ORIGINAL ARTICLE

ACTIVE LEARNING AND GRADE 10 ACHIEVEMENT: A QUASI-EXPERIMENTAL POSTTEST COMPARISON

Razlani[1*] Shelly Samual[2] Cherys Fomy Laloan[3]
[123]Universitas Negeri Manado, Indonesia
*Corresponding Author: xiirazlan53@gmail.com

ABSTRACT
Teacher-centred instruction can restrict opportunities for secondary students to explain, apply, and collaboratively refine knowledge. This study compared immediate cognitive posttest achievement between two intact Grade 10 classes at SMA Negeri 1 Tompaso, Indonesia: one class received a broadly defined active-learning condition and one received conventional instruction. A quantitative non-equivalent control-group design involved 60 students, with 30 students in each class. The source record states that pretests were administered and that the achievement test underwent item-validity and internal-consistency checks; however, baseline statistics, psychometric coefficients, the academic subject, intervention duration, teacher allocation, and implementation-fidelity evidence were unavailable. Posttest scores were analysed using descriptive statistics and a two-tailed independent-samples t test after reported normality and homogeneity checks. The active-learning class obtained a higher mean ($M = 83.56$, $SD = 3.69$) than the conventional class ($M = 72.40$, $SD = 4.49$), $t(58) = 9.906$, $p < .001$, mean difference = 11.16, 95% CI [8.90, 13.41], Hedges' $g = 2.68$. The result documents a large site-specific posttest association but cannot separate instructional condition from baseline class differences or other class-level influences. The study therefore offers promising contextual evidence rather than a population-wide causal estimate. Future research should use multiple classes per condition, baseline-adjusted analyses, replicable intervention protocols, validated measures, fidelity checks, and delayed assessments of retention and transfer.

Keywords: Active learning, Cognitive achievement, Non-equivalent control-group design, Secondary education, Student-centred instruction

INTRODUCTION

Background and International Evidence

In the past decade, much research has focused on how instructional design can move learners from passive reception toward purposeful cognitive participation. This concern is especially important in secondary education, where students are expected not only to remember information but also to interpret evidence, solve unfamiliar problems, communicate reasoning, and transfer knowledge across contexts. Conventional teacher-centred lessons can efficiently organise content, but when they dominate classroom time they may provide too few opportunities for students to retrieve prior knowledge, explain ideas, test interpretations, receive feedback, and revise misconceptions. Active learning addresses this limitation by requiring learners to do intellectually consequential work during instruction rather than merely listen or copy information (Prince, 2004; Michael, 2006). Its defining feature is not movement or entertainment; it is the deliberate activation of cognitive processes through tasks such as guided questioning, peer explanation, problem solving, discussion, inquiry, application, and reflection (Chi & Wylie, 2014; Driessen et al., 2020).

The international evidence base is substantial but not uniform. Prescribed active-learning designs have improved performance in introductory biology (Freeman et al., 2007), and a meta-analysis of 225 science, technology, engineering, and mathematics studies reported higher examination performance and lower failure rates under active learning (Freeman et al., 2014). Positive average effects have also been reported in humanities and social-science courses (Kozanitis & Nenciovici, 2023), while systematic evidence on achievement indicates that instructional structure, feedback, and opportunities for cognitive engagement are important correlates of learning (Schneider & Preckel, 2017). Earlier interactive-engagement research in physics likewise found stronger conceptual gains than traditional instruction (Hake, 1998). These findings align with cooperative, problem-based, project-based, inquiry, and guided-discovery approaches when the activities require explanation, reasoning, feedback, and individual accountability rather than visible activity alone (Alfieri et al., 2011; Chen & Yang, 2019; Gijbels et al., 2005; Slavin, 2014; Springer et al., 1999).

The theoretical rationale is equally strong. The ICAP framework predicts that learning generally improves as engagement moves from passive to active, constructive, and interactive modes, because learners increasingly generate inferences and co-construct understanding (Chi & Wylie, 2014; Driessen et al., 2020). From a self-determination perspective, learning environments that support competence, autonomy, and relatedness can strengthen sustained engagement and internalised motivation (Deci & Ryan, 2000; Ryan & Deci, 2020). Engagement research likewise conceptualises behavioural, emotional, and cognitive participation as interconnected pathways to educational success (Fredricks et al., 2004; Kahu, 2013). Active learning can activate these pathways when students perceive tasks as meaningful, manageable, and socially safe.

The Indonesian secondary-school literature demonstrates sustained interest in active and student-centred instruction, but much of the available evidence is tied to a specific strategy, subject, or instructional medium. For example, Arwalembun et al. (2024) examined a Learning Starts With a Question strategy in Grade 7 science, Oktavia et al. (2025) studied Wordwall-supported geography instruction in senior high school, and Aidah and Masitoh (2026) evaluated blended microlearning in a Grade 11 geography practicum. These studies are relevant because they show that quasi-experimental comparisons are feasible in Indonesian schools; however, their intervention-specific designs do not establish how a broadly defined active-learning package should be interpreted in another school context. The present study therefore addresses a narrower question: whether two Grade 10 classes at SMA Negeri 1 Tompaso displayed different immediate posttest outcomes under the reported instructional conditions. The available record identifies the school as a public high school but does not provide evidence for classifying the setting as rural, semi-rural, or non-metropolitan.

Unresolved Problems and Research Gap

Despite the generally positive literature, it remains unclear why active learning produces strong gains in some classrooms but weak or negligible gains in others. Andrews et al. (2011), for example, found that self-reported use of active learning by typical instructors was not necessarily associated with better learning. This apparent contradiction suggests that the label “active learning” is insufficient: instructional quality depends on the intellectual demand of tasks, the degree of guidance, the quality of peer interaction, alignment with assessment, and the teacher’s ability to diagnose and respond to emerging understanding. Student resistance can also undermine participation when learners do not understand the purpose of active tasks or perceive them as inefficient (Cavanagh et al., 2016; Deslauriers et al., 2019; Tharayil et al., 2018).

A second problem concerns the tendency to treat student-centred instruction and teacher guidance as opposites. Pure discovery can overload novice learners and allow misconceptions to persist, whereas guided discovery and inquiry are more consistently effective (Alfieri et al., 2011; Kirschner et al., 2006; Lazonder & Harmsen, 2016; Mayer, 2004). Thus, the general solution is not the removal of the teacher but the redesign of the teacher’s role—from primary transmitter to architect of tasks, facilitator of reasoning, provider of timely feedback, and monitor of participation. Black and Wiliam (2009) emphasised that formative assessment is central to such adaptive teaching because evidence of student thinking must inform the next instructional move.

A third problem concerns contextual transfer and research design. Much of the highest-profile evidence comes from higher education and STEM courses, whereas secondary-school instruction is shaped by adolescent development, timetables, curriculum coverage, assessment pressures, teacher authority, and class-level organisation. School-based syntheses report positive average effects for inquiry-oriented science instruction, especially when guidance is explicit (Furtak et al., 2012; Minner et al., 2010; Mueller et al., 2015; Schroeder et al., 2007), but local estimates remain dependent on the subject, teacher, class composition, outcome measure, and fidelity of implementation. Consequently, a comparison involving one intact class per condition can provide useful site-specific evidence, but it cannot by itself isolate a general treatment effect from class-level influences.

The literature supports three bounded conclusions. First, active learning has a positive average association with achievement, but implementation quality is a major moderator (Andrews et al., 2011; Freeman et al., 2014; Prince, 2004). Second, student-centred activity is most defensible when combined with teacher guidance and formative feedback rather than minimal guidance (Black & Wiliam, 2009; Kirschner et al., 2006; Lazonder & Harmsen, 2016). Third, problem-based, project-based, inquiry-based, discovery-oriented, cooperative, and flipped approaches share mechanisms such as generation, interaction, feedback, and application, but they are not operationally interchangeable (Chi & Wylie, 2014; Wijnia et al., 2024; Yew & Goh, 2016). A study that uses the umbrella term active learning must therefore state which features were actually implemented and which mechanisms remain theoretical.

The present gap is not whether active learning can work in principle; that question is already supported by a large international literature. The narrower gap concerns transparent, context-sensitive reporting of a broad active-learning condition in an Indonesian public secondary school. Existing Indonesian studies frequently examine a named strategy or technology in a defined subject (Aidah & Masitoh, 2026; Arwalembun et al., 2024; Oktavia et al., 2025), whereas the supplied record describes a combination of discussion, problem solving, group projects, discovery-oriented activity, and application tasks. The record also permits calculation of a standardised posttest difference and confidence interval, but it does not contain the subject domain, intervention duration, baseline statistics, teacher allocation, instrument coefficients, implementation-fidelity evidence, or ethics documentation. The contribution of this reconstruction is therefore contextual and reporting-oriented: it presents the available evidence transparently, quantifies the observed contrast, and makes the unresolved design boundaries explicit rather than claiming major theoretical novelty.

Design Principles and Theoretical Framework

Scientific literature points to a family of design principles rather than a single universally superior technique. First, active tasks should require generative processing: students should explain, compare, justify, predict, solve, create, or evaluate rather than merely repeat facts. Linton et al. (2014) showed that writing and peer discussion can contribute to conceptual learning when students externalise reasoning and encounter alternative ideas. Second, active learning should be guided. Meta-analyses of inquiry and discovery indicate that prompts, scaffolds, feedback, and strategic teacher intervention increase effectiveness (Alfieri et al., 2011; Lazonder & Harmsen, 2016). The inquiry cycle described by Pedaste et al. (2015) also illustrates that orientation, conceptualisation, investigation, conclusion, and discussion require coherent sequencing rather than unguided activity. Third, collaboration should include individual accountability and promotive interaction because unstructured group work can divide labour without producing substantive learning (Slavin, 2014; Springer et al., 1999).

Fourth, activities should be aligned with intended learning outcomes and assessment. Problem-based learning tends to show stronger effects when assessment captures application and understanding rather than isolated factual recall (Dochy et al., 2003; Gijbels et al., 2005; Hmelo-Silver, 2004). Project-based learning can support achievement and complex competencies when projects are authentic, cognitively demanding, and adequately scaffolded (Bell, 2010; Chen & Yang, 2019; Guo et al., 2020; Kokotsaki et al., 2016). Fifth, active learning should be embedded in a coherent lesson architecture that includes preparation, activation of prior knowledge, task engagement, monitoring, feedback, and consolidation. Evidence on increased course structure and active participation suggests that the design and consistency of implementation are at least as important as the label attached to the method (Eddy & Hogan, 2014). Flipped-classroom evidence similarly indicates that learning improves when class time is reorganised for purposeful activity and feedback, not merely when content is moved outside class (Strelan et al., 2020; van Alten et al., 2019).

These principles imply a practical model for secondary classrooms: the teacher introduces a focused problem or question, activates prior knowledge, organises students for explanation or collaborative problem solving, monitors reasoning, provides prompts and corrective feedback, and closes with synthesis and individual accountability. The interpretive model in Figure 2 presents the expected pathway from active-learning features through cognitive and motivational mechanisms to achievement.

Active learning is best understood as an umbrella construct that organises learning around student thinking. Its effectiveness depends on the type and quality of cognitive activity. In the ICAP taxonomy, passive engagement involves receiving information; active engagement involves manipulating information; constructive engagement requires generating outputs that extend beyond the presented material; and interactive engagement entails reciprocal co-construction (Chi & Wylie, 2014; Driessen et al., 2020). This hierarchy predicts why superficially active tasks can fail: copying notes, completing routine worksheets, or dividing project labour may create visible activity without constructive reasoning. Effective active learning instead elicits explanations, predictions, comparisons, arguments, and revisions.

Constructivist accounts explain that learners interpret new information through prior knowledge, but they do not imply that novices should discover disciplinary structures without support. Mayer (2004) and Kirschner et al. (2006) cautioned that minimally guided instruction can overload working memory and produce inefficient search. Meta-analytic evidence reconciles the debate: unassisted discovery is often inferior to direct instruction, whereas guided discovery can outperform other approaches when learners receive prompts, feedback, examples, or scaffolds (Alfieri et al., 2011). Inquiry learning follows a similar pattern; guidance is positively associated with learning outcomes, particularly for younger or less experienced learners (Lazonder & Harmsen, 2016).

Motivational theory provides a complementary explanation. Active learning can support autonomy by giving students meaningful choices and intellectual responsibility, competence by making progress visible through feedback, and relatedness through collaborative work (Deci & Ryan, 2000; Ryan & Deci, 2020). Yet these benefits are not automatic. Students may resist unfamiliar activities, particularly when they believe listening is easier or when the teacher does not explain how the activity supports learning. Deslauriers et al. (2019) found that students in an active classroom learned more while feeling that they learned less, illustrating a metacognitive mismatch between effort and perceived effectiveness. Cavanagh et al. (2016) showed that student buy-in is associated with engagement and performance, highlighting the importance of transparent rationale, supportive facilitation, and appropriately challenging tasks.

The present study therefore uses the framework as an interpretive model rather than as a tested mediation model. The active-learning condition is expected, in theory, to support cognitive engagement, retrieval and elaboration, peer-mediated feedback, and motivation or agency, but the study directly measures only the immediate cognitive posttest outcome. H1 is consequently restricted to the between-class difference in mean posttest scores; compatibility with ICAP, self-determination theory, engagement theory, guided instruction, or formative assessment must be discussed as a plausible explanation rather than empirical confirmation.

Study Objective and Hypothesis

The objective was to compare immediate cognitive posttest scores between one Grade 10 class taught under the reported active-learning condition and one Grade 10 class taught under conventional instruction at SMA Negeri 1 Tompaso. The study contributes site-specific evidence from an Indonesian public high school, a transparent re-analysis that reports the magnitude and precision of the observed posttest contrast, and an explicit account of information that could not be recovered from the supplied record. It does not propose or validate a new theory of active learning.

The directional hypothesis was H1: the active-learning class would obtain a higher mean posttest score than the conventional class. The hypothesis was informed by evidence that structured explanation, problem solving, interaction, and formative feedback can promote deeper processing (Chi & Wylie, 2014; Freeman et al., 2014; Linton et al., 2014). The empirical scope was limited to two intact classes, one class per condition, and to an immediate cognitive test. Engagement, motivation, critical thinking, retention, transfer, and implementation fidelity were not measured. Because baseline statistics and class-level replication were unavailable, the analysis tests an observed posttest contrast and should not be interpreted as an unbiased population-wide causal estimate.

METHOD

Research Design and Participants

The study used a quantitative quasi-experimental non-equivalent control-group design. Two intact Grade 10 classes were assigned at the class level to different instructional conditions: one class received active learning and one received conventional instruction. The source record states that pretests and posttests were administered, but only post-intervention group statistics were available for the present reconstruction. Accordingly, the inferential analysis is a student-level comparison of posttest scores rather than a baseline-adjusted estimate. Because there was only one intact class per condition, treatment status is completely confounded with class membership and may also be confounded with teacher, timetable, peer-composition, classroom-climate, and other class-level influences. The source record does not state whether the same teacher taught both classes or whether instructional time and assessment conditions were equivalent.

The population comprised Grade 10 students at SMA Negeri 1 Tompaso. Purposive sampling was used to select two intact classes, each containing 30 students, for a total analytic sample of 60. The criteria used to select the classes and judge their comparability were not reported. The supplied record also does not provide age, gender, prior-achievement profile, programme or stream, socioeconomic background, attendance, exclusions, attrition, or missing-data information. It is therefore not possible to determine whether every enrolled student completed every session and the posttest, whether scores were excluded, or whether differential participation affected the observed contrast.

Instructional Conditions

The source record describes the experimental condition as active learning involving student participation, discussion, problem solving, group projects, discovery-oriented activity, and application of knowledge. Conventional instruction is described as predominantly teacher-centred and one-directional. These descriptions identify broad pedagogical features but do not constitute a replicable intervention protocol. The academic subject, learning objectives, curriculum unit, number and duration of lessons, detailed sequence, group size, teacher preparation, materials, time on task, formative-assessment procedures, and control-condition lesson structure were not available. No fidelity-monitoring procedure, contamination check, or record of deviations and co-interventions was supplied.

Instrument and Data Collection

Student learning outcomes were measured with a cognitive achievement test reportedly administered as a pretest and posttest. The source record states that item validity was examined using product-moment correlations and that internal consistency was assessed using Cronbach's alpha before use. It does not report the number of items, academic content, blueprint, response format, score range, item examples, validation sample, item coefficients, decision thresholds, alpha coefficient, standard error of measurement, or evidence of expert content review. The psychometric adequacy and curricular alignment of the outcome measure therefore cannot be independently evaluated. The score categories previously labelled "Good" and "Sufficient" have been removed because the underlying cut-off system was not documented.

Data Analysis

Descriptive statistics summarised group size, mean, and standard deviation. The source reports calculated L values compared with a critical value of .1607, which is consistent with a Lilliefors-type normality assessment, although the exact procedure, p values, software, and software version were not reported. Homogeneity of variance was evaluated using Levene's test. H1 was evaluated with a two-tailed independent-samples t test at $\alpha = .05$; the use of a two-tailed test is retained even though the hypothesis was directional. Hedges' g was calculated from the reported means, standard deviations, and sample sizes using the pooled standard deviation and the small-sample correction $J = 1 - 3/[4(n1 + n2) - 9]$. The approximate 95% confidence interval for g used a normal approximation with $SE(g) = \sqrt{[(n1 + n2)/(n1n2) + g^2/\{2(n1 + n2 - 2)\}]}$. Student-level raw data were unavailable, so outlier diagnostics, score ranges, quartiles, distribution plots, robust analyses, and non-parametric sensitivity tests could not be conducted.

Validity, Reliability, and Ethical Considerations

The comparison group and common posttest provide a basis for describing the observed between-class difference, but internal validity remains vulnerable to selection, class-level confounding, teacher effects, history, maturation, differential exposure, testing, instrumentation, and implementation-fidelity threats. With one class per condition, the instructional condition cannot be statistically separated from the class effect. Statistical conclusion validity is supported by the internal consistency of the reported mean difference, standard error, t statistic, confidence interval, and effect size, but it is constrained by the absence of raw scores and baseline-adjusted analysis. External validity is limited to analytic transfer from a single-school, two-class context whose subject and implementation characteristics were not reported.

The supplied study record did not include an institutional ethics approval or exemption identifier, school-permission statement, parental or guardian consent procedure, student assent procedure, confidentiality plan, data-storage location, access-control arrangement, retention period, or disposal procedure. These matters remain unresolved and require author verification and editor-verifiable documentation before submission or publication. No personally identifying student information is reported in the manuscript.

Figure 1 summarises the research procedure and the analytical sequence supported by the available record.

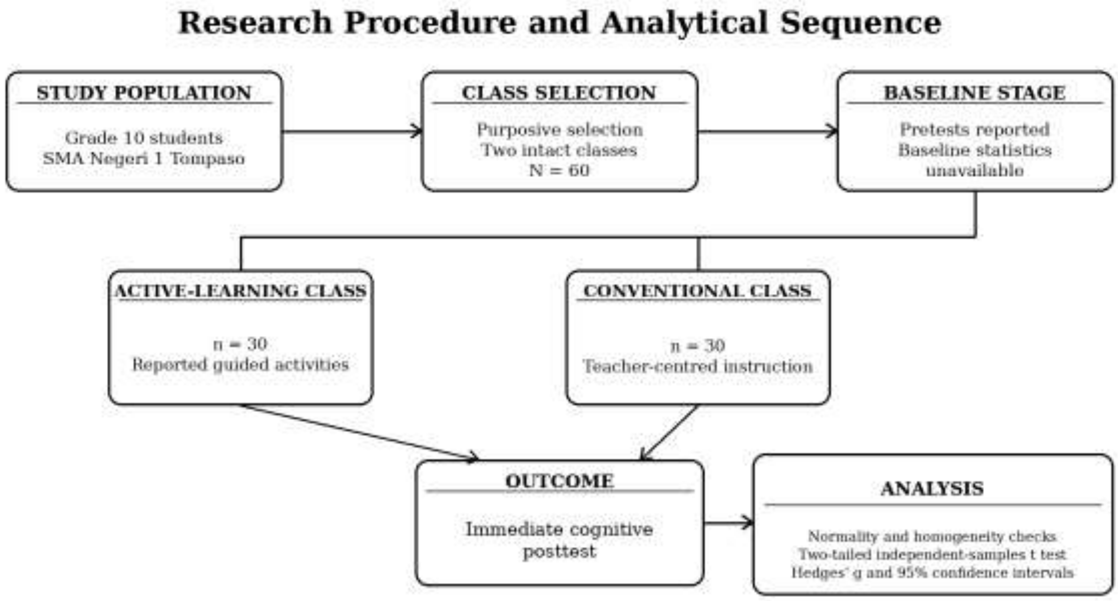

Figure 1. Research procedure and analytical sequence

RESULTS AND DISCUSSION

Posttest Achievement

The available post-intervention statistics show a large observed difference between the two classes. The 30 students in the active-learning class obtained a mean score of 83.56 (SD = 3.69), whereas the 30 students in the conventional class obtained a mean score of 72.40 (SD = 4.49). The raw mean difference was 11.16 points in favour of the active-learning class. The experimental-class scores were also less dispersed, although Levene's test did not indicate a statistically consequential variance difference.

The posttest separation is large relative to the within-class standard deviations. Nevertheless, its educational meaning cannot be fully interpreted because the subject, test blueprint, score range, proficiency standards, and baseline scores were not available. The analysis therefore treats the scores as continuous outcomes and does not assign qualitative achievement categories.

Table 1. Descriptive Statistics for Post-Intervention Learning Outcomes

Group	n	Mean	SD
Active learning	30	83.56	3.69
Conventional instruction	30	72.40	4.49

Assumption Tests and Hypothesis Test

The source record reports normality statistics of .1207 for the active-learning class and .1183 for the conventional class. Both values were below the reported critical value of .1607, and the source therefore treated the distributions as sufficiently normal for parametric analysis. Because the test was not explicitly named, it is described here as Lilliefors-type reporting rather than confirmed as a specific software implementation. Levene's test produced $F(1, 58) = 1.622$, $p = .208$, so the equal-variances specification was retained. Raw scores were not available for independent verification through histograms, quantile plots, outlier checks, or alternative distributional tests.

Table 2. Assumption-Test Results

Test	Group/Statistic	Calculated Value	Critical Value/p	Decision
Normality	Active learning	.1207	.1607	Normal
Normality	Conventional instruction	.1183	.1607	Normal
Levene homogeneity	F(1, 58)	1.622	p = .208	Equal variances retained

The two-tailed independent-samples t test indicated a statistically significant posttest difference between the two classes, $t(58) = 9.906$, $p < .001$. The estimated mean difference was 11.158 points (SE = 1.126), with a 95% confidence interval of approximately [8.90, 13.41]. Thus, the available posttest data support H1 at the descriptive between-class level: the active-learning class had the higher mean. They do not, however, establish that the instructional condition caused the difference because baseline equivalence and class-level replication were unavailable.

The pooled standard deviation was 4.11, Cohen's d was approximately 2.72, and the small-sample-corrected estimate was Hedges' $g = 2.68$ with an approximate 95% confidence interval of [1.98, 3.38]. This is an exceptionally large standardised contrast. The calculation is internally coherent with the reported summary statistics, but the magnitude must be interpreted as a description of these two classes rather than as an unbiased treatment effect. Plausible contributors include pre-existing class differences, teacher or timetable effects, differential exposure, close alignment between the intervention and test, novelty, measurement characteristics, or a genuinely strong instructional contrast.

Table 3. Independent-Samples Comparison and Standardised Posttest Contrast

t	df	p	Mean diff.	SE	95% CI	Hedges' g	95% CI for g
9.906	58	< .001	11.158	1.126	[8.90, 13.41]	2.68	[1.98, 3.38]

Interpretation and Comparison with Previous Research

The principal empirical finding is an immediate posttest contrast: the class exposed to the reported active-learning condition had a substantially higher mean than the class exposed to conventional instruction. This pattern is consistent with the central proposition of active-learning research that learning can benefit when students retrieve, generate, discuss, apply, and revise knowledge rather than primarily receive explanations (Michael, 2006; Prince, 2004). It is also compatible with the ICAP framework, which predicts stronger learning when engagement becomes constructive and interactive (Chi & Wylie, 2014; Driessen et al., 2020). Compatibility is not equivalent to mechanism confirmation, however, because the study did not measure cognitive engagement, discourse quality, feedback, motivation, or other proposed mediators.

A plausible explanation is that discussion, problem solving, and student explanation increased retrieval and elaboration, made misconceptions visible, and created opportunities for corrective feedback. Research on writing and peer discussion supports the value of externalising reasoning (Linton et al., 2014), and formative-assessment theory explains why visible student thinking can guide subsequent instruction (Black & Wiliam, 2009). In this manuscript, these mechanisms remain theoretically plausible interpretations only. The available evidence shows a score difference; it does not show which active-learning feature generated the difference or whether the reported activities were implemented consistently.

The observed direction is consistent with international evidence. Freeman et al. (2014) reported an average advantage for active learning in STEM, Hake (1998) documented stronger gains under interactive engagement, and Freeman et al. (2007) and Haak et al. (2011) showed that prescribed structure and active participation can improve course performance. Meta-analytic evidence on flipped instruction likewise identifies structured active learning as a central design feature (Strelan et al., 2020). School-level inquiry syntheses likewise report positive average effects when conceptual emphasis and guidance are present (Furtak et al., 2012; Minner et al., 2010; Mueller et al., 2015; Schroeder et al., 2007). The present result extends the descriptive pattern to two Grade 10 classes in one Indonesian public high school, but it does not provide the design leverage needed to claim the same average effect as multi-study syntheses.

Within Indonesian secondary education, the result is broadly compatible with studies of more specific interventions. Arwalembun et al. (2024) reported improved science activity and outcomes under a Learning Starts With a Question strategy; Oktavia et al. (2025) examined a Wordwall-supported geography condition; and Aidah and Masitoh (2026) reported higher outcomes under blended microlearning in a geography practicum. These comparisons help establish the national literature position, but they also sharpen the distinction of the present manuscript: the supplied record concerns a broad package of active-learning features rather than one named strategy, technology, or subject-specific design. Because the subject and protocol were not recovered, the present finding cannot be directly attributed to any one component of that package.

Problem-based, project-based, discovery, inquiry, and cooperative learning provide relevant explanatory literatures, but they should not be treated as if all were implemented in full. The source record supports only the presence of selected features—discussion, problem solving, group projects, discovery-oriented activity, participation, and application. Evidence from problem-based learning emphasises scaffolding and assessment alignment (Dochy et al., 2003; Gijbels et al., 2005; Hmelo-Silver, 2004; Yew & Goh, 2016); project-based learning requires coherent disciplinary structure (Bell, 2010; Chen & Yang, 2019; Guo et al., 2020; Kokotsaki et al., 2016); and guided discovery is more effective than unassisted discovery (Alfieri et al., 2011). The study should therefore be read as an evaluation of the reported combination, not as a test of each named model.

The exceptionally large Hedges' g should remain central to interpretation. Andrews et al. (2011) found that active learning as typically implemented was not necessarily associated with stronger learning, demonstrating that the label alone has little explanatory value. In the present study, a large posttest separation could reflect highly effective implementation, but it could also arise from baseline differences, class composition, teacher effects, unequal time or attention, test alignment, or other uncontrolled factors. The absence of fidelity and baseline evidence prevents adjudication among these explanations.

External research also indicates that students may learn more in active classrooms while feeling that they learn less because effortful processing is experienced as disfluency (Deslauriers et al., 2019), and that student buy-in is associated with engagement and performance (Cavanagh et al., 2016). These findings are useful for designing future implementation, but motivation, perception, and engagement were not measured in the present study. They should therefore be presented as issues for subsequent evaluation rather than as outcomes demonstrated at SMA Negeri 1 Tompaso.

Interpretive Theoretical Framework

The theoretical contribution is interpretive rather than confirmatory. Figure 2 organises the reported instructional features around possible pathways involving cognitive engagement, retrieval and elaboration, peer feedback, and motivation or agency. The observed posttest contrast is compatible with those pathways, but no mediator was measured and no mediation analysis was conducted. The manuscript therefore does not validate ICAP, self-determination theory, engagement theory, or formative-assessment mechanisms; it uses them to formulate testable explanations for future studies.

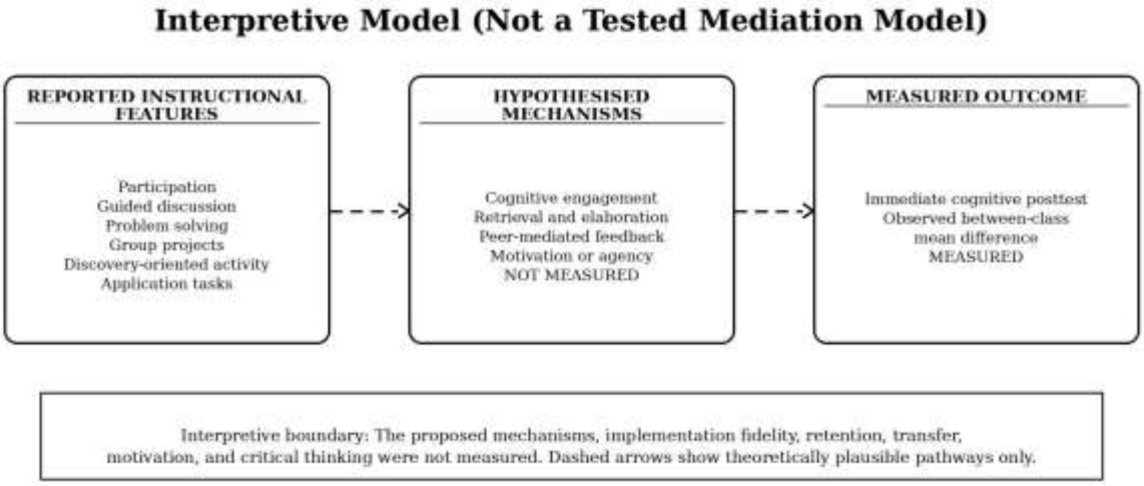

Figure 2. Interpretive Model for the Observed Posttest Contrast

The framework also supports a qualified view of constructivist instruction. Learners benefit from opportunities to construct and revise understanding, but teacher guidance remains essential. Evidence against minimally guided instruction (Kirschner et al., 2006; Mayer, 2004) and evidence for guided inquiry (Lazonder & Harmsen, 2016) suggest that effective instruction combines concise explanation with cognitively demanding activity, diagnostic prompts, and timely feedback. Whether that balance was achieved in the present intervention cannot be verified because the lesson sequence and fidelity data were not supplied.

Practical Implications and Transferability

For classroom practice, the study provides a provisional rationale for piloting structured active-learning routines in comparable Grade 10 settings, not a mandate for wholesale adoption. Discussion, retrieval questions, collaborative problem solving, peer explanation, application tasks, and individual synthesis are plausible components when they are aligned with a specific learning objective. Schools should evaluate such routines with baseline measures, common assessments, observation of implementation, and evidence that all students—not only the most vocal participants—engage in the intended cognitive work.

For school leadership, the most defensible implication is to support careful instructional development and local evaluation. Teachers may need collaborative planning time, subject-specific examples, formative-assessment tools, observation, and coaching. Because the source record does not report the teacher, duration, resources, or implementation demands, the study cannot estimate feasibility, workload, cost, or sustainability. These dimensions should be documented before the approach is scaled.

At policy level, the result should be treated as hypothesis-generating local evidence. Curriculum and assessment systems can encourage explanation, application, and problem solving, but a single-school comparison cannot establish the effectiveness of a policy across subjects, regions, or student groups. Broader recommendations require multisite studies with multiple classes per condition, baseline adjustment, intervention protocols, and implementation data.

Transferability can be assessed only after the missing context is specified. The finding may be informative for settings that resemble the participating school in grade level, subject, curriculum demands, class size, teacher expertise, resources, assessment, and instructional culture, but those scope conditions are not currently documented. Accordingly, the manuscript supports analytic transfer as a question for replication rather than statistical generalisation to Indonesian secondary education as a whole.

Alternative Explanations, Limitations, and Future Research

The magnitude of the difference is the most striking feature of the analysis. Hedges' g = 2.68 is far larger than typical average effects reported in broad syntheses. Five alternative explanations require attention: pre-existing differences between the two intact classes; complete confounding of treatment with class and possibly teacher; close alignment between the test and active-learning activities; unequal time, attention, or novelty; and measurement characteristics or restricted score variability. Without pretest statistics, raw data, teacher information, and fidelity records, the contribution of these factors cannot be estimated.

The reported summary statistics are internally coherent: the mean difference, pooled variability, standard error, t statistic, confidence interval, and standardised effect agree mathematically. Reported normality and homogeneity checks also support the selected t-test specification. Nonetheless, statistical coherence cannot overcome the design limitation of one class per condition or substitute for missing intervention and measurement documentation. The result is best characterised as a promising, large, site-specific posttest association that warrants replication.

The study is limited by a sample of 60 students from two intact classes in one school, purposive class selection, no random assignment, one class per condition, unavailable baseline statistics, and unreported participant characteristics and missing-data procedures. The subject, curriculum unit, intervention duration, teacher allocation, activity sequence, time on task, fidelity, instrument specifications, score range, psychometric coefficients, and ethics and data-governance documentation were also unavailable. Only immediate cognitive posttest performance was analysed; engagement, motivation, critical thinking, retention, transfer, implementation process, feasibility, and student experience were not directly measured.

Future research should use multiple classes per condition, preferably across several schools, with random or carefully matched assignment. Pretest-adjusted regression or ANCOVA should be used when assumptions are satisfied, and multilevel modelling should account for students nested within classes when enough clusters are available. Researchers should preregister primary outcomes, publish a replicable intervention protocol, report participant flow and missing data, document validity and reliability, measure fidelity and contamination, and retain raw-score diagnostics. Delayed posttests and transfer tasks should assess durability and application, while mixed-methods process data could examine teacher moves, peer discourse, participation equity, student buy-in, and contextual feasibility.

CONCLUSION

This study compared immediate cognitive posttest achievement in two intact Grade 10 classes at SMA Negeri 1 Tompaso, one taught under a broadly defined active-learning condition and one under conventional instruction. The active-learning class obtained a substantially higher mean, and the reported summary statistics yielded a statistically significant, very large standardised posttest contrast. The study's contribution is contextual and methodological: it documents a local result, relates it cautiously to mechanism-based active-learning literature, reports effect magnitude and precision, and makes the limits of the available record explicit. Because baseline statistics, class-level replication, intervention fidelity, psychometric coefficients, teacher information, and ethics documentation were unavailable, the result should not be interpreted as a definitive causal estimate or generalised beyond comparable settings. Future research should use multiple classes and sites, baseline-adjusted analyses, replicable protocols, validated measures, fidelity checks, delayed outcomes, and direct measures of the proposed learning mechanisms.

REFERENCES

Aidah, Y., & Masitoh, F. (2026). The effect of blended microlearning soil analysis practicum on students' learning outcomes in Geography. *Indonesian Journal of Educational Development (IJED)*, *6*(4), 1471–1482. https://doi.org/10.59672/ijed.v6i4.5730

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? Journal of Educational Psychology, 103(1), 1–18. https://doi.org/10.1037/a0021017

Andrews, T. M., Leonard, M. J., Colgrove, C. A., & Kalinowski, S. T. (2011). Active learning not associated with student learning in a random sample of college biology courses. *CBE—Life Sciences Education*, *10*(4), 394–405. https://doi.org/10.1187/cbe.11-07-0061

Arwalembun, A., Megawati, R., & Akobiarek, M. (2024). Pengaruh strategi pembelajaran aktif tipe Learning Starts With A Question terhadap keaktifan dan hasil belajar IPA siswa kelas VII SMP YPK Hedam tahun ajaran 2018/2019. *Sulawesi Tenggara Educational Journal*, *4*(2), 60–67. https://doi.org/10.54297/seduj.v4i3.817

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

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment*, *Evaluation* and Accountability, 21(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Cavanagh, A. J., Aragón, O. R., Chen, X., Couch, B. A., Durham, M. F., Bobrownicki, A., Hanauer, D. I., & Graham, M. J. (2016). Student buy-in to active learning in a college science course. *CBE—Life Sciences Education*, *15*(4), ar76. https://doi.org/10.1187/cbe.16-07-0212

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

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

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

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, *116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, *13*(5), 533–568. https://doi.org/10.1016/S0959-4752(02)00025-7

Driessen, E. P., Knight, J. K., Smith, M. K., & Ballen, C. J. (2020). Demystifying the meaning of active learning in postsecondary biology education. *CBE—Life Sciences Education*, *19*(4), ar52. https://doi.org/10.1187/cbe.20-04-0068

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? CBE—Life Sciences Education, 13(3), 453–468. https://doi.org/10.1187/cbe.14-03-0050

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

Freeman, S., O'Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., Dirks, C., & Wenderoth, M. P. (2007). Prescribed active learning increases performance in introductory biology. *CBE—Life Sciences Education*, *6*(2), 132–139. https://doi.org/10.1187/cbe.06-09-0194

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*(3), 300–329. https://doi.org/10.3102/0034654312457206

Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research*, *75*(1), 27–61. https://doi.org/10.3102/00346543075001027

Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. *International Journal of Educational Research*, *102*, 101586. https://doi.org/10.1016/j.ijer.2020.101586

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, *332*(6034), 1213–1216. https://doi.org/10.1126/science.1204820

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

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

Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, *38*(5), 758–773. https://doi.org/10.1080/03075079.2011.598505

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

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

Kozanitis, A., & Nenciovici, L. (2023). Effect of active learning versus traditional lecturing on the learning achievement of college students in humanities and social sciences: A meta-analysis. *Higher Education*, *86*, 1377–1394. https://doi.org/10.1007/s10734-022-00977-8

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, *86*(3), 681–718. https://doi.org/10.3102/0034654315627366

Linton, D. L., Pangle, W. M., Wyatt, K. H., Powell, K. N., & Sherwood, R. E. (2014). Identifying key features of effective active learning: The effects of writing and peer discussion. *CBE—Life Sciences Education*, *13*(3), 469–477. https://doi.org/10.1187/cbe.13-12-0242

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? American Psychologist, 59(1), 14–19. https://doi.org/10.1037/0003-066X.59.1.14
Michael, J. (2006). Where's the evidence that active learning works? Advances in Physiology Education, 30(4), 159–167. https://doi.org/10.1152/advan.00053.2006
Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—What is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474–496. https://doi.org/10.1002/tea.20347
Mueller, A. L., Knobloch, N. A., & Orvis, K. S. (2015). Exploring the effects of active learning on high school students' outcomes and teachers' perceptions of biotechnology and genetics instruction. *Journal of Agricultural Education*, *56*(2), 138–152. https://doi.org/10.5032/jae.2015.02138
Oktavia, A. D., Tabbu, M. A. S., & Simpuruh, I. (2025). Learning with Wordwall: A quasi-experimental study on student engagement and learning outcomes in geography at a senior high school in Indonesia. *Information Technology Education Journal*, *4*(2), 94–105. https://doi.org/10.59562/intec.v4i2.7779
Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003
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
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*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860
Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565–600. https://doi.org/10.1037/bul0000098
Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T.-Y., & Lee, Y.-H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching*, *44*(10), 1436–1460. https://doi.org/10.1002/tea.20212
Slavin, R. E. (2014). Cooperative learning and academic achievement: Why does groupwork work? Anales de Psicología, 30(3), 785–791. https://doi.org/10.6018/analesps.30.3.201201
Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, *69*(1), 21–51. https://doi.org/10.3102/00346543069001021
Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educational Research Review*, *30*, 100314. https://doi.org/10.1016/j.edurev.2020.100314
Tharayil, S., Borrego, M., Prince, M., Nguyen, K. A., Shekhar, P., Finelli, C. J., & Waters, C. (2018). Strategies to mitigate student resistance to active learning. *International Journal of STEM Education*, *5*, Article 7. https://doi.org/10.1186/s40594-018-0102-y
van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review*, *28*, 100281. https://doi.org/10.1016/j.edurev.2019.05.003

Wijnia, L., Noordzij, G., Arends, L. R., Rikers, R. M. J. P., & Loyens, S. M. M. (2024). The effects of problem-based, project-based, and case-based learning on students' motivation: A meta-analysis. *Educational Psychology Review*, *36*(1), Article 29. https://doi.org/10.1007/s10648-024-09864-3

Yew, E. H. J., & Goh, K. (2016). Problem-based learning: An overview of its process and impact on learning. *Health Professions Education*, *2*(2), 75–79. https://doi.org/10.1016/j.hpe.2016.01.004