Journal of General Education and Humanities
Vol. 5, No. 4, August 2026, pp. 5595 – 5612, https://doi.org/10.58421/gehu.v5i4.1963
ISSN 2963-7147

The Pedagogical Role of Verification in Supporting Mathematical Justification within the ACTIVER Learning Model

Sri Lestari[1], Kartono[2], Masrukan[3], Emi Pujiastuti[4]
[1,2,3,4]Universitas negeri Semarang, Semarang Indonesia

Article Info

Article history:

Received 2026-08-01
Revised 2026-08-11
Accepted 2026-08-26

Keywords:

Differentiated Instruction
Vocational Education
Mathematical Justification
Mathematical Reasoning
Verification

ABSTRACT

Mathematical justification requires students to examine and defend mathematical claims rather than merely produce correct answers. However, vocational students often experience difficulty constructing defensible mathematical arguments, while verification may be limited to checking final answers. This study examines the pedagogical role of Verify Understanding within ACTIVER, a differentiated mathematical reasoning learning model developed using the ADDIE framework. The broader study was conducted from March 2025 to May 2026. The comparative implementation involved 216 Grade X vocational high school students across six intact classes and was delivered over eight meetings. Quantitative evidence included posttest achievement, mastery proportions, and normalized gain, while implementation evidence was examined in relation to predefined justification indicators. In the limited trial, the ACTIVER group achieved a higher posttest mean than the comparison group (83.86 vs. 64.39), $t(70) = 17.232$, $p < .001$, $d = 4.06$. In the broader trials, experimental means of 84.44 and 85.39 exceeded comparison means of 69.56 and 71.00, respectively ($p < .001$). Mastery proportions and learning gains consistently favored ACTIVER. Verification provided structured opportunities to examine evidence, respond to critique, and revise mathematical reasoning. Its contribution should nevertheless be interpreted within the integrated ACTIVER sequence rather than as an independently tested intervention.

Corresponding Author:

Kartono
Doctoral Program in Mathematics Education, Universitas Negeri Semarang, Semarang, Indonesia
Email: mailto:kartono.mat@mail.unnes.ac.id

1. INTRODUCTION

Mathematical reasoning is a fundamental component of mathematical competence because it enables students to identify relevant information, recognize relationships, select appropriate strategies, formulate generalizations, evaluate procedures, and establish why mathematical conclusions should be accepted [1], [2].

 https://doi.org/10.58421/gehu.v5i4.1963

Mathematics learning should therefore extend beyond procedural accuracy and producing correct numerical answers. Students also need opportunities to understand mathematical relationships, evaluate the validity of procedures, and communicate defensible reasons for their conclusions. Within this broader competence, mathematical justification occupies an important position. In the present study, mathematical reasoning refers to the broader process through which students analyse information, construct mathematical relationships, formulate generalizations, and draw defensible conclusions. Mathematical justification refers to establishing why a mathematical claim, procedure, representation, generalization, or conclusion is acceptable by connecting it with relevant evidence and logical reasoning [3], [4]. Justification is therefore treated as both a component of mathematical reasoning and an observable classroom practice.

Several related concepts need to be distinguished. Mathematical argumentation refers to the communicative process through which claims and supporting reasons are proposed, questioned, evaluated, defended, and revised through interaction. Proof represents a more formal form of mathematical justification that satisfies accepted standards of logical validity and generality. Verification, as conceptualized in this study, refers to the instructional process through which students examine the correctness, consistency, applicability, and logical support of mathematical strategies, solutions, and generalizations. This distinction is pedagogically important because obtaining a correct answer does not necessarily demonstrate valid mathematical reasoning. Students may obtain an expected numerical result through an inappropriate procedure, an unsupported assumption, or an unnoticed computational error. A sequence of calculations describes what a student has done, whereas mathematical justification explains why the procedure and conclusion should be accepted. Research consequently emphasizes learning environments in which students are required to explain, test, question, and defend mathematical ideas rather than depend solely on teacher confirmation [5], [4], [6].

The problem is particularly relevant in vocational education. Vocational high school students are expected to use mathematics in occupational, technical, economic, and everyday contexts. Such situations require students not only to perform calculations but also to interpret quantitative information, identify relationships, select appropriate representations, evaluate whether solutions are reasonable, and justify decisions mathematically. Previous studies involving vocational students have nevertheless identified difficulties in interpreting mathematical situations, constructing relevant relationships, drawing valid conclusions, and demonstrating adaptive mathematical reasoning [7], [8], [9].

The preliminary analysis underlying the development of ACTIVER similarly indicated that students' mathematical reasoning was not yet optimal, particularly in generalization and justification. The instructional challenge was therefore not merely to increase students' ability to produce correct answers, but to create structured opportunities for them to examine why mathematical conclusions were valid and defend their reasoning when questioned.

Previous studies offer several relevant perspectives. Comparing mathematical cases can help students identify similarities, differences, and structures from which generalizations emerge. However, identifying a pattern and explaining why it should be accepted involve different reasoning demands. Mathematical justification requires students to move beyond recognizing that a relationship works in selected examples toward explaining the mathematical structure that supports its validity [10], [11].

Collaborative mathematical argumentation provides another important perspective. When students publicly communicate their reasoning, others can examine their claims. Students may question assumptions, request clarification, compare representations, challenge unsupported conclusions, and revise incomplete arguments. Such interaction can support the co-construction and refinement of mathematical reasoning [12], [13], [14].

Teacher actions are equally influential. Questions that require students to explain strategies, compare cases, generalize patterns, examine assumptions, and defend conclusions can facilitate mathematical reasoning [15], [16], [17]. Fransisco further demonstrates that collective mathematical argumentation requires deliberate teacher support [18]. However, support needs to preserve students' intellectual responsibility. Teacher intervention can become counterproductive when teachers supply mathematical warrants too quickly and the authority for establishing validity returns to the teacher rather than remaining with students' evidence and reasoning [19].

Student heterogeneity creates an additional instructional challenge. Students enter mathematics classrooms with different levels of prerequisite knowledge, readiness, communication ability, interests, and ways of accessing mathematical information. Differentiated instruction responds to this diversity by adjusting access to information, learning processes, representations, scaffolding, and response formats while maintaining meaningful common learning objectives [20], [21]. Within mathematical justification, differentiated support can therefore include guiding questions, visual representations, structured argument prompts, peer support, or extension tasks without reducing the intellectual requirement to provide valid mathematical reasoning.

Although previous research has investigated mathematical reasoning, proof, argumentation, teacher questioning, retrieval practice, and differentiated instruction, these strands are frequently addressed separately. Less attention has been given to an instructional architecture in which verification is explicitly positioned after conceptual construction, generalization, and retrieval and is integrated with differentiated support for mathematical reasoning, particularly within vocational mathematics education.

This gap matters because verification can easily be reduced to checking whether a final numerical answer is correct. Verification as conceptualized in the present study extends beyond answer checking. It involves examining evidence, testing generalizations, explaining mathematical transformations, responding to questions, defending claims, and revising incomplete reasoning.

The ACTIVER Learning Model was developed to address this need. ACTIVER consists of six interconnected phases: Initial Affirmation, Create Understanding, Transform Meaning, Internalize Memory, Verify Understanding, and Evaluate and Reflect. The originality claimed for ACTIVER does not lie in treating questioning, checking, retrieval, differentiation, argumentation, or reflection as individually new practices. Rather, its contribution lies in systematically sequencing them for instruction.

Analysis is operationalized through identifying important facts and information (A1), determining variables and relationships among variables (A2), and selecting appropriate strategies (A3). Generalization involves observing and comparing cases (G1), identifying underlying patterns (G2), and formulating general mathematical statements (G3). Justification involves presenting evidence systematically and logically (J1) and defending mathematical arguments in response to questions or criticism (J2).

Within this sequence, Verify Understanding occupies a particular pedagogical position. Before students are expected to defend mathematical conclusions, they have activated prerequisite knowledge, developed conceptual Understanding, analysed information, compared cases, constructed generalizations, and retrieved relevant concepts and procedures. Verification therefore does not occur in a conceptual vacuum. Students use previously constructed and reactivated knowledge as resources for judging whether mathematical claims should be accepted.

Retrieval in ACTIVER is not intended merely to reinforce memorization of formulas. During Internalize Memory, students reactivate relevant concepts, relationships, and procedures so they can later use them in reasoning and verification. Research on retrieval practice indicates that retrieving previously learned information can strengthen subsequent access to knowledge and may support performance on complex mathematical tasks [22], [23]. Because the present study did not employ a delayed posttest, however, no claim is made that ACTIVER conclusively established long-term retention.

Accordingly, this study examines the pedagogical role of Verify Understanding within the integrated ACTIVER Learning Model and addresses three research questions:

1. How is mathematical justification operationalized through Verify Understanding?
2. What verification activities provide opportunities for students to examine, defend, and revise mathematical arguments?
3. How do implementation and effectiveness findings inform the pedagogical contribution of verification within the integrated ACTIVER sequence?

The study contributes theoretically by positioning verification as a bridge among conceptual Understanding, generalization, knowledge retrieval, and mathematical justification. Pedagogically, it provides a structured pathway for shifting classroom attention from solution production to examining mathematical validity. Practically, it offers teachers a framework for combining questioning, differentiated scaffolding, peer critique, evidence evaluation, and argument revision. The study does not claim that Verify Understanding was experimentally isolated from the other ACTIVER phases; it interprets its contribution within the integrated learning sequence.

2. METHOD

2.1 Research Design

This article reports a focused secondary mixed-method product-evaluation analysis derived from a broader research and development study. The broader study employed the Analysis, Design, Development, Implementation, and Evaluation (ADDIE) framework to develop and evaluate the ACTIVER Learning Model [24].

During the Analysis phase, the study examined students' initial mathematical reasoning, learning readiness, instructional needs, curriculum requirements, and existing classroom practices. During Design, the rationale, objectives, syntax, social system, principles of reaction, supporting system, instructional effects, and accompanying effects of ACTIVER were formulated. During development, the model and supporting instructional materials were produced, validated by experts, and revised. Implementation included limited and broader classroom trials. Evaluation was conducted throughout development to examine the validity, practicality, and effectiveness of the model.

The present article combines quantitative product-effectiveness evidence with implementation evidence specifically related to Verify Understanding. Individual ACTIVER phases were not removed or experimentally manipulated; consequently, the study evaluates ACTIVER as an integrated instructional model rather than estimating the independent causal effect of verification.

2.2 Research Setting and Timeline

The research was conducted from March 2025 to May 2026. The preliminary analysis of students' mathematical reasoning ability was conducted in March 2025. The model design and expert-validation stages were conducted from May to September 2025. Classroom implementation, including limited and broader trials, evaluation, dissemination, and implementation in additional schools, was conducted from October 2025 to May 2026.

The primary classroom trials were conducted at SMKN 1 Demak, Central Java, Indonesia. Dissemination was subsequently conducted through the Vocational High School Mathematics Teachers' Association (MGMP Matematika SMK) of Demak Regency, followed by implementation at SMK Negeri 2 Demak and SMK Negeri 1 Sayung.

ACTIVER was implemented over eight classroom meetings, with each meeting lasting approximately 90 minutes. The mathematical content included systems of linear equations in three variables and systems of linear inequalities in two variables.

2.3 Participants, Sampling, and Instructional Personnel

The comparative classroom trials involved 216 Grade X vocational high school students from six intact classes, with 36 students in each class. Because the intervention was implemented within regular school schedules, the intact class rather than the individual student constituted the unit of classroom participation. Existing class membership was retained, and individual students were not randomly reassigned across classes.

In the limited trial, X AKL 3 served as the experimental group and X AKL 1 as the comparison group. The experimental class received ACTIVER, whereas the comparison class received Think Pair Share instruction.

Large-scale Trial 1 involved X PM1 as the ACTIVER group and X PM2 as the comparison group receiving Jigsaw instruction. Large-scale Trial 2 involved X MPLB1 as the ACTIVER group and X MPLB3 as the comparison group receiving Problem-Based Learning. ACTIVER was therefore compared with active instructional approaches rather than exclusively with teacher-centred instruction.

Instructional personnel included the researcher, who implemented the model in the limited trial; two mathematics teachers who participated in the broader classroom trials; two mathematics teachers who participated in dissemination and subsequent implementation; and two mathematics teachers who served as classroom observers. Six expert validators were involved in evaluating the model and its supporting materials.

2.4 ACTIVER Instructional Procedure

ACTIVER integrates differentiated learning with the development of analysis, generalization, and justification.

Table 1. ACTIVER phases and their primary instructional functions

Phase	Main activities	Relationship with mathematical reasoning
Initial Affirmation	Initial preassessment, creating a positive classroom atmosphere, affirmation, and motivation	Establishes readiness and provides information for differentiated support
Create Understanding	Concept mapping, differentiated access to information, activation of prerequisite knowledge, and acquisition of new information	Supports A1–A3 through identification and interpretation of relevant mathematical information
Transform Meaning	Group investigation, analysis and processing of information, concept generalization, and scaffolding	Supports analysis and G1–G3
Internalize Memory	Retrieval practice, individual exercises, and reinforcement of previously constructed knowledge	Reactivates conceptual resources required for subsequent reasoning
Verify Understanding	Presentation, feedback, examination of conceptual correctness, testing of solutions and generalizations, and strengthening of justification	Provides direct opportunities for J1 and J2
Evaluate and Reflect	Formative assessment, reflection on learning strategies, reinforcement, and appreciation	Supports evaluation and metacognitive awareness

Differentiated support was integrated throughout the sequence by varying access to information, representations, scaffolding, guiding questions, grouping arrangements, and response formats according to students' readiness while maintaining common mathematical reasoning objectives.

2.5 Mathematical Reasoning Indicators

The principal quantitative outcome was mathematical reasoning, operationalized through three dimensions and eight research subindicators. The focused analysis reported in this article gives particular attention to J1 and J2 because these indicators are directly associated with Verify Understanding.

Table 2. Mathematical reasoning indicators

Code	Dimension	Research subindicator
A1	Analysis	Identifying important facts and information from a problem
A2	Analysis	Determining variables and relationships among variables
A3	Analysis	Selecting an appropriate mathematical strategy or formula
G1	Generalization	Observing and comparing several cases or examples
G2	Generalization	Identifying underlying patterns or similarities
G3	Generalization	Formulating a general rule, formula, or statement based on the identified pattern
J1	Justification	Presenting mathematical evidence or proof systematically and logically
J2	Justification	Defending an argument in response to questions or criticism

2.6 Data Sources and Research Instruments

Data were collected using mathematical reasoning tests, expert-validation sheets, classroom implementation observation sheets, teacher and student response questionnaires, interviews, and instructional documentation.

The mathematical reasoning test was designed to assess analysis, generalization, and justification according to indicators A1–A3, G1–G3, and J1–J2. Classroom observations documented implementation of the ACTIVER phases and relevant teacher–student learning activities. Teacher and student questionnaires and interviews examined practicality and responses to implementation. Instructional documentation provided additional evidence concerning solution checking, mathematical explanation, testing of generalizations, questioning, peer critique, and argument revision.

The final mathematical reasoning test consisted of 20 scored items or subitems that had undergone empirical item-quality analysis before use in the classroom trials.

2.7 Instrument Validity and Reliability

The mathematical reasoning test was subjected to empirical analysis covering item validity, overall reliability, difficulty level, and discrimination power. Item validity was examined using Product Moment correlations. Across the 20 scored items, correlation coefficients ranged from .92 to .99, indicating strong relationships between individual item scores and overall test performance. The instrument obtained an overall reliability coefficient of .99, indicating very high reliability.

Item-difficulty indices ranged from .29 to .56. Nineteen items were categorized as having moderate difficulty, while one item was categorized as difficult. Discrimination indices ranged from .54 to .76, and all 20 items were classified as having good discrimination and were retained for use in the study.

Table 3. Summary of mathematical reasoning test quality

Psychometric aspect	Result	Interpretation
Number of scored items	20	—
Product Moment correlations	.92–.99	Strong item validity
Overall reliability coefficient	.99	Very high reliability
Difficulty index	.29–.56	19 moderate; 1 difficult
Discrimination index	.54–.76	Good
Final item decision	20 retained	Acceptable for use

The results indicate that the test demonstrated adequate empirical quality for measuring students' mathematical reasoning in the classroom trials.

2.8 Product Validity and Practicality

Six expert validators assessed the ACTIVER Learning Model and its supporting instructional materials using a five-point rating scale. The model obtained an average validity score of 4.32 out of 5, while its supporting instructional materials obtained scores ranging from 4.20 to 4.49. These results indicated that the model and its supporting materials met the established validity criteria.

The study examined practicality through implementation of the ACTIVER syntax and teacher and student responses during classroom trials. The model was categorized as highly practical because the instructional phases could be implemented as intended and teacher and student responses to the learning process were positive.

2.9 Analysis of Verification Evidence

Evidence related specifically to verification was examined from classroom implementation records, observation notes, interview information, and instructional documentation available from the broader development study.

The analysis focused on instructional episodes associated with J1 and J2 and was organized descriptively around five pedagogical functions:

1. checking solutions against original mathematical conditions;
2. explaining the validity of mathematical transformations;
3. examining patterns and mathematical generalizations;
4. responding to teacher or peer questions and criticism; and
5. revising incomplete or unsupported mathematical reasoning.

This analysis aimed to identify how mathematical justification was operationalized within the instructional sequence, not to construct a longitudinal discourse score for individual students. Because the original development study was not designed as a fine-grained classroom-discourse investigation, the available records did not permit systematic student-level coding of argumentation development across all eight meetings or calculation of intercoder reliability. Qualitative findings are therefore interpreted as implementation evidence rather than as a quantified developmental trajectory of individual students' arguments.

2.10 Quantitative Data Analysis

A score of 80 was used as the mastery criterion. Posttest differences between experimental and comparison groups were analysed using independent-samples t-tests after examining assumptions of normality and homogeneity. Shapiro–Wilk tests were used to examine normality, and Levene's test was used only to examine equality of variances; Levene's significance values were not interpreted as evidence of instructional effectiveness. Differences in mastery proportions were examined using Pearson's Chi-Square and Fisher's Exact Test where appropriate. For the limited trial, the N-Gain distribution in the experimental group did not satisfy the normality assumption; therefore, the Mann–Whitney

U test was used to compare learning gains. For both broader trials, N-Gain scores were compared using independent-samples t-tests. Statistical significance was evaluated at $\alpha = .05$. Cohen's d was used descriptively to represent standardized between-group differences where relevant. Because the research involved intact classes rather than individual random assignment, interpret effect-size estimates cautiously.

2.11 Ethical Considerations

The researchers obtained formal research permission from the participating schools through official institutional letters before data collection and classroom implementation. The researchers coordinated research activities with school authorities and mathematics teachers to ensure implementation remained compatible with the regular instructional program. The researchers maintained participant confidentiality throughout data collection, analysis, and reporting. Researchers replaced students' identities with codes in research records and reported data so individual participants could not be identified. Findings are reported at group or classroom level rather than through personally identifiable student information.

3. RESULTS

3.1 Validity and Practicality of ACTIVER

Expert validation indicated that ACTIVER satisfied the established validity criteria. The model obtained an average validity score of 4.32 out of 5, while supporting materials received scores ranging from 4.20 to 4.49. Classroom implementation across eight meetings and positive teacher and student responses indicated that the model was highly practical in the participating instructional settings. These product-evaluation results provided the basis for proceeding to limited and broader effectiveness trials.

3.2 Verification Activities Supporting Mathematical Justification

Five recurring instructional functions were identified within Verify Understanding: checking mathematical solutions, explaining transformations, examining generalizations, responding to critique, and revising arguments.

Table 4. Verification activities and mathematical justification

Verification activity	Expected reasoning activity	Indicator
Checking solutions against original conditions	Establishing whether a proposed solution satisfies all relevant mathematical conditions	J1
Explaining transformations	Connecting procedures with mathematical properties or equivalent operations	J1
Examining generalizations	Determining whether evidence adequately supports a broader mathematical claim	J1
Responding to questions or criticism	Clarifying evidence and defending the logic of a conclusion	J2
Revising mathematical arguments	Correcting unsupported claims, logical gaps, or incomplete explanations	J1–J2

These activities created explicit opportunities for students to examine why mathematical solutions and generalizations should be accepted rather than treating a correct final answer as the sole criterion for completion.

However, the available implementation records did not allow reliable quantification of how many individual students progressed from procedural checking to more sophisticated forms of justification across the eight meetings.

3.3 **Limited-Trial Effectiveness**

The limited trial involved 36 students in the experimental class and 36 students in the comparison class.

The experimental group obtained a posttest mean of 83.86 ($SD = 5.066$), whereas the comparison group obtained a mean of 64.39 ($SD = 4.506$). The difference was statistically significant, $t(70) = 17.232, p < .001$. The standardized between-group difference was very large, $d \approx 4.06$.

Mastery attainment showed the same direction. Thirty of 36 students in the experimental group (83.3%) achieved the established mastery criterion, whereas no student in the comparison group reached the criterion. The difference in mastery proportions was statistically significant, $\chi^2(1) = 51.429, p < .001$.

The experimental-group N-Gain distribution did not satisfy the normality assumption. Consequently, a Mann–Whitney *U* test was used. The ACTIVER group obtained a mean rank of 54.50, compared with 18.50 in the comparison group. The difference was statistically significant, $U = 0, Z = -7.314, p < .001$.

Thus, the limited-trial findings showed a consistent pattern across final achievement, mastery attainment, and measured learning improvement.

3.4 Large-Scale Trial 1

Large-scale Trial 1 compared X PM1, which received ACTIVER, with X PM2.

Table 5. Posttest descriptive statistics for Large-scale Trial 1

Group	N	Mean	SD
Experimental (X PM1)	36	84.44	2.999
Comparison (X PM2)	36	69.56	3.102

Levene's test yielded $p = .739$, indicating that the equal-variance assumption was satisfied. The independent-samples t-test showed a statistically significant difference, $t(70) = 20.701, p < .001$, with a mean difference of approximately 14.89 points. Mastery attainment was 94.4% in the experimental group and 0% in the comparison group. The difference in mastery proportions was significant, $\chi^2(1) = 64.421, p < .001$.
Mean N-Gain was 0.7575 in the ACTIVER group and 0.5322 in the comparison group.
The difference was statistically significant, $t(70) = 32.839, p < .001$.

3.5 Large-Scale Trial 2

Large-scale Trial 2 compared X MPLB1, which received ACTIVER, with X MPLB3.

Table 6. Posttest descriptive statistics for Large-scale Trial 2

Group	N	Mean	SD
Experimental (X MPLB1)	36	85.39	2.567
Comparison (X MPLB3)	36	71.00	2.878

Levene's test yielded p = .508, supporting the equal-variance assumption. The independent-samples comparison produced a statistically significant difference, $t(70) = 22.386, p < .001$, with a mean difference of approximately 14.39 points.

Mastery attainment reached 97.2% in the experimental class and 0% in the comparison class. The difference was statistically significant, $\chi^2(1) = 68.108, p < .001$. Mean N-Gain was 0.7694 in the ACTIVER group compared with 0.5486 in the comparison group. The difference was statistically significant, $t(70) = 36.945, p < .001$.

3.6 Summary of Effectiveness Evidence

Table 7. Summary of ACTIVER effectiveness

Trial	Experimental posttest	Comparison posttest	Experimental mastery	Comparison mastery	Learning-gain evidence
Limited trial	83.86	64.39	83.3%	0%	Mean rank 54.50 vs. 18.50; *p* < .001
Large-scale Trial 1	84.44	69.56	94.4%	0%	0.7575 vs. 0.5322; *p* < .001
Large-scale Trial 2	85.39	71.00	97.2%	0%	0.7694 vs. 0.5486; *p* < .001

Across all three comparative settings, ACTIVER groups demonstrated higher posttest performance, higher mastery attainment, and stronger measured learning gains than their respective comparison groups.

3.7 Dissemination and Subsequent Implementation

Following the classroom trials, dissemination was conducted through the Vocational High School Mathematics Teachers' Association (*MGMP Matematika SMK*) of Demak Regency. Subsequent implementation at SMK Negeri 2 Demak and SMK Negeri 1 Sayung indicated that the planned learning objectives could also be achieved outside the primary development school.

Because these dissemination settings were not included in the principal comparative statistical Design, these findings are interpreted as evidence of applicability and implementation feasibility, rather than as additional experimental effectiveness tests.

4. DISCUSSION

The Verification of Understanding phase is a distinctive component of the ACTIVER Learning Model that gives students explicit opportunities to develop mathematical justification. Through this phase, students examine strategies, test solutions and generalizations, provide evidence, respond to questions, defend conclusions, and revise incomplete arguments.

4.1 Interpretation of the Main Findings

Three findings are central to the study.

First, Verify Understanding provided a designated instructional space in which mathematical validity became an explicit learning concern. Students were provided opportunities to examine solutions, explain transformations, test generalizations, present evidence, respond to questions, and revise mathematical arguments. Second, the integrated ACTIVER model consistently produced higher overall mathematical reasoning outcomes than the comparison instruction across the limited and broader trials. Third, the evidence for effectiveness did not depend on a single statistical indicator. Posttest achievement, mastery proportions, and learning gains consistently showed the same direction.

The convergence of these results suggests that ACTIVER supported not only final performance but also a broader distribution of mastery and stronger measured development during instruction.

4.2 From Procedural Verification to Mathematical Justification

The pedagogical significance of verification lies in the distinction between confirming an answer and establishing why it is mathematically valid.

Substituting values into an equation, for example, can identify whether an obtained result satisfies an original equation. However, this remains largely procedural unless students also explain why satisfaction of all relevant conditions validates the solution. Similarly, testing several numerical examples can increase confidence in a conjecture, but empirical confirmation does not establish a universal mathematical statement. Students need to connect examples to mathematical structure to move toward more defensible generalization.

This distinction aligns with research indicating that students can formulate claims and identify patterns but have greater difficulty constructing the warrants needed to justify them [3], [10], [25].

The available implementation evidence should not be interpreted as showing that every student automatically progressed from procedural checking to sophisticated deductive reasoning. Rather, Verify Understanding created instructional opportunities to support such progression.

4.3 Relationship among Comparing, Generalizing, and Justifying

The ACTIVER sequence establishes a deliberate relationship between generalization and justification. During Transform Meaning, students compare cases, identify mathematical structures, and formulate broader relationships. During Verify Understanding, students examine those emerging generalizations.

This relationship aligns with Öz and Çiftci, who conceptualize comparing, generalizing, and justifying as interconnected reasoning activities [11]. Widjaja similarly demonstrates that comparison can support the formation of generalizations, but justification requires students to articulate why an identified relationship should be accepted. The pedagogical shift is therefore from asking only what pattern has been identified toward asking why the pattern is mathematically defensible [10].

4.4 Questioning, Peer Critique, and Public Argumentation

Teacher questioning provides one mechanism for making implicit reasoning explicit. Questions about the appropriateness of a strategy, the mathematical properties underlying a transformation, the adequacy of evidence, the applicability of a generalization, or possible counterexamples require students to make the basis of their reasoning visible.

This interpretation is consistent with Mata-Pereira and da Ponte and Mukuka, who emphasize teacher actions that facilitate mathematical reasoning, generalization, and justification [15], [16], [17]. Francisco similarly highlights deliberate teacher support in collective mathematical argumentation [18].

However, the quality of intervention is crucial. Bredow and Knipping show that teacher actions can shape mathematical argumentation in complex ways. When teachers provide a mathematical warrant immediately after students hesitate, intellectual responsibility can shift away from the learner [19].

Verification therefore requires a balance between support and challenge. The teacher needs to scaffold reasoning while allowing students to retain responsibility for establishing the validity of a mathematical claim. Peer critique adds a social dimension. When students present mathematical solutions publicly, their reasoning becomes available for examination. Peer questions may require clarification, justification of a chosen procedure, reconsideration of assumptions, or revision of unsupported conclusions. This aligns with research showing that collaborative interaction can support the construction and refinement of mathematical arguments [26], [14], [14].

4.5 Differentiated Support for Mathematical Justification

Students do not enter verification with identical readiness. Some students may communicate reasoning verbally but struggle to represent it symbolically. Others may require guiding questions, visual representations, or partially structured arguments before identifying appropriate mathematical evidence.

ACTIVER addresses this heterogeneity through differentiated access, representations, scaffolding, and response formats while preserving common reasoning objectives.

Students requiring more support may receive guiding questions, diagrams, structured claim–evidence prompts, or opportunities to rehearse reasoning before presentation. Students demonstrating greater readiness may be challenged to develop counterexamples, compare alternative justifications, analyse invalid arguments, or evaluate the strength of competing mathematical evidence.

The central principle is that differentiation changes the pathway and level of support, not the intellectual standard for mathematical justification. This interpretation is consistent

 https://doi.org/10.58421/gehu.v5i4.1963

with differentiated-instruction research emphasizing responsiveness to learner readiness while maintaining substantive learning goals [20], [21].

The high mastery proportions observed in the ACTIVER groups are compatible with this interpretation. Nevertheless, the study did not experimentally isolate differentiated support, so the results cannot establish its independent contribution.

4.6 Retrieval as a Resource for Verification

The position of Internalize Memory immediately before Verify Understanding provides another theoretical explanation for the ACTIVER sequence.

Students cannot effectively evaluate a mathematical claim when the relevant definitions, properties, or relationships are inaccessible. Retrieval activities reactivate previously constructed mathematical knowledge so that it can function as a resource during verification.

Roediger and Karpicke established the broader learning benefits of retrieval practice [22], while Anna and Csaba reported advantages of retrieval practice in university mathematics, including complex mathematical problem solving [23].

In ACTIVER, retrieval is therefore not intended merely as formula repetition. Its pedagogical function is to make conceptual knowledge available when students need to evaluate a strategy, generalize, or draw a mathematical conclusion.

Because no delayed posttest was administered, however, the present findings should not be interpreted as evidence that ACTIVER independently improved long-term memory.

4.7 Interpreting the Effectiveness Evidence

The effectiveness findings were consistent across three comparative settings.

In the limited trial, ACTIVER was associated with higher posttest achievement, higher mastery attainment, and stronger N-Gain than the comparison instruction. The pattern was reproduced in both broader trials.

Replicating the same direction across different classes and comparison approaches strengthens the product-evaluation evidence because the finding was not confined to a single classroom.

Mastery proportions provide an interpretation that mean scores alone cannot offer. A higher mean can potentially be influenced by a smaller number of high-performing students, whereas the high proportions reaching mastery in the ACTIVER groups indicate that successful learning outcomes were distributed across a substantial part of each class.

N-Gain provides an additional perspective because it captures measured improvement rather than final achievement alone.

Nevertheless, the exceptionally large standardized differences require caution. The study used intact classes rather than individual random assignment. Relatively low within-class variability, class composition, teacher characteristics, implementation fidelity, comparison-treatment differences, novelty, or differences in instructional attention may have contributed to the magnitude of the effects.

Consequently, the results support ACTIVER's effectiveness as an integrated instructional model in the settings investigated, but they do not establish that Verify Understanding independently caused the observed differences.

4.8 Theoretical Contribution

The study's principal theoretical contribution is to position verification as a bridge among conceptual Understanding, generalization, retrieval, and mathematical justification.

Verification is not conceptualized merely as a final step for checking numerical answers. It is the point at which students are asked to use mathematical knowledge as evidence for determining whether a claim should be accepted.

This positioning also connects private mathematical reasoning with public argumentation. Reasoning that appears convincing to an individual learner needs to be communicated in a form that teachers and peers can question and evaluate.

Thus, Verify Understanding creates a designated instructional space in which mathematical authority is intended to rest increasingly on evidence and relationships rather than exclusively on teacher confirmation, consistent with sociomathematical perspectives on acceptable mathematical explanation [5].

4.9 Practical Implications

The findings suggest several implications for mathematics instruction.

Verification should be allocated explicit instructional time rather than reduced to checking final answers. Teachers can structure verification around recurring questions: What is the mathematical claim? What evidence supports it? Why does the evidence support the conclusion? Does the solution satisfy all original conditions? Is there an alternative strategy or counterexample? What part of the argument needs revision?

Assessment should also distinguish answer correctness from quality of reasoning. A correct numerical answer produced through unsupported reasoning should not necessarily be treated as equivalent to a correct answer supported by systematic mathematical justification.

Differentiated scaffolding can help students with varying readiness engage in these demands. The form of support may differ, but the reasoning objective remains common.

For vocational mathematics, this orientation is particularly relevant because occupational and applied contexts often require learners to explain why a quantitative decision is reasonable rather than merely calculate a value.

4.10 Implementation Constraints and Alternative Explanations

The implementation of explicit verification also has practical costs.

Verification requires classroom time for explanation, questioning, critique, and revision. Teachers need sufficient mathematical and pedagogical knowledge to recognize incomplete warrants and formulate productive questions without replacing students' reasoning.

Differentiated scaffolding increases planning demands, especially in large classes, and evaluation of written or oral justification may require more assessment time than checking numerical answers.

Not every student can necessarily defend a solution publicly during a single lesson. Teachers may therefore need combinations of group presentations, peer review, short written justifications, and targeted questioning.

The parent study did not quantify teacher workload or the amount of instructional time devoted specifically to verification. Therefore, although successful implementation across eight meetings indicates feasibility in the participating setting, no claim is made that the approach carries negligible implementation costs.

Alternative explanations must also be acknowledged. Because intact classes were used, pre-existing class characteristics may have contributed to the differences. Teacher enthusiasm, novelty effects, instructional fidelity, differences among comparison models, and repeated exposure to related mathematical tasks cannot be fully excluded.

4.11 Limitations and Directions for Future Research

Several limitations frame the conclusions of the study.

First, Verify Understanding was not experimentally isolated. The quantitative effectiveness findings represent the complete ACTIVER sequence.

Second, the principal quantitative outcome was overall mathematical reasoning rather than justification alone. Although J1 and J2 were explicitly defined, separate experimental estimates of improvement in these two justification indicators were not available.

Third, intact classes were used instead of individual random assignment.

Fourth, comparison instruction differed across trials. This increases contextual breadth but complicates direct comparisons of treatment-effect magnitude.

Fifth, the original development study was not designed as a detailed classroom-discourse investigation. The implementation evidence identifies opportunities for verification but cannot quantify each student's progression from procedural or empirical verification to more sophisticated mathematical justification.

Sixth, no delayed posttest was administered.

Seventh, the study did not formally measure teacher workload, instructional-time costs, or assessment burden associated specifically with verification and differentiation.

Future research should therefore employ phase-comparison or phase-ablation designs, analyse J1 and J2 separately, conduct detailed coding of written and oral arguments, use classroom discourse analysis, administer delayed posttests, replicate the intervention across additional mathematical domains and school contexts, and compare ACTIVER directly with other reasoning-oriented instructional models

5. CONCLUSION

This study examined the pedagogical role of Verify Understanding within the integrated ACTIVER Learning Model.

In response to the first research question, mathematical justification was operationalized through activities requiring students to check solutions against mathematical conditions, explain transformations, examine generalizations, present evidence, and establish why mathematical conclusions should be accepted.

In response to the second research question, verification provided instructional opportunities through teacher questioning, peer critique, presenting mathematical evidence, defending claims, and revising incomplete reasoning. These activities positioned justification as an explicit mathematical practice rather than an optional explanation after obtaining an answer.

In response to the third research question, the classroom trials showed that ACTIVER, as an integrated model, was consistently associated with higher mathematical reasoning performance, mastery attainment, and learning gains than the respective comparison instruction. The same overall pattern occurred in the limited trial and both broader trials.

The study’s theoretical contribution lies in positioning verification as a bridge connecting conceptual Understanding, generalization, retrieval of relevant mathematical knowledge, and justification. Its practical contribution lies in providing teachers with a structured pathway for requiring students not only to solve mathematical problems but also to explain why their solutions deserve mathematical acceptance.

These conclusions should nevertheless remain within the boundaries of the research design. The study does not establish that Verify Understanding independently caused the observed improvements, and overall reasoning results cannot be treated as a separate statistical estimate of improvement specifically in justification.

Future research should isolate the contribution of verification more directly, examine students’ written and oral justification longitudinally, analyse J1 and J2 separately, investigate the role of teacher questioning and differentiated scaffolding, and determine whether reasoning gains persist across mathematical domains and over longer periods

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to the promoters, six expert validators, mathematics teachers, classroom observers, participating students, the Vocational High School Mathematics Teachers’ Association (MGMP Matematika SMK) of Demak Regency, and the vocational schools that supported the development, validation, implementation, and dissemination of the ACTIVER Learning Model.

REFERENCES

[1] M. Niss and T. Højgaard, “Mathematical competencies revisited,” *Educ. Stud. Math.*, 2019, doi: https://doi.org/10.1007/s10649-019-09903-9 Mathematical.

[2] OECD, *PISA 2022 Assessment and Analytical Framework*. 2023.

[3] A. J. Stylianides, “Proof and proving in school mathematics,” *J. Res. Math. Educ.*, vol. 38, no. 3, pp. 289–321, 2007.

[4] M. E. Staples, J. Bartlo, and E. Thanheiser, “The Journal of Mathematical Behavior Justification as a teaching and learning practice : Its (potential) multifacted role in middle grades mathematics classrooms,” *J. Math. Behav.*, vol. 31, no. 4, pp. 447–462, 2012, doi: 10.1016/j.jmathb.2012.07.001.

[5] E. Yackel and P. Cobb, “Sociomathematical Norms, Argumentation, And Autonomy In Mathematics,” *J. Res. Math. Educ.*, vol. 27, no. 4, pp. 458–477, 2011.

[6] Y. A. Sarumaha and I. Rizkianto, “PROMOTING MATHEMATICAL JUSTIFICATION THROUGH,” *J. Pendidik. Mat.*, vol. 5, no. 2, pp. 83–94, 2022.

[7] A. T. Fatimah, P. Studi, P. Matematika, and U. G. Ciamis, “Mathematical Reasoning of Vocational High School Students on Mathematical Tasks in the Law of Demand Context,” *J. Math. Educ.*, vol. 7, no. November, pp. 101–113, 2021.

[8] D. Septi, N. Afifah, M. Syauqy, and M. I. Nafián, “Adaptive Reasoning Characteristics of Vocational School Students in Solving Mathematical Problems,” *Al-Ishlah J. Pendidik.*, vol. 15, pp. 3028–3039,

2023, doi: 10.35445/alishlah.v15i3.3748.

[9] R. T. Ronny, Harry D. Putra, and H. Hendriana, "Analysis Of Vocational School Student ' S Mathematical Reasoning In Quadratic," *J. Innov. Math. Learn.*, vol. 7, no. 2, pp. 145–152, 2024.

[10] W. Widjaja, C. Vale, S. Herbert, E. Y. Loong, and L. A. Bragg, "Linking comparing and contrasting, generalising and justifying : a case study of primary students ' levels of justifying," *Math. Educ. Res. J.*, 2020.

[11] T. Öz and Z. ÇİftcȈ, "Mathematical Reasoning Activity : Compare, Generalize and Justify," *Necatibey Fac. Educ. Electron. J. Sci. Math. Educ.*, vol. 18, no. 2, pp. 291–323, 2024, doi: 10.17522/balikesirnef.1506921.

[12] M. F. Mueller, "The co-construction of arguments by middle-school students," *J. Math. Behav.*, vol. 28, pp. 138–149, 2009, doi: 10.1016/j.jmathb.2009.06.003.

[13] J. Hiltunen, K. Jokiranta, S. Lehesvuori, and P. Nieminen, "Learning, Culture and Social Interaction Students ' dialogic and justifying moves during dialogic argumentation in mathematics and physics ¨ hki o," *Learn. Cult. Soc. Interact. J.*, vol. 33, no. August 2020, 2022, doi: 10.1016/j.lcsi.2022.100608.

[14] A. G. Uzun, "The development of mathematical argumentation : A case study on two mathematics classrooms," *Int. Electron. J. Math. Educ.*, vol. 19, no. 2, 2024.

[15] J. Mata-pereira and J. Ponte, "Enhancing students ' mathematical reasoning in the classroom : teacher actions facilitating generalization and justification," *Educ Stud Math*, 2017, doi: 10.1007/s10649-017-9773-4.

[16] J. Mata-Pereira and J. P. Da Ponte, "Promoting Students' Mathematical Reasoning: A design-based research," *Bolema - Math. Educ. Bull.*, vol. 32, no. 62, pp. 781–801, 2018, doi: 10.1590/1980-4415v32n62a02.

[17] A. Mukuka, S. Balimuttajjo, and V. Mutarutinya, "Teacher efforts towards the development of students' mathematical reasoning skills," *Heliyon*, vol. 9, no. 4, p. e14789, 2023, doi: 10.1016/j.heliyon.2023.e14789.

[18] J. M. Francisco, "Teacher support of collective argumentation in mathematics classrooms : insights from an afterschool program," *J. Math. Teach. Educ.*, vol. 28, no. 1, pp. 211–237, 2025, doi: 10.1007/s10857-024-09651-5.

[19] F. Bredow and C. Knipping, "Teacher actions that shape mathematical argumentation in the classroom," *J. Math. Teach. Educ.*, no. 0123456789, 2025, doi: 10.1007/s10857-025-09728-9.

[20] A. E. Smale-Jacobse, A. Meijer, M. Helms-Lorenz, and R. Maulana, "Differentiated Instruction in Secondary Education: A Systematic Review of Research Evidence," *Front. Psychol.*, vol. 10, no. November, 2019, doi: 10.3389/fpsyg.2019.02366.

[21] M. Afifi, B. Setambah, M. Ikhwan, and M. Saad, "A comprehensive synthesis of differentiated instruction approaches in mathematics," *Int. J. Math Sci. Technol. Educ.*, vol. 12, no. 4, 2025.

[22] H. L. Roediger and J. D. Karpicke, "Test-Enhanced Learning Taking Memory Tests Improves Long-Term Retention," *Psychol. Sci.*, vol. 17, no. 3, pp. 249–255, 2006.

[23] L. Anna and T. S. Csaba, "Investigating the efficacy of retrieval practice in university mathematics," 2024.

[24] R. M. Branch, *Instructional Design: The ADDIE Approach*. Springer, New York, Dordrecht, Heidelberg, London, 2009. doi: 10.1007/978-3-319-19650-3_2438.

[25] M. Marlina, R. Rohati, A. Kumalasari, and S. Winarni, "Levels of Mathematical Argumentation : An Analysis of Junior High School Students ' Reasoning in Number Pattern Tasks," *J. Gen. Educ. Humanit.*, vol. 5, no. 2, pp. 2457–2468, 2026.

[26] M. F. Mueller, "The Journal of Mathematical Behavior The co-construction of arguments by middle-school students," *J. Math. Behav.*, vol. 28, pp. 138–149, 2009, doi: 10.1016/j.jmathb.2009.06.003.