

Improvement in Junior High School Students' Mathematical Critical Thinking Ability through the Teams Games Tournament (TGT) Learning Model Assisted by Wordwall on the Topic of Two-Variable Linear Equation Systems (SPLDV)

Asrah Nabila Purba¹, Yahfizham²

^{1,2}Universitas Islam Negeri Sumatera Utara Medan, Indonesia

Article Info

Article history:

Received 2026-09-14

Revised 2026-10-10

Accepted 2026-10-16

Keywords:

Mathematical Critical Thinking Ability

Teams Games Tournament

Two-Variable Linear Equation-System

Wordwall

ABSTRACT

This research was motivated by junior high school students' low mathematical critical thinking ability in Two-Variable Linear Equation Systems (SPLDV), shown by their difficulties interpreting word problems, constructing mathematical models, and drawing logical conclusions. This study aims to describe the change in students' mathematical critical thinking ability on SPLDV after learning with the Teams Games Tournament (TGT) model assisted by Wordwall. Using a quantitative pre-experimental approach with a One Group Pretest-Posttest Design, the sample comprised 28 eighth-grade students (class VIII-B) of SMP Negeri 1 Sipispis, selected through purposive sampling. The instrument consisted of five essay items, one for each of five critical thinking indicators, scored with a 0–4 rubric. A pilot test with 25 students showed item-total correlations of 0.59–0.85 ($r_{table} = 0.396$) and a Cronbach's Alpha of 0.76. Data were analyzed using normality tests on the difference scores, a paired-samples t-test ($\alpha = 0.05$), effect size, and normalized gain. The mean score increased from 36.61 (pre-test) to 75.71 (post-test); the mean difference of 39.11 points was statistically significant (95% CI [33.33, 44.89]; $t(27) = 13.88$, $p < 0.001$; Cohen's $d_z = 2.62$), with a normalized gain of 0.62 (moderate), and all 28 students improved. Scores rose on all five indicators. Because the design had no control group, the improvement cannot be attributed solely to the treatment and should be confirmed with a quasi-experimental design.

This is an open-access article under the [CC BY-SA](#) license.



Corresponding Author:

Yahfizham

Faculty of Tarbiyah and Teacher Training, Universitas Islam Negeri Sumatera Utara Medan, Indonesia

Email: yahfizham@uinsu.ac.id

1. INTRODUCTION

Twenty-first-century education requires learners to master the 4C competencies, namely critical thinking and problem solving, creativity, communication, and collaboration [1], [2], [3], [4], [5]. This requirement is consistent with Regulation of the Minister of

Education and Culture of the Republic of Indonesia Number 20 of 2016 concerning Graduate Competency Standards for Basic and Secondary Education [6], which affirms that every graduate of a basic or secondary education unit must possess creative, productive, critical, independent, collaborative, and communicative competencies. In light of this regulation, mathematical critical thinking ability becomes an aspect that warrants serious attention in mathematics instruction at school [7], [8], [9], [10], [11], [12].

Critical thinking has been described as reasonable, reflective thinking focused on deciding what to believe or do [13]. Within the context of mathematics learning, it plays a central role because mathematical content is organized through interrelated systems, concepts, principles, and relationships arranged in a structured manner. Critical thinking can also be understood as an individual's ability to obtain accurate information, examine and process that information, and subsequently use it as a basis for decision-making [14].

A range of research findings indicate that Indonesian students' mathematical critical thinking ability remains relatively low, particularly on the topic of Two-Variable Linear Equation Systems (SPLDV). Agustin and Effendi [15] state that in a small qualitative study, it was found that the mathematical critical thinking ability of junior high school students in Purwakarta Regency on the SPLDV topic fell into the low category. Even at a higher educational level, Benyamin et al. [16] reported that the critical thinking ability of tenth-grade senior high school students in solving three-variable linear equation system (SPLTV) problems reached only 43.01%, with the explanation aspect being the weakest.

The recurring pattern of low mathematical critical thinking ability on the SPLDV topic observed across various regions strengthens the assumption that a similar problem also occurs at SMP Negeri 1 Sipispis. Interviews with the eighth-grade mathematics teacher at SMP Negeri 1 Sipispis revealed that most students experienced difficulty when confronted with SPLDV problems that could not be solved directly using previously provided formulas or examples. An analysis of class VIII-B students' work further showed that students made errors in identifying the information given in the problem, failed to equalize variable coefficients during the elimination process, and misinterpreted the elimination results, leading to incorrect conclusions. These errors indicate that students continue to struggle with interpreting problems, devising appropriate solution strategies, and drawing logical conclusions, the three principal indicators of mathematical critical thinking ability.

Beyond the cognitive dimension, students' interest in learning mathematics at this school also remains relatively low, as reflected in their limited attention and enthusiasm during lessons, where students tend to be passive, rarely ask questions, and show little motivation in completing tasks. This low learning interest is further compounded by a lack of variety in teaching methods and media, and point to the role of interactive learning media in raising students' interest and motivation, such that less innovative learning has the potential to diminish student motivation [17]. This condition was likewise observed at SMP Negeri 1 Sipispis, where the use of technology-based media in mathematics learning remains very limited.

The Teams Games Tournament (TGT) cooperative learning model offers one alternative for creating enjoyable learning while cultivating collaboration among students in mastering the material. In this model, students are not only required to master the subject

matter individually but are also expected to work together in groups to earn team points through a series of games and academic tournaments. The games and tournament stages of TGT require students to solve problems individually in front of their group and defend their answers under competitive conditions, thereby training them to analyze problems and articulate the reasoning behind their solution strategies. Wordwall, meanwhile, is a game-based learning application offering features such as quizzes, matching activities, and other interactive games that can be freely accessed through a browser. The interactive quiz feature in Wordwall provides immediate feedback (correct/incorrect scores) as soon as students respond, prompting them to re-evaluate their solution strategy whenever an answer proves incorrect. It is this combination of team-based competition in TGT and instant feedback in Wordwall that is expected to gradually cultivate students' mathematical critical thinking indicators, particularly on the SPLDV topic, which requires constructing mathematical models from word problems.

A number of prior studies bear directly on the use of TGT and Wordwall for critical thinking. Ali et al. [17] used a one-group pretest-posttest design and reported a significant difference (paired t-test, Sig. < 0.05) in fourth-grade students' critical thinking in mathematics after learning with TGT assisted by Wordwall. Ainurrohman et al. [18] likewise reported a significant paired-sample result for the use of Wordwall on critical thinking in elementary mathematics. Rohman [19] compares TGT assisted by Wordwall with Problem-Based Learning among eighth-grade students at MTs Negeri 3 Jenepono in a quasi-experimental design, and reported a significant difference in mathematical critical thinking between the two classes, with a normalized gain in the "fairly effective" category for the TGT–Wordwall class. The combination of TGT and Wordwall for mathematical critical thinking is therefore not in itself new. The present study adds to this literature in three limited ways: it focuses on SPLDV in a public junior high school in North Sumatra; it maps five critical thinking indicators [13] to one essay item each at cognitive levels C3–C5 and reports indicator-level results; and it reports the instrument, effect size, and the inferential limits of a one-group design transparently. It does not claim stronger causal evidence than the controlled study of Rohman [19].

Based on the foregoing discussion, this study aims to describe the improvement in junior high school students' mathematical critical thinking ability on the topic of Two-Variable Linear Equation Systems (SPLDV) from before to after learning with the Teams Games Tournament (TGT) model assisted by Wordwall, and to assess the statistical significance and magnitude of that improvement. The findings of this study are expected to contribute theoretically to the development of mathematics education research, particularly regarding the application of game-based cooperative learning models, as well as practically by offering teachers an innovative and enjoyable alternative strategy for improving students' mathematical critical thinking ability.

2. METHOD

This study employed a quantitative approach using a pre-experimental research design [20]. Specifically, a One Group Pretest-Posttest Design was used, in which a single group of subjects was administered an initial test (pre-test), subsequently given treatment

consisting of learning through the Teams Games Tournament (TGT) model assisted by Wordwall, and finally administered a final test (post-test). This design was intended to determine the change in students' mathematical critical thinking ability before and after treatment, denoted by the notation $O1 \times O2$, where $O1$ represents the pre-test, X represents the implementation of the TGT model assisted by Wordwall, and $O2$ represents the post-test. Because this design has no comparison group, it can only show the change between pre-test and post-test and cannot establish cause and effect [21].

The study was conducted at SMP Negeri 1 Sipispis, Serdang Bedagai Regency, North Sumatra, during the odd semester of the 2026/2027 academic year. The research population comprised all students of SMP Negeri 1 Sipispis in the 2026/2027 academic year, consisting of nine classes with a total of 238 students. The research sample was determined using purposive sampling, a technique in which sample selection is based on specific considerations or criteria relevant to the research objective [20]. Class VIII-B, consisting of 28 students, was selected as the sample based on the following considerations: (1) interview data and daily grade records from the mathematics teacher indicated that this class exhibited relatively low mathematical critical thinking ability on the SPLDV topic; (2) class VIII-B had not previously received instruction using a cooperative model supported by interactive media, allowing the effect of the treatment to be observed more clearly; and (3) the school granted permission to conduct the research in this class without disrupting the academic schedule of other classes.

The research was carried out over four sessions: one pre-test session, two learning sessions employing the TGT model assisted by Wordwall, and one post-test session. During the treatment sessions, students were divided into heterogeneous groups to discuss material, complete Student Worksheets (LKPD), participate in educational games via Wordwall during the games stage, and compete in an inter-group academic tournament during the tournament stage, in accordance with the TGT syntax comprising material presentation, group study, games, tournament, and group recognition. The instructional materials in the form of a teaching module and Student Worksheets were developed based on the TGT syntax assisted by Wordwall on the SPLDV topic and were validated by a mathematics lecturer and a mathematics teacher prior to implementation.

The research instrument consisted of five essay items assessing mathematical critical thinking ability, administered as a pre-test and a post-test with parallel items (the same indicators and cognitive levels but different story contexts) so that pre- and post-treatment results could be compared. Item development was grounded in five indicators of mathematical critical thinking ability [13], with one item per indicator: item 1, providing elementary clarification (C3); item 2, building basic support (C3); item 3, drawing inferences (C4); item 4, providing advanced clarification (C5); and item 5, employing strategies and tactics (C5). Each item was scored with a holistic 0–4 rubric (4 = interprets the problem, builds the model, solves the system correctly, interprets the result correctly, and checks the answer logically and completely; 3 = correct model and solution but incomplete interpretation or evaluation of the result; 2 = model constructed but errors in the solution process lead to an inaccurate result; 1 = shows little understanding but cannot yet build an appropriate model; 0 = no answer or irrelevant answer). The maximum total score is 20,

converted to a 0–100 scale as $(\text{total score} \div 20) \times 100$. Content validation used a ten-aspect form rated on a 1–5 scale (alignment of the items with the indicators and with the blueprint, SPLDV content, cognitive level, clarity of wording, limits of the question and expected answer, contextual form, clarity of instructions, language, and time allocation). The lecturer validator (Universitas Islam Negeri Sumatera Utara) rated all aspects 3 or 4 (mean 3.60), and the teacher validator (a mathematics teacher at a private junior high school) likewise rated all aspects 3 or 4 (mean 3.40); both concluded that the instrument could be used without revision. All answers were scored by the first author using this rubric; no second rater was involved. The instrument was then piloted on 25 students. Validity was tested with the Pearson product-moment correlation (item-total correlation); all five items were valid, as the calculated r -values (0.59 to 0.85) exceeded the critical r -value of 0.396 ($n = 25$, $\alpha = 0.05$). The corrected item-total correlations (with the item removed from the total) ranged from 0.39 to 0.74, with item 5 at the threshold (0.392). Reliability, estimated using Cronbach's Alpha, was 0.76, which falls in the high category (0.60–0.79) of the criteria used and is acceptable for a five-item essay test, although below 0.80. Item difficulty ranged from difficult (item 4) to easy (item 1). Only the pre-test item set was piloted; the equivalence of the post-test set rests on the matching of indicators and cognitive levels and was not tested empirically.

Data analysis comprised descriptive statistics, assumption testing, hypothesis testing, effect size, and normalized gain. Normality was examined using the Lilliefors test at a significance level of $\alpha = 0.05$ (data are considered normally distributed when $L_{\text{count}} < L_{\text{table}}$, with L_{table} approximated by $0.886/\sqrt{n}$) and the Shapiro-Wilk test, applied to the pre-test, post-test, and difference scores (post-test – pre-test), since the paired t -test assumes normally distributed differences. Hypothesis testing used a two-tailed paired sample t -test at $\alpha = 0.05$, reported with a 95% confidence interval; H_0 is rejected when the significance value (Sig.) < 0.05 or $t_{\text{count}} > t_{\text{table}}$. Effect size was reported as Cohen's d_z (mean difference divided by the standard deviation of the differences; 0.2 = small, 0.5 = medium, 0.8 = large; [22], and normalized gain was computed following [23] as $(\text{post-test} - \text{pre-test}) \div (100 - \text{pre-test})$, categorized as low (< 0.3), moderate (0.3–0.7), or high (> 0.7). All data processing was carried out using Microsoft Excel 2024 for Mac.

3. RESULTS AND DISCUSSION

3.1. Description of Pre-Test and Post-Test Results

Research data were obtained by administering a mathematical critical thinking ability test to 28 students of class VIII-B at SMP Negeri 1 Sipispis, both before (pre-test) and after (post-test) the treatment consisting of learning through the Teams Games Tournament (TGT) model assisted by Wordwall on the topic of Two-Variable Linear Equation Systems (SPLDV). A summary of the descriptive statistics for both sets of data is presented in Table 1.

As shown in Table 1, the pre-test results of students' mathematical critical thinking ability prior to treatment indicate a mean score of 36.61, with a minimum score of 15 and a maximum score of 55. A variance of 77.88 and a standard deviation of 8.82 reveal considerable variation in students' critical thinking ability before treatment. These results

indicate that students' mathematical critical thinking ability on the SPLDV topic was relatively low and in need of improvement, consistent with the initial problem underlying this study; no student reached a score of 75 or above.

Table 1. Descriptive Statistics of Pre-Test and Post-Test Results

Descriptive Statistic	Pre-Test	Post-Test
Mean	36.61	75.71
Variance	77.88	106.88
Standard Deviation	8.82	10.34
Minimum Score	15	55
Maximum Score	55	95
Number of Students (N)	28	28

Source: Pre-test and post-test scores of class VIII-B, SMP Negeri 1 Sipispis (per-student raw scores converted to a 0–100 scale).

Following the treatment involving learning through the TGT model assisted by Wordwall, the post-test results show an increase in the mean score to 75.71, with a minimum score of 55 and a maximum score of 95; 19 of the 28 students reached a score of 75 or above. The increase in variance to 106.88 and standard deviation to 10.34 likewise indicates variation in students' achievement after treatment; nevertheless, the overall post-test results demonstrate an improvement in students' mathematical critical thinking ability compared to the pre-test results, and no student's score decreased (the smallest increase was 10 points and the largest 65 points). These pre-test and post-test data were subsequently used in the assumption testing and hypothesis testing to examine the change in scores.

Normality Test

A normality test was conducted as a prerequisite prior to hypothesis testing using parametric statistics. The tests were performed using the Lilliefors test (L_0) at a significance level of $\alpha = 0.05$, applying the criterion that data are declared normally distributed when $L_{count} < L_{table}$, and the Shapiro-Wilk test. Both were applied to the pre-test scores, the post-test scores, and the difference scores (post-test – pre-test), the latter being the distribution assumed by the paired t-test. The results are presented in Table 2.

Table 2. Results of the Normality Tests (Lilliefors and Shapiro-Wilk)

Data	Lcount	Ltable	Shapiro-Wilk p
Pre-test	0.150	0.167	0.387
Post-test	0.173	0.167	0.120
Difference (post – pre)	0.119	0.167	0.375

Source: Lilliefors and Shapiro-Wilk tests at a significance level of $\alpha = 0.05$ (L_{count} is the larger of the two one-sided differences between the empirical and normal distribution functions; L_{table} approximated by $0.886/\sqrt{n}$ for $n = 28$).

As shown in Table 2, the difference scores satisfy the normality assumption on both tests ($L_{\text{count}} = 0.119 < L_{\text{table}} \approx 0.167$; Shapiro-Wilk $p = 0.375$), as do the pre-test scores. For the post-test scores, L_{count} (0.173) slightly exceeds L_{table} , whereas the Shapiro-Wilk test does not reject normality ($p = 0.120$); this is plausible because the scores are discrete with many tied values. Since the paired t-test requires normally distributed differences, hypothesis testing proceeded using the paired sample t-test, and the Wilcoxon signed-rank test was added as a sensitivity check.

Hypothesis Testing (Paired Sample T-Test)

Hypothesis testing in this study was conducted using a paired sample t-test with the assistance of Microsoft Excel 2024 for Mac at a significance level of $\alpha = 0.05$, reported with a 95% confidence interval, effect size, and normalized gain. This test was employed because the study involved only a single group of subjects who completed both the pre-test and post-test. The results of the hypothesis test are presented in Table 3.

Table 3. Results of the Paired Sample T-Test

Description	Post-Test	Pre-Test
Mean	75.71428571	36.60714286
Variance	106.8783069	77.87698413
Observations (N)	28	28
Mean difference (post – pre)	39.10714286	
Hypothesized Mean Difference	0	
df	27	
t Stat (t count)	13.87856109	
t Critical two-tail (t table)	2.051830516	
P (Sig.) two-tailed	8.31E-14 ($p < 0.001$)	
SD of the differences	14.91	
Standard error of the difference	2.82	
95% CI of the mean difference	[33.33, 44.89]	
Cohen's dz (95% CI)	2.62 [1.84, 3.40]	
Normalized gain (N-gain, from the means)	0.62 (moderate)	
Wilcoxon signed-rank test	W = 0, $p < 0.001$	

Source: Results of the paired sample t-test using Microsoft Excel 2024 for Mac on student-matched pre-test and post-test scores; SD of the differences, standard error, confidence interval, dz, N-gain, and Wilcoxon test computed from the same per-student data.

As shown in Table 3, the mean difference between post-test and pre-test was 39.11 points (SD of the differences = 14.91; SE = 2.82), with a 95% confidence interval of [33.33, 44.89]. The t-count value was 13.88 with 27 degrees of freedom (df), while the t-table value at a significance level of 0.05 was 2.052. Since the t-count exceeds the t-table value and the significance value is below 0.001, H_0 is rejected, and H_1 is accepted. The Wilcoxon signed-rank test led to the same conclusion. The effect size was Cohen's $dz = 2.62$ (95% CI

approximately [1.84, 3.40]), a very large effect by the conventional benchmarks, although the wide interval reflects the small sample. The mean normalized gain was 0.62, in the moderate category; at the individual level, 11 students fell in the high category, 16 in the moderate category, and 1 in the low category. It can therefore be concluded that there is a significant difference between students' mathematical critical thinking scores before and after learning with the Teams Games Tournament (TGT) model assisted by Wordwall, indicating an improvement in students' mathematical critical thinking ability on the SPLDV topic. However, the paired t-test and the effect size describe only the magnitude of the pretest-to-posttest change; they do not show that the change was caused solely by the treatment.

Table 4. Mean Item Score per Critical Thinking Indicator (maximum score = 4)

Indicator (item)	Pre-test mean	Post-test mean	N-gain
Elementary clarification (1, C3)	1.46	2.86	0.55
Basic support (2, C3)	1.57	3.18	0.66
Inference (3, C4)	1.54	3.07	0.62
Advanced clarification (4, C5)	1.29	2.93	0.61
Strategies and tactics (5, C5)	1.46	3.11	0.65

Source: Item scores of class VIII-B, SMP Negeri 1 Sipispis ($n = 28$). $N\text{-gain} = (\text{post-test mean} - \text{pre-test mean}) \div (4 - \text{pre-test mean})$.

As shown in Table 4, mean item scores increased on all five indicators, from 1.29–1.57 at pre-test to 2.86–3.18 at post-test, with moderate normalized gains of 0.55–0.66. The largest gains in raw score were on advanced clarification and strategies and tactics (1.64 points each) and the smallest on elementary clarification (1.39 points). The change was least uniform on inference, where 20 students improved, 4 stayed the same, and 4 declined. Because each indicator was measured by a single item and the pre-test and post-test items differ in context, these per-indicator results are descriptive and should not be taken as evidence that any particular indicator was affected more than the others.

3.2. Discussion

The findings of this study show an improvement in students' mathematical critical thinking ability on the SPLDV topic after learning with the Teams Games Tournament (TGT) model assisted by Wordwall, as evidenced by the increase in the mean score from 36.61 in the pre-test to 75.71 in the post-test and by the paired sample t-test, which rejected H_0 at a significance level of 0.05. These findings are consistent with the expectation that a team-competition-based cooperative learning model combined with an interactive, instant-feedback game-based medium can support the development of students' mathematical critical thinking ability. However, because this study used a one-group design without a

control group, the improvement cannot be attributed solely to the TGT model assisted by Wordwall.

This improvement may be explained, tentatively and at a theoretical level, by the characteristics of the TGT model itself, which is designed to motivate students to support one another and actively master the learning material in order to earn points for their group[24]. This condition may encourage students to understand the material in depth before participating in the academic tournament, thereby possibly training them to interpret problems, analyze issues, draw conclusions, and evaluate answers logically, activities that correspond to indicators of mathematical critical thinking ability. Meanwhile, the use of Wordwall media may reinforce this process by creating an interactive, engaging, and non-monotonous learning atmosphere through quizzes, games, and visualizations capable of enhancing students' interest and engagement, and by giving immediate feedback that allows students to correct errors such as failing to equalize coefficients during elimination, which was found in the preliminary study. These explanations were not tested in this study; the per-indicator results (Table 4) show gains on all five indicators but cannot show which component of the treatment produced them.

These findings are consistent in direction with earlier reports.[17] , using a similar one-group design, and [18], using Wordwall alone, both reported significant pretest-to-posttest differences in critical thinking in elementary mathematics, and [19] reported a significant difference in mathematical critical thinking in favor of a TGT–Wordwall class over a Problem-Based Learning class among eighth-grade students. The size of the gain is also of a comparable order: the experimental class in [19] fell in the "fairly effective" normalized-gain category, whereas the mean normalized gain here was 0.62, in the moderate category by Hake's criteria, although the two classification schemes and designs are not directly comparable. Unlike [19], this study had no comparison class, so it cannot show whether the gain exceeds what an alternative approach would produce; and unlike the elementary-school studies, it was conducted in junior high school on a single topic, SPLDV. Taken together, these results are compatible with the view that combining the TGT cooperative learning model with the interactive Wordwall medium may support students' mathematical critical thinking ability on the topic of Two-Variable Linear Equation Systems (SPLDV), but stronger evidence requires a control-group design.

Several limitations should be noted for consideration in future research. First, this study employed a one-group pretest-posttest design without a control group, so the observed improvement may also reflect other events at school during the study (history), students' natural development (maturation), and familiarity with similar tests (testing). Although the pre-test and post-test used parallel items, their equivalence in difficulty was not tested (instrumentation). Second, the class was selected because its initial ability was relatively low, so part of the gain may reflect regression toward the mean, and students new to game-based learning may have been more enthusiastic because of novelty. Third, some students felt confused at the beginning of the learning process, as they were accustomed to teacher-centered instruction, and limited instructional time within a single session affected the optimal execution of the games and tournament stages and made it difficult to monitor less active students within their groups. Fourth, the study involved a single class ($n = 28$) on a

single topic (SPLDV), so the results cannot necessarily be generalized to other topics or classes. Fifth, each indicator was measured by a single item, so the per-indicator results are descriptive and cannot support conclusions about specific indicators. Sixth, all answers were scored by a single rater (the first author) without a second rater, so scoring consistency was not checked, and scorer bias cannot be excluded. Further research employing a control-group design (for example, TGT without Wordwall or conventional instruction), broader material coverage, and larger samples is therefore needed.

4. CONCLUSION

Among the 28 students of class VIII-B at SMP Negeri 1 Sipispis, mathematical critical thinking scores on the SPLDV topic increased from a mean of 36.61 in the pre-test to 75.71 in the post-test following learning with the Teams Games Tournament (TGT) model assisted by Wordwall. The mean difference of 39.11 points was statistically significant ($t(27) = 13.88$, $p < 0.001$; 95% CI [33.33, 44.89]; Cohen's $d_z = 2.62$), with a normalized gain of 0.62 (moderate); all 28 students improved, and scores rose on all five indicators. Because the study used a one-group design without a comparison group, these findings show that improvement occurred during and after implementation of the model, but they do not allow the conclusion that the model was its cause.

SUGGESTIONS

For mathematics teachers, the TGT learning model assisted by Wordwall can be tried as an innovative and enjoyable alternative learning strategy, particularly for topics requiring analytical and reasoning skills such as SPLDV, while giving due attention to time management so that the games and tournament stages can proceed optimally and to monitoring less active students within groups. For future researchers, it is recommended that subsequent studies employ a design incorporating a control group (such as a two-group quasi-experimental design, for example comparing TGT with and without Wordwall), use several items per indicator so that indicator-level changes can be analyzed, test the equivalence of the pre-test and post-test item sets, broaden the scope of topics and educational levels examined, and investigate other relevant variables such as learning motivation or self-efficacy, in order to obtain a more comprehensive picture of the potential of the TGT learning model assisted by Wordwall in mathematics instruction.

REFERENCES

- [1] L. Serafini, A. Cammarota, G. Baldi, and M. Giannelli, "Building long-term consumer relationships with GenAI: The 4C Framework," 2026, *Elsevier*. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0007681326000753>
- [2] N. Nasran and M. I. Zakaria, "Implementation of 4C skills through problem solving: a study among mathematics teachers in Malaysia," *Journal of Education and ...*, 2025, [Online]. Available: <http://edulearn.intelektual.org/index.php/EduLearn/article/view/22621>
- [3] M. I. Zakaria, C. Hanri, S. H. Noer, and ..., "4C skills teaching activities for mathematics teachers: application of modified nominal group technique," ... *of education and ...*, 2025, [Online]. Available: <http://edulearn.intelektual.org/index.php/EduLearn/article/view/22337>
- [4] L. Serafini, A. Cammarota, G. Baldi, and M. Giannelli, "Building long-term consumer relationships with GenAI: The 4C Framework," 2026, *Elsevier*. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0007681326000753>

- [5] A. M. Amin, F. Karmila, R. Adiansyah, M. K. Mustami, and ..., "The Potential of the Integrated Project-Based and 4C-Scaffolding (Pj4CS) Learning Model to Increase Prospective Biology Teachers' Global 4C Skills," *Journal of Turkish ...*, 2026, [Online]. Available: <https://eric.ed.gov/?id=EJ1517402>
- [6] Kementerian Pendidikan dan Kebudayaan Republik Indonesia, *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 20 Tahun 2016 tentang Standar Kompetensi Lulusan Pendidikan Dasar dan Menengah*. Indonesia, 2016.
- [7] S. Syamsinar, F. N. Sa'diyyah, and M. Mulyana, "Influence of Digital Literacy on the Critical Thinking Ability of Mathematics Education Students in Evaluating Digital Information," *Kognitif: Jurnal Riset HOTS ...*, 2026, [Online]. Available: <https://etdci.org/journal/kognitif/article/view/4877>
- [8] B. B. M. Suryonegoro, W. Wardono, and ..., "The Analysis of Mathematical Critical Thinking Ability: A Meta-Analysis Review of Mathematical Disposition in Mathematics Learning," *Jurnal Riset Pendidikan ...*, 2026, [Online]. Available: <https://jurnal.uny.ac.id/index.php/jrpm/article/view/93872>
- [9] S. Syamsinar, F. N. Sa'diyyah, and M. Mulyana, "Influence of Digital Literacy on the Critical Thinking Ability of Mathematics Education Students in Evaluating Digital Information," *Kognitif: Jurnal Riset HOTS ...*, 2026, [Online]. Available: <https://etdci.org/journal/kognitif/article/view/4877>
- [10] B. B. M. Suryonegoro, W. Wardono, and ..., "The Analysis of Mathematical Critical Thinking Ability: A Meta-Analysis Review of Mathematical Disposition in Mathematics Learning," *Jurnal Riset Pendidikan ...*, 2026, [Online]. Available: <https://jurnal.uny.ac.id/index.php/jrpm/article/view/93872>
- [11] P. Güner and H. Yıldızlı, "The interplay between critical and mathematical thinking: mediating role of mathematics self-efficacy among pre-service mathematics teachers," *Innovation: The European Journal of Social ...*, 2026, doi: 10.1080/13511610.2026.2657562.
- [12] B. B. M. Suryonegoro, W. Wardono, and ..., "The Analysis of Mathematical Critical Thinking Ability: A Meta-Analysis Review of Mathematical Disposition in Mathematics Learning," *Jurnal Riset Pendidikan ...*, 2026, [Online]. Available: <https://jurnal.uny.ac.id/index.php/jrpm/article/view/93872>
- [13] R. H. Ennis, "A Logical Basis for Measuring Critical Thinking Skills," *Educational Leadership*, vol. 43, pp. 44–48, 1985.
- [14] Rendi Rendi, Marni Marni, Tia Neonane, and Mozes Lawalata, "Peran Logika Dalam Berfikir Kritis Untuk Membangun Kemampuan Memahami Dan Menginterpretasi Informasi," *Sinar Kasih: Jurnal Pendidikan Agama dan Filsafat*, vol. 2, no. 2, pp. 82–98, Apr. 2024, doi: 10.55606/sinarkasih.v2i2.313.
- [15] Y. Agustin and K. N. S. Effendi, "Analisis Kemampuan Berpikir Kritis Matematis Siswa SMP Pada Materi SPLDV," *Transformasi: Jurnal Pendidikan Matematika dan Matematika*, vol. 6, no. 2, pp. 121–132, 2022, doi: <https://ejournal.unibabwi.ac.id/index.php/transformasi/article/view/2222>.
- [16] B. Benyamin, Abd. Qohar, and I. M. Sulandra, "Analisis Kemampuan Berpikir Kritis Siswa SMA Kelas X Dalam Memecahkan Masalah SPLTV," *Jurnal Cendekia : Jurnal Pendidikan Matematika*, vol. 5, no. 2, pp. 909–922, Apr. 2021, doi: 10.31004/cendekia.v5i2.574.
- [17] I. Ainurrohman, T. Y. E. Siswono, and W. Wiryanto, "Pengaruh Model Pembelajaran TGT Berbantu Media Wordwall terhadap Kemampuan Berpikir Kritis Siswa," *Ideas: Jurnal Pendidikan, Sosial, dan Budaya*, vol. 10, no. 2, pp. 267–274, May 2024, doi: 10.32884/ideas.v10i2.1725.
- [18] I. F. Rohman, A. S. Anwar, and L. Nur'aeni, "Pengaruh Penggunaan Media Wordwall dalam Pembelajaran Matematika Terhadap Kemampuan Berpikir Kritis Di Sekolah Dasar," *Pendas: Jurnal Ilmiah Pendidikan Dasar*, vol. 9, no. 4, pp. 322–330, 2024, doi: <https://doi.org/10.23969/jp.v9i04.19104>.
- [19] M. Masni, A. I. P. Abrar, N. Yuliany, L. A. Mattoliang, and A. F. Majid, "Efektivitas Model Pembelajaran Temas Games Tournament (TGT) Berbantuan Aplikasi Wordwall Terhadap Kemampuan Berpikir Kritis Matematika Siswa," *Pedagogy: Jurnal Pendidikan Matematika*, vol. 11, no. 2, pp. 424–435, Jun. 2026, doi: 10.30605/npfeta15.
- [20] Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*, 2nd ed. Bandung: Alfabeta, 2020.
- [21] W. R. Shadish, T. D. Cook, and D. T. Campbell, *Experimental and Quasi-Experimental Design for Generalized Causal Inference*. Boston: Houghton Mifflin, 2002.
- [22] J. Cohen, *Statistical Power Analysis For The Behavioral Sciences (2nd ed.)*. New York: Academic Press, 1988.
- [23] R. R. Hake, "Analyzing Change/ Gain Scores," 1999.
- [24] N. I. Lestari, A. Razak, L. Lufri, Z. Zulyusri, and F. Arsih, "Pengaruh Model Pembelajaran Inkuiri Terbimbing Terhadap Keterampilan Berpikir Kritis Peserta Didik Pada Pembelajaran Biologi," *Bioilmi: Jurnal Pendidikan*, vol. 8, no. 1, pp. 17–30, Jan. 2022, doi: 10.19109/bioilmi.v8i1.12917.