes hasil belajar pilihan ganda sebelum dan sesudah pembelajaran. Rerata pretest kelas kontrol sebesar 55,50 dan kelas eksperimen 58,00; perbedaannya tidak signifikan ($p = 0{,}125$). Rerata posttest meningkat menjadi 66,64 pada kelas kontrol dan 82,91 pada kelas eksperimen. Kenaikan rerata masing-masing sebesar 11,14 dan 24,91 poin, dengan selisih kenaikan 13,77 poin. Uji beda posttest menunjukkan perbedaan signifikan ($p < 0{,}001$). Uji normalitas dan homogenitas yang dilaporkan memenuhi kriteria parametrik. Temuan mengindikasikan bahwa penerapan pembelajaran mendalam berbasis proyek berkaitan dengan capaian kognitif lebih tinggi dibandingkan pembelajaran konvensional pada sampel yang diteliti. Namun, pemilihan kelas tanpa pengacakan membatasi penarikan kesimpulan kausal dan generalisasi. Penelitian ini mendukung pengembangan pembelajaran Fikih Muamalah yang partisipatif dan berorientasi pemahaman.

Kata kunci: Pembelajaran Mendalam, Pembelajaran Berbasis Proyek, Hasil Belajar

Abstract

This study analyzed baseline achievement, learning gains, and differences in Islamic Religious Education outcomes following a deep-learning pedagogical approach supported by project-based learning in Fiqh Muamalah. A quantitative, nonequivalent-control-group pretest–posttest quasi-experimental design was employed. Purposive sampling selected 45 eighth-grade students, comprising 22 control-group and 23 experimental-group participants. Cognitive learning outcomes were assessed with multiple-choice tests before and after instruction. The control and experimental groups obtained pretest means of 55.50 and 58.00, respectively; the difference was not statistically significant ($p = .125$). Their posttest means increased to 66.64 and 82.91, corresponding to gains of 11.14 and 24.91 points and a between-group difference in gains of 13.77 points. The reported posttest comparison was statistically significant ($p < .001$). Reported normality and homogeneity tests satisfied the stated criteria for parametric analysis. These results suggest that project-supported deep learning was associated with higher cognitive achievement than conventional instruction in the observed classes. However, nonrandom class allocation limits causal interpretation and generalizability. The study offers empirical support for developing more participatory, contextually grounded approaches to teaching Fiqh Muamalah.

Keywords: Deep Learning, Project-Based Learning, Learning Outcomes

INTRODUCTION

Islamic Religious Education is expected to cultivate more than the ability to reproduce religious terminology. At the lower-secondary level, students should interpret principles, reason about ethical choices, and relate Islamic teachings to decisions made in families, communities, and everyday transactions. This expectation is especially relevant to Fiqh Muamalah, where concepts such as fair exchange, debt, and the prohibition of usury require learners to distinguish acceptable practices from problematic conduct. A lecture can introduce rules efficiently, but an instructional sequence restricted to explanation and recall may give students insufficient opportunities to analyze situations in which those rules apply. Meaningful achievement therefore requires activities that invite students to connect conceptual knowledge, moral reasoning, and practical examples. Project-based learning provides one possible route because it organizes instruction around tasks that demand inquiry, collaboration, and demonstrable products (Bell, 2010; Blumenfeld et al., 1991).

The term deep learning in this article denotes a pedagogical approach emphasizing meaningful understanding, reflection, conscious participation, and application; it does not denote machine-learning algorithms. In the investigated classroom it was combined with project-based learning (PjBL) as an instructional strategy. These labels should not be interpreted as two separately manipulated treatments: the experiment compares a combined instructional arrangement with conventional instruction. The conceptual rationale is that sustained questioning and authentic cases can promote connections among rules, reasons, and consequences, whereas copying definitions alone may encourage surface-level performance. Barron et al. (1998) argue that understanding in project-based curricula requires aligned goals, scaffolding, formative assessment, and opportunities for revision. Similarly, Hmelo-Silver (2004) shows how guided problem-centered learning can support flexible knowledge and self-directed inquiry. Although problem-based and project-based learning are distinct traditions, both emphasize learners' active construction and application of knowledge.

At SMP IT Ad-Durrah Medan Marelan, the thesis reports preliminary observations that instruction in Islamic Religious Education relied considerably on teacher explanation and that some students experienced limited participation and insufficient home learning support. Those observations motivated a classroom-based comparison focusing on eighth-grade Fiqh Muamalah. They are contextual descriptions rather than independently tested determinants of achievement. The chosen topic was educationally relevant because students must reason through exchanges, borrowing, and responsible economic behavior, not merely remember isolated definitions. A project may help a learner investigate a familiar transaction, identify its relevant principles, and justify a decision. Yet the presence of a project by itself does not guarantee conceptual understanding: the quality of guidance, task design, discussion, and assessment remains critical (Kokotsaki et al., 2016).

Evidence from broader educational research supports examining this approach empirically. Chen and Yang (2019) found a favorable overall association between project-based learning and academic achievement across studies, while also identifying differences associated with contextual moderators. Zhang and Ma (2023), using a meta-analysis of experimental and quasi-experimental evidence, similarly reported positive effects on academic achievement and other outcomes. These syntheses concern heterogeneous subjects and populations; their aggregate estimates cannot substitute for observations in a particular Islamic education classroom. Rather, they justify a focused comparison using directly measured outcomes. A local study is necessary

because instruction addressing religious legal reasoning may involve different assessment demands and cultural expectations from projects in science or general education.

Research situated closer to Islamic Religious Education further clarifies the need for context-sensitive implementation. Sari and Handayani (2025) describe planning and collaborative participation in project-based Islamic education at the elementary level. Nasripin et al. (2026) emphasize teachers' mediation of inquiry, discussion, presentation, and religious habituation in Islamic education classrooms. Pradipta et al. (2026) examine project-based elements alongside gamification in high-school Islamic education and highlight engagement and value internalization. Such studies illuminate plausible pedagogical processes but differ in educational level, treatment combination, method, or outcome. Consequently, they do not directly establish whether a combined deep-learning and project-based arrangement produces a higher cognitive test score than conventional teaching for eighth-grade Fiqh Muamalah. This difference of subject matter, participant level, and comparative quantitative design defines the present study's distinction rather than an unsupported claim that project-based learning has never been studied in Islamic education.

The original thesis identifies three concerns: the learning condition before intervention, student achievement following instruction, and the extent of improvement under the instructional approach. These concerns are operationalized here as three objectives. First, the article describes and compares the two classes' initial cognitive achievement before the intervention. Second, it examines within-class changes from pretest to posttest and the magnitude of the descriptive gain under the contrasting teaching arrangements. Third, it evaluates whether post-intervention achievement differs statistically between the experimental and control classes. The associated hypothesis is that posttest achievement will differ between groups, with higher scores expected in the experimental class. The design tests between-group differences rather than independently estimating the contribution of each feature of the combined intervention.

The article contributes a transparent account of comparative achievement in a specific Islamic Religious Education unit, using reported means, measurement checks, and hypothesis tests rather than inferring success exclusively from students' apparent enthusiasm.

METHOD

The study employed a quantitative quasi-experimental approach using a nonequivalent-control-group pretest–posttest design. The independent condition was instructional assignment: conventional teacher-led learning in one intact class versus a deep-learning pedagogical approach implemented through project-based learning in another. The dependent variable was cognitive achievement in the Fiqh Muamalah unit of Islamic Religious Education, operationalized as scores on a multiple-choice achievement test. Because the research did not randomly allocate individual students or classes, the design permits a comparative estimate of learning outcomes but does not remove all potential selection or classroom confounding.

The research took place at SMP IT Ad-Durrah, Medan Marelan, North Sumatra. The target population comprised 111 eighth-grade students across five classes. Purposive selection identified class VIII Biduri, with 22 students, as the control group and class VIII Diamond, with 23 students, as the experimental group, yielding 45 participants. Both classes studied the same subject matter and completed pretests and posttests using the reported test format. The experimental instruction encouraged case analysis, group discussion, project completion, presentation, and connections between Islamic transaction rules and daily-life situations; the comparison class received conventional lecture-based teaching. The thesis does not establish a

detailed standardized fidelity rubric or random treatment assignment, so these features are not claimed here.

Achievement was measured through a reported 20-item multiple-choice instrument. The thesis states that item validity was examined using Pearson item correlations and that all items met its decision criterion. Reported internal-consistency coefficients were Cronbach's alpha = .917 for the pretest instrument and .978 for the posttest instrument. These figures indicate high reported internal consistency, although the available summary does not supply an independent audit of the scoring procedure. Scores were summarized by group using means, standard deviations, and observed minimum and maximum values. Pretest comparability was assessed before treatment, and changes within each group were described by subtracting the respective pretest mean from its posttest mean.

For inferential analysis, the thesis reports Shapiro–Wilk normality checks and Levene homogeneity tests, followed by independent-samples t-tests comparing the two groups at pretest and posttest. The decision threshold was alpha = .05. Shapiro–Wilk p-values were .533 and .371 for the control and experimental pretests, and .174 and .114 for the corresponding posttests. Levene p-values were .511 at pretest and .641 at posttest. These results did not provide evidence against the selected normality and equal-variance assumptions at the stated level. A reported p-value of .000 is expressed in this article as $p < .001$ rather than as a zero probability.

RESULTS AND DISCUSSION

Baseline Islamic Religious Education Achievement

The first objective was to characterize baseline cognitive achievement in Fiqh Muamalah before the instructional conditions were compared. The two intact classes comprised 22 participants in VIII Biduri and 23 in VIII Diamond. At pretest, the control class obtained a mean of 55.50, with a standard deviation of 5.72, whereas the experimental class obtained a mean of 58.00, with a standard deviation of 5.00. The observed difference in means was 2.50 points in favor of the experimental class. This difference is small in relation to the subsequent posttest difference and indicates that the groups began the instructional sequence at somewhat different, but broadly comparable, levels of demonstrated knowledge. It would nevertheless be inaccurate to call their baseline achievement identical.

The thesis reports an independent-samples comparison of baseline scores yielding $p = .125$. At the conventional alpha = .05 threshold, this result does not reject the null hypothesis of equal group means. This finding provides a useful diagnostic check for the planned comparison, because a large initial disparity could make final achievement differences more difficult to interpret. However, a nonsignificant p-value is not a formal equivalence test and does not prove that the groups were exchangeable with respect to motivation, prior academic experience, teacher exposure, or family support. In a purposively selected intact-class design, such unmeasured factors can differ even when the observed pretest means are statistically indistinguishable.

Distributional checks offer further context. Reported Shapiro–Wilk significance values for the control and experimental pretests were .533 and .371, respectively. The pretest Levene test yielded $p = .511$. None of these tests led to rejection of the corresponding diagnostic assumption at the .05 threshold. Accordingly, the original study proceeded with an independent-samples t-test. The decisions follow conventional statistical logic, but assumption-checking should not be overinterpreted: normality tests have limited power in modest samples, and homogeneity does not resolve potential differences in pre-intervention educational experiences.

Educationally, the baseline means indicate substantial room for improvement in the measured unit. Both groups had already encountered some ideas relevant to Islamic economic interactions, yet their average test performance before the unit remained below the levels observed afterward. The design cannot identify which particular concept—fair exchange, debt, or usury—accounted for baseline errors, because the supplied summary does not present an item-by-item domain breakdown. Consequently, the descriptive finding supports a need for stronger conceptual teaching but cannot identify a specific misconception or claim that all participants lacked ethical understanding.

The contrast between initial test performance and the broader goals of Islamic education is significant. The subject aims to cultivate principled decision-making, whereas the observed outcome indicator captures achievement on a 20-item cognitive measure. Motivation, ethical conduct, and long-term transfer were not independently measured as primary outcomes. Fredricks et al. (2004) describe academic engagement as behavioral, emotional, and cognitive; these aspects should not be inferred from test performance alone. Likewise, Pintrich and De Groot (1990) distinguish motivational and self-regulatory processes from observed classroom grades and examination performance. These distinctions protect the study from treating a baseline achievement score as a complete diagnosis of students' religious development.

Methodological literature also cautions against conflating a stable starting point with randomized assignment. Prince (2004) shows that active-learning effects must be understood through the instructional forms and evidence employed; broad claims require careful matching between interventions and assessment. The present pretest therefore serves a narrower purpose: documenting the observable starting achievement of two existing classes, not validating all possible background characteristics. The reported means and p-value are consistent with proceeding to a comparative posttest analysis, while preserving the possibility of residual confounding.

Table 1. Pretest achievement and assumption checks

Indicator	Control (n = 22)	Experimental (n = 23)
Pretest mean	55.50	58.00
Pretest standard deviation	5.72	5.00
Shapiro–Wilk p-value	.533	.371
Between-group t-test p-value	.125	
Levene homogeneity p-value	.511	

Source: Thesis research results; p-values reported by the original study.

Changes in Achievement Following the Two Instructional Approaches

The second objective was to describe the achievement gains observed after the contrasting instructional arrangements. The control group mean increased from 55.50 at pretest to 66.64 at posttest, an improvement of 11.14 score points. The experimental group mean increased from 58.00 to 82.91, an improvement of 24.91 points. The difference between these two descriptive gains was 13.77 points. The posttest mean also exceeded the control mean by 16.27 points. These calculations describe the central tendency of each class and directly reflect the results reported in the thesis. They do not require assuming that every student improved or that the gain was uniform within the group.

The posttest spread was not identical. The control group had a reported standard deviation of 6.58, compared with 7.24 for the experimental group. The respective observed posttest ranges were 58–81 and 70–93. Consequently, the higher experimental mean was accompanied by somewhat greater score variation, not a collapse of individual differences. Although the ranges overlap, their position and central tendencies suggest a consistent class-level

shift toward higher achievement in the experimental group. The available descriptive summary does not support claims about which specific individuals benefited most, nor does it establish achievement gains by gender, prior ability category, or particular Fiqh Muamalah concept.

The instructional explanation offered by the thesis emphasizes discussion, case analysis, collaborative project work, and links between Islamic legal rules and daily transactions. These activities could have encouraged students to elaborate concepts, justify decisions, and test their understanding against familiar situations. They also potentially increased opportunities for correction before the posttest. Such mechanisms are pedagogically plausible but were not independently isolated by an observational fidelity study or a formal mediation model. Therefore, they should be presented as interpretations of the treatment design, not as statistically verified pathways through which the intervention produced a measured effect.

Research on active learning provides useful interpretive parallels. Freeman et al. (2014) synthesized evidence of improved performance under active instruction in STEM higher education, but the differences in educational level and subject prohibit transferring their effect estimate to Islamic Religious Education. Reeve et al. (2004) showed that teachers' autonomy-supportive practice can promote engagement, suggesting why greater decision-making within guided projects might matter. Similarly, Ryan and Deci (2020) explain that autonomy, competence, and relatedness are relevant conditions for sustained motivation. The thesis, however, measured cognitive test achievement rather than these psychological needs, so the present data cannot establish increases in autonomous motivation or engagement as outcomes.

A distinguishing feature of the intervention was its applied content. In Fiqh Muamalah, learners can connect abstract rules to realistic decisions about sales, credit, borrowing, fair dealing, and prohibited practices. A project should encourage accurate interpretation rather than simply produce an attractive artifact. Zimmerman (2002) argues that self-regulated learning involves setting goals, monitoring progress, and reflecting on performance. These practices are compatible with structured project activities and may help explain why guided application produces deeper consolidation than passively receiving information. Again, the study did not collect separate self-regulation scores, and that proposed mechanism must remain theoretical.

The large descriptive difference in gains is educationally important, yet its evidential meaning must be stated carefully. Mean gain comparisons are not the same as a statistical test of the difference between individual change scores. No standard deviation of paired change, paired-test result, or adjusted difference-in-differences model is reproduced in the available thesis summary. Accordingly, this subsection reports the exact descriptive changes but does not attach a newly calculated significance level or confidence interval to the 13.77-point gain contrast. The stronger inferential claim concerns the posttest between-group comparison examined in the third objective.

Table 2. Descriptive pretest–posttest learning outcomes

Measure	Control (n = 22)	Experimental (n = 23)
Pretest mean	55.50	58.00
Posttest mean	66.64	82.91
Mean gain (points)	11.14	24.91
Posttest standard deviation	6.58	7.24
Posttest minimum–maximum	58–81	70–93

Source: Thesis research results. Between-group difference in descriptive gains = 13.77 points.

Difference in Posttest Achievement Between Experimental and Control Classes

The third objective was to evaluate whether post-intervention learning outcomes differed between the deep-learning/PjBL and conventional-instruction classes. The reported independent-samples posttest comparison produced a significance value displayed as .000 by the statistical software; in standard reporting, this is written as $p < .001$. Since this value is below the prespecified alpha of .05, the null hypothesis of equal posttest means was rejected. In the observed sample, students in the experimental class achieved a higher mean (82.91) than those in the control class (66.64), with a difference of 16.27 points. The result is consistent with an advantage for the combined deep-learning and project-supported approach under the studied classroom conditions.

The inference is supported by the reported testing sequence rather than by a single uncontextualized comparison. At baseline, the two classes had pretest means of 58.00 and 55.50, and the independent-samples test yielded $p = .125$. The posttest normality significance values were .114 in the experimental group and .174 in the control group, and the reported Levene result was $p = .641$. These diagnostics did not contradict the stated assumptions used for the equal-variance t-test. The contrast between a nonsignificant baseline test and a significant posttest test supports interpreting an instructional-group difference as emerging during the study period. Nonetheless, a change from nonsignificant to significant is not itself a formal test of differences in change; nor does baseline nonsignificance eliminate all alternative explanations.

The intervention was multidimensional: students experienced a learning approach promoting reflection and meaningful understanding, operationalized through project-based work. Because these features were introduced together, the data do not identify whether the observed difference resulted primarily from projects, teacher guidance, peer discussion, classroom enthusiasm, additional practice, or their combination. Likewise, the study does not quantify teacher effects, instructional time, project fidelity, or home support. Consequently, the term effect is used in the limited experimental-comparison sense documented by the thesis, while the contribution of individual mechanisms remains unresolved. Unmeasured differences between intact classes may also have influenced the observed scores.

This cautious reading is consistent with educational research on comparative pedagogy. Strobel and van Barneveld (2009) show that relative advantages of problem-based instruction depend on the outcome being measured; long-term retention and applied competence may differ from immediate standardized-test performance. Deci and Ryan (2000) provide a theoretical account of meaningful internalization but do not establish that the current participants experienced it. Roorda et al. (2011) document associations between teacher–student relationships and engagement or achievement across broader educational settings, reminding us that interpersonal instructional conditions may be relevant. None of these external results is used as a substitute for the local data or as confirmation of mechanisms the thesis did not measure.

For Islamic Religious Education, the finding suggests that teachers might supplement explanations of Fiqh Muamalah with carefully guided scenarios, collaborative inquiry, and reflective discussion in which students justify legal and ethical reasoning. The aim should be to strengthen understanding of principles rather than merely to assign work labeled a project. The teacher's responsibilities include ensuring doctrinal accuracy, clarifying misconceptions, supporting groups that struggle, and aligning the achievement test with the stated learning objectives. This implication follows from the observed higher test scores and the instructional design; it does not establish that students subsequently practiced Islamic economic ethics more faithfully outside school.

Several constraints should frame interpretation. The total sample included 45 students from two purposively chosen classes in one school; the resulting statistical finding does not directly

generalize to other grade levels or institutions. Neither random allocation nor a fully documented statistical adjustment for baseline achievement was undertaken.

Table 3. Summary of posttest statistical comparison

Indicator	Reported result
Control posttest mean	66.64
Experimental posttest mean	82.91
Difference between posttest means	16.27 points
Independent-samples t-test	$p < .001$
Posttest Shapiro–Wilk p-values	Control .174; experimental .114
Posttest Levene p-value	.641

Source: Thesis research results. The software display ".000" is reported as $p < .001$.

CONCLUSION

This quantitative quasi-experimental study examined the achievement of eighth-grade students studying Fiqh Muamalah through a deep-learning pedagogical approach implemented with project-based learning. The initial achievement means of the control and experimental classes were 55.50 and 58.00, respectively, and the reported pretest comparison was not statistically significant ($p = .125$). Following instruction, the control mean rose to 66.64, whereas the experimental mean reached 82.91. Their respective descriptive gains were 11.14 and 24.91 points, giving a 13.77-point difference in mean gains. The posttest comparison indicated a statistically significant difference between classes ($p < .001$), favoring the experimental arrangement. These findings suggest that contextually guided project activities can support cognitive understanding of Islamic transaction principles more effectively than the conventional instruction used in the comparison class. Nevertheless, the intervention combined several instructional features, and the purposively selected intact classes were not randomly assigned. Therefore, the evidence does not identify the contribution of each pedagogical component or justify unrestricted causal claims. The outcomes concern short-term test achievement, not independently measured changes in moral behavior, motivation, or long-term retention. Further studies should strengthen treatment documentation, expand school participation, adjust for baseline achievement, and evaluate durable understanding and practical application.

REFERENCES

Barron, B. J. S., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., & Bransford, J. D. (1998). Doing with understanding: Lessons from research on problem- and project-based learning. *The Journal of the Learning Sciences*, *7(3–4), 271–311*. https://doi.org/10.1080/10508406.1998.9672056

Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *83(2), 39–43*. https://doi.org/10.1080/00098650903505415

Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, *26(3–4), 369–398*. https://doi.org/10.1080/00461520.1991.9653139

Chen, C.-H., & Yang, Y.-C. (2019). Revisiting the effects of project-based learning on students' academic achievement: A meta-analysis investigating moderators. *Educational Research Review*, *26, 71–81*. https://doi.org/10.1016/j.edurev.2018.11.001

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11(4), 227–268*. https://doi.org/10.1207/S15327965PLI1104_01

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74(1), 59–109*. https://doi.org/10.3102/00346543074001059

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111(23), 8410–8415*. https://doi.org/10.1073/pnas.1319030111

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16(3), 235–266*. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools*, *19(3), 267–277*. https://doi.org/10.1177/1365480216659733

Nasripin, N., Beddu, M. J., & Miswanto, M. (2026). Teacher strategy as pedagogical mediation in project-based learning: Enhancing motivation and learning achievement in Islamic Religious Education. *Muaddib: Islamic Education Journal*, *9(1), 59–74*. https://doi.org/10.19109/muaddib.v9i1.37104

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82(1), 33–40*. https://doi.org/10.1037/0022-0663.82.1.33

Pradipta, M. A. S., Susetyo, M. B., Budiyanti, N., Mubarokah, S. F., & Helmy, H. A. (2026). The effectiveness of integrating gamification and project-based learning to enhance student engagement and value internalization in high school Islamic education. *Atthulab: Islamic Religion Teaching and Learning Journal*, *11(1), 21–35*. https://doi.org/10.15575/atthulab.v11i2.53875

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93(3), 223–231*. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion*, *28(2), 147–169*. https://doi.org/10.1023/B:MOEM.0000032312.95499.6f

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, *81(4), 493–529*. https://doi.org/10.3102/0034654311421793

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61, Article 101860*. https://doi.org/10.1016/j.cedpsych.2020.101860

Sari, N., & Handayani, R. P. (2025). Exploration of project-based learning in Islamic Religious Education in elementary schools. *Ahlussunnah: Journal of Islamic Education*, *4(3), 895–903*. https://doi.org/10.58485/jie.v4i3.723

Strobel, J., & van Barneveld, A. (2009). When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problem-Based Learning*, *3(1), 44–58*. https://doi.org/10.7771/1541-5015.1046

Zhang, L., & Ma, Y. (2023). A study of the impact of project-based learning on student learning effects: A meta-analysis study. *Frontiers in Psychology*, *14, Article 1202728*. https://doi.org/10.3389/fpsyg.2023.1202728

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41(2), 64–70*. https://doi.org/10.1207/s15430421tip4102_2