EDUMATSAINS
JURNAL PENDIDIKAN MATEMATIKA DAN SAINS
E-ISSN:2527-7235
Volume 11 Issue. 2, October 2026, pp 422-438
http://ejournal.uki.ac.id/index.php/edumatsains

Wayground-Integrated Team Assisted Individualization and Students' Mathematical Literacy Skills in Statistics

Selviana Anggreani[1*], Lola Mandasari[2]

[1,2]Department of Mathematics Education, Faculty of Tarbiyah and Teacher Training, Universitas Islam Negeri Sumatera Utara

e-mail: selviana0305222043@uinsu.ac.id , lolamandasari@uinsu.ac.id

Article Info
Article history:
Received: September 29th, 2026
Revised: October 29th, 2026
Accepted: October 30th, 2026
Available online: October 31st, 2026

https://doi.org/10.33541/edumatsains.v11i2.8764

Abstract

This research examines whether the Wayground-integrated Team Assisted Individualization (TAI) cooperative learning model is associated with differences in the mathematical literacy of eighth-grade students at SMP IT Darul Muqomah Al-Khairiyyah, focusing specifically on statistics content. A quasi-experimental method with a Pretest-Posttest Control Group Design was employed for this quantitative investigation. From a population of 277 students spread across nine classes, 60 participants were selected by cluster sampling and evenly split between an experimental group (class VIII-C) and a control group (class VIII-D), with 30 students in each. To measure mathematical literacy, an essay test instrument that had undergone validity, reliability, difficulty-index, and discrimination-index testing was administered. Following the intervention, the experimental group posted a mean posttest score of 53.67, exceeding the control group's 46.93. Because the two classes differed at pretest (33.33 vs. 36.17), the main analysis was an ANCOVA with the pretest as covariate, which showed a significant difference in adjusted posttest means (53.73 vs. 46.87; $F(1, 57) = 27.72$, $p < 0.001$; partial $\eta^2 = 0.33$). A supplementary unadjusted dummy-variable regression gave a consistent result ($F = 29.380$, $p < 0.001$; $R^2 = 0.336$, $f^2 \approx 0.51$). Given the quasi-experimental design and the combined delivery of TAI and Wayground, the findings are interpreted as differences associated with the combined intervention, not as causal effects of either component.

Keywords: Team Assisted Individualization (TAI), Wayground, Mathematical Literacy, Statistics.

1. Introduction

Quality human resource development relies heavily on the role played by education capable of competing in the global era. Education is intended to give learners the chance to develop their full

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

potential, spanning knowledge, skills, and attitudes (Fajartiani et al., 2024). Law Number 20 of 2003 on the National Education System frames education's purpose in two parts: building up abilities, and forging the kind of character and civilization befitting a nation with dignity (Sari et al., 2021). Therefore, the learning process in schools needs to be designed systematically and meaningfully so that it can develop learners' higher-order thinking skills.

In facing the demands of the 21st century, learning does not only emphasize mastery of material, but also the development of various important competencies, such as critical thinking, problem solving, communication, collaboration, and literacy (Asri et al., 2023). Mathematical literacy stands out as a particularly important competency in learning mathematics; it refers to a person's capacity to formulate, apply, and interpret mathematics in various everyday situations (Rambe et al., 2023). This ability enables students to understand real problems and make logical decisions based on mathematical reasoning (Fauziah et al., 2021).

The Programme for International Student Assessment (PISA) framework of the OECD (2023) defines mathematical literacy as an individual's capacity to reason mathematically and to formulate, employ, and interpret mathematics to solve problems in a variety of real-world contexts. The framework describes the construct through three components: the mathematical processes (formulating situations mathematically; employing mathematical concepts, facts, and procedures; and interpreting, applying, and evaluating mathematical outcomes), the content categories (quantity, uncertainty and data, change and relationships, and space and shape), and the contexts in which problems are situated (personal, occupational, societal, and scientific) (OECD, 2023). The uncertainty and data content, which is closely related to statistics material, is one aspect that requires special attention, given that the ability to process, present, and interpret data is a skill that is increasingly needed in today's digital era (Rahma & Reflina, 2023).

However, various research results show that Indonesian students' mathematical literacy skills are still relatively low (Mashuri et al., 2023). Indonesian students' achievement in mathematical literacy remains below the international average, with a mean of 366 points in PISA 2022 compared with the OECD average of 472 points (OECD, 2023), especially when it comes to context-based problem solving (Putrawangsa & Hasanah, 2022), which suggests that mathematics learning in schools has not yet fully brought students' mathematical literacy skills to an optimal level (Anwar, 2018). One factor that may contribute to this is that teaching still leans heavily on conventional, teacher-centered methods, which may leave students less accustomed to thinking critically and reflectively (Mandasari, 2016).

The problem of low mathematical literacy is also evident in mathematics learning at the Junior High School (SMP) level, particularly on statistics material (Maryati et al., 2019). Statistics plays a significant role in building mathematical literacy, since it connects directly to the ability to process, present, and interpret data across different life contexts (Oktaviranda & Asmara, 2021). In reality, however, many students still experience difficulty in understanding word problems, interpreting data presented in tables or diagrams, and drawing conclusions based on available data (Islamiati & Putra, 2025).

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

The low ability of students in statistics material cannot be separated from the lack of variation in the learning models used (Latifah & Afriansyah, 2021). One-directional learning tends to make students passive and gives them little opportunity to construct their own understanding. One learning model that can encourage student activeness, cooperation, and attention to individual differences in ability is the Team Assisted Individualization (TAI) cooperative learning model, which combines individual learning and group work so that students can help one another and learn according to their respective abilities (Endriani & Rakhmawati, 2019).

The Team Assisted Individualization (TAI) cooperative learning model was developed as a combination of individual and group learning, in which students first study independently using provided worksheets, then discuss their work in small, heterogeneous groups (Slavin, 2010). Through these stages, higher-ability students can help group members who are experiencing difficulty, creating a mutually beneficial learning interaction without neglecting individual responsibility for mastering the material. These characteristics make the TAI model well suited to statistics material, since students can practice interpreting and processing data individually before refining their understanding through group discussion.

In addition to selecting an appropriate learning model, the use of technology in mathematics learning also needs attention, as it can create more engaging and interactive learning (Sari & Munir, 2024). One application that can be utilized is Wayground (formerly Quizizz), a game-based application that can increase student motivation and engagement. The use of Wayground allows students to work on questions interactively and receive immediate feedback, which helps students understand the material better (Singgi et al., 2026). Beyond motivation and engagement, several Wayground features used in this study are theoretically relevant to the processes of mathematical literacy. The Lesson feature (presentation slides) allowed the teacher to present statistical situations in texts, tables, and images before students worked on problems, which may support the formulating process (OECD, 2023); individual practice items (fill-in-the-blank and multiple-choice) with instant feedback allowed students to check and revise their solution steps, which may support the employing process; the placement test provided the information used to form heterogeneous teams in the TAI sequence; and the Reports feature allowed the teacher to identify the items and processes with which students struggled.

The use of gamification-based applications such as Wayground in mathematics learning reflects the underlying principle of joyful learning, in which game elements such as scores, time limits, and leaderboards can increase students' intrinsic motivation in working on mathematical literacy problems. The instant feedback feature of Wayground also allows students to immediately known whether their answers are correct, enabling them to reflect directly on the mistakes they made. By integrating the TAI cooperative learning model with the Wayground application, this approach is expected to foster a livelier, more collaborative, and more enjoyable classroom environment, thereby enabling students' mathematical literacy in statistics to develop to its full potential.

Although several previous studies have examined the effectiveness of the Team Assisted Individualization (TAI) model on students' learning outcomes or mathematical abilities (Sabillah

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

et al., 2019), these studies have generally concentrated on achievement or learning outcomes rather than on mathematical literacy. A broader look at the related literature shows three separate strands. First, TAI has mostly been examined for its effect on learning outcomes and achievement (Indriyani & Sujana, 2021; Sabillah et al., 2019; Tarim & Akdeniz, 2008). Second, game-based digital platforms such as Quizizz have been studied mainly in relation to learning interest and mathematical reasoning, often together with other cooperative models rather than TAI (Fauziah et al., 2021; Singgi et al., 2026). Third, research on mathematical and statistical literacy has largely documented students' difficulties or developed literacy items (Latifah & Afriansyah, 2021; Maryati et al., 2019; Oktaviranda & Asmara, 2021) rather than testing instructional interventions. Within the studies reviewed here, none combined TAI with a gamification-based platform to improve mathematical literacy in statistics. To close this gap, the present research examines a Wayground-integrated TAI intervention on grade VIII students' mathematical literacy skills in statistics, treating the intervention as one combined package.

Preliminary observations and interviews with two mathematics teachers at SMP IT Darul Muqomah Al-Khairiyyah indicated that grade VIII students' mathematical literacy skills in statistics appeared to be low (the pretest of this study points in the same direction: when the raw scores, with a maximum of 60, were converted to a 0–100 scale, 58 of the 60 students (96.7%) scored below the minimum mastery criterion of 72, namely 28 of 30 students in the control class and all 30 students in the experimental class). Teachers reported that students struggled to understand contextual problems, interpret data, and draw conclusions from the available data, and that cooperative learning models and technology-based media had not been used to their full potential. This description forms the basis for the present study, which aims to determine whether the Team Assisted Individualization (TAI) cooperative learning model integrated with Wayground is associated with a difference in students' mathematical literacy skills in statistics among grade VIII students of SMP IT Darul Muqomah Al-Khairiyyah relative to conventional learning, after adjusting for pretest scores, and if so, how large that difference is.

2. Methods

This research relied on quantitative data and was structured as a quasi-experiment, since the researcher could not control all conditions and external variables affecting students during the learning process in the field (Sugiyono, 2019). The study employed the Pretest-Posttest Control Group Design, a design built around two groups: one experimental group taught using the TAI cooperative learning model integrated with the Wayground application, and a control group taught using conventional learning. A pretest was administered to both groups prior to treatment, followed by a posttest afterward, allowing the researcher to compare the change in students' mathematical literacy skills between the two groups. The research design is illustrated as follows.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Group	Pretest	Treatment	Posttest
Experimental	O_1	X	O_2
Control	O_3	-	O_4

Notes:
O_1 = experimental class pretest;
O_2 = experimental class posttest;
O_3 = control class pretest;
O_4 = control class posttest;
X = TAI cooperative learning model integrated with the Wayground application;
– = conventional learning.

This study was conducted at SMP IT Darul Muqomah Al-Khairiyyah, Jl. Besar Glugur Rimbun Dusun 1 Gg. Seri, Kec. Pancur Batu, Kab. Deli Serdang, North Sumatra, in the odd semester of the 2026/2027 academic year, using statistics material. The population of the study consisted of all grade VIII students of SMP IT Darul Muqomah Al-Khairiyyah, totaling 277 students from 9 classes. The researcher selected the sample by applying a cluster sampling technique, involving 60 students divided into class VIII-C as the experimental class and class VIII-D as the control class, each consisting of 30 students.

Research ethics. Permission to conduct the research at the school was requested through a letter from the Faculty of Tarbiyah and Teacher Training, Universitas Islam Negeri Sumatera Utara Medan (No. B-5504/Un.11/ITK.IV.2/PP.00/07/2026, dated 28 July 2026), and the study was carried out with the permission of the school. After the research period (28 July to 14 August 2026), the school principal issued a reply letter (No. 210/SMP-DMAK/VIII/2026, dated 14 August 2026). The students took part as part of their regular classroom learning. Through the class messaging group, the students and their parents were informed of the purpose of the research, that the data would be used for an article, that students' names would not be published, and that students needed to bring mobile phones for the research activities, and both the students and their parents expressed their agreement in that group. Parental consent and student assent were therefore obtained informally rather than through signed forms, and no approval from an institutional ethics committee was obtained; this is acknowledged as a limitation. Students' names are not reported in this article, and the data were used only for research purposes.

To measure mathematical literacy skills, the researcher used a four-item essay test grounded in four indicators, namely formulating problems, applying mathematical concepts and procedures, interpreting the results of problem solving, and mathematically communicating those results on statistics material, with the first three indicators corresponding to the PISA mathematical processes of formulating, employing, and interpreting (OECD, 2023) and the fourth reflecting the

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

communication of results. The fourth indicator is not one of the three PISA processes; it was added as a study-specific indicator, in line with the mathematical communication construct put forward by Daulay et al. (2025). The instrument was examined for validity, reliability, item difficulty, and discriminating power with the assistance of IBM SPSS Statistics 27, as reported below. Each item consisted of five sub-questions on one statistics topic (items 1–4: mean; median and mode; range; and quartiles, respectively): (a) understanding the problem context (2 points), (b) formulating the problem mathematically (3 points), (c) employing mathematical concepts and procedures (2 points), (d) interpreting the results (3 points), and (e) communicating the results (5 points), so that each item was scored 0–15 and the maximum total score was 60. Sub-questions (b) to (e) correspond to the four indicators above, whereas sub-question (a) is a preparatory step in which students show that they understand the situation. All scores reported in this article are raw total scores out of 60.

The test instrument was first checked by expert lecturers through expert judgment before being applied in the study, with the review focused on how well the items reflected the mathematical literacy indicators and on the precision of the language and context used. Having been confirmed as content-valid, the instrument moved to a pilot stage involving students not included in the sample, aimed at determining empirical validity through Pearson's Product Moment correlation and reliability through Cronbach's Alpha, difficulty level, and discriminating power for each item. Item-level results are presented in Table 1. Of the five items that were piloted ($n = 30$; r-table = 0.361 at $\alpha = 0.05$), item 5 was invalid ($r = 0.352$) and was discarded, so that four items were used in the study. For the four retained items, the item-total correlations were 0.861, 0.797, 0.881, and 0.860; the difficulty indices were 0.753 (easy), 0.609 (moderate), 0.493 (moderate), and 0.293 (difficult); and the discrimination indices (corrected item-total correlations) were 0.755, 0.658, 0.755, and 0.743 (good to very good). Cronbach's alpha of the four-item test was 0.870 (very reliable). Responses were scored with an analytic rubric that assigned maximum scores of 2, 3, 2, 3, and 5 points to the five sub-questions of each item, with descriptors for every score level (0–2 for sub-questions (a) and (c), 0–3 for (b) and (d), and 0–5 for (e)); for example, the top score on sub-question (e) required a complete, clear, and sequential conclusion that used appropriate mathematical terms and fitted the problem context, whereas a score of 0 meant that no conclusion was written or that it was entirely irrelevant. The full rubric and answer key can be provided as supplementary material. Students had 80 minutes to complete the test. All responses were scored by the researcher, who also taught both classes, so inter-rater agreement could not be estimated.

Table 1. Item Validity, Difficulty Index, and Discrimination Index of the Mathematical Literacy Test

Item	Topic	r	Validity	Difficulty (P)	Discrimination (D)
1	Mean	0.861	Valid	0.753 (easy)	0.755 (very good)
2	Median and mode	0.797	Valid	0.609 (moderate)	0.658 (good)
3	Range	0.881	Valid	0.493 (moderate)	0.755 (very good)
4	Quartiles	0.860	Valid	0.293 (difficult)	0.743 (very good)
5	Discarded	0.352	Not valid	0.569 (moderate)	0.047 (low)

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Note. n = 30 (pilot); r = Pearson correlation between the item score and the total score (r-table = 0.361); for items 1–4 the total is the four-item total and for item 5 the five-item total; D = corrected item-total correlation. Item 5 was discarded. Cronbach's alpha of items 1–4 = 0.870.

Several stages made up the process of carrying out this research. The preparation stage involved the researcher designing learning tools, including a Lesson Plan (RPP), Student Worksheets (LKPD), and interactive questions on the Wayground application adapted to the syntax of the TAI learning model. At the implementation stage, both classes began with a pretest designed to measure students' starting mathematical literacy. From there, the experimental class moved into a round of instruction shaped by the TAI stages, namely a placement test, team formation (teams), delivery of material by the teacher (teaching group), group study and discussion (team study) followed by interactive practice through the Wayground application, scoring and team recognition (team scores and team recognition), and whole-class discussion (whole-class units). By contrast, teaching in the control class followed the conventional approach — mostly lecturing punctuated by question-and-answer exchanges. In the final stage, both classes were given a posttest using the same instrument to measure students' mathematical literacy skills after treatment. The treatment lasted five meetings of 120 minutes (3 × 40 minutes) each, from 28 July 2026 to 14 August 2026. Both classes were taught by the researcher and covered the same statistics material, and the control class likewise had five meetings. Implementation followed the lesson plan prepared for each meeting, but no formal observation checklist or fidelity percentage was recorded, and this is acknowledged as a limitation. Because TAI and Wayground were implemented together as a single package, this design cannot separate the contribution of each component, and the results are interpreted at the level of the combined intervention.

Data analysis began with descriptive statistics (mean, standard deviation, minimum value, and maximum value) for the pretest and posttest data of both classes. Baseline equivalence of the two classes was examined with an independent-samples t-test of the pretest scores, which showed that the experimental class started significantly lower than the control class ($t(58) = -2.16$, $p = 0.035$; Cohen's $d = -0.56$). Because the pretest means differed (Table 2), the main analysis was an analysis of covariance (ANCOVA), implemented as a multiple regression of posttest scores on group and pretest scores: where is the posttest score, is the group dummy (experimental class coded 1 and control class coded 0), and is the pretest score, centered at the grand mean. Thus is the difference in adjusted posttest means between the two classes. For transparency, an unadjusted simple regression of posttest scores on the dummy variable, is also reported. Regression assumptions were checked on the model residuals rather than on the raw scores of each group: normality of the standardized residuals (Shapiro-Wilk test and normal Q-Q plot), homoscedasticity (residuals-versus-predicted plot and Levene's test of equality of error variances), linearity of the pretest-posttest relationship (scatterplot), and homogeneity of regression slopes (a non-significant group × pretest interaction). Independence of residuals could not be guaranteed by the design because students were taught in two intact classes, so possible dependence within classes is acknowledged as a limitation. All analyses used SPSS version 27 at a significance level of. The statistical hypotheses were (there is no difference in adjusted posttest means between the two classes) and (there is a difference), with the criterion that if Sig. < 0.05, then is rejected and is accepted. Effect

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

sizes were reported as partial η^2 and f^2 = partial $\eta^2/(1 -$ partial $\eta^2)$ for the ANCOVA, and as $f^2 = R^2/(1 - R^2)$ for the unadjusted model, together with adjusted means and the 95% confidence interval of their difference.

3. Results and Discussion

3.1.Results

This study sets out to examine the difference in mathematical literacy skills associated with the Team Assisted Individualization (TAI) cooperative learning model integrated with the Wayground application among grade VIII students of SMP IT Darul Muqomah Al-Khairiyyah on statistics material. The research sample consisted of class VIII-C as the experimental class and class VIII-D as the control class, each consisting of 30 students.

Table 2. Comparison of Pretest and Posttest Results

Class	Test	N	Mean	SD	Min	Max
Experimental	Pretest	30	33.33	4.61	21	42
	Posttest	30	53.67	4.54	43	60
Control	Pretest	30	36.17	5.49	27	46
	Posttest	30	46.93	5.07	36	57

Table 2 shows that after learning was provided, the two classes had different mean mathematical literacy skills. All scores are raw scores out of a maximum of 60, and Table 2 also reports the standard deviation and the lowest and highest scores of each class. The experimental class obtained a mean posttest score of 53.67, higher than the control class at 46.93. The increase (pretest–posttest difference) was 20.34 in the experimental class and only 10.76 in the control class. This indicates that the experimental class, which started from a lower pretest mean, showed a greater increase in mathematical literacy skills than the control class; whether this increase can be attributed to the intervention is examined in the pretest-adjusted analysis below. The standard deviations were similar in the two classes (4.54 to 5.49), and the highest posttest score in the experimental class (60) equals the maximum possible score, so a ceiling effect cannot be ruled out.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Figure 1.
Mean Pretest and Posttest Score Comparison for the Experimental and Control Classes

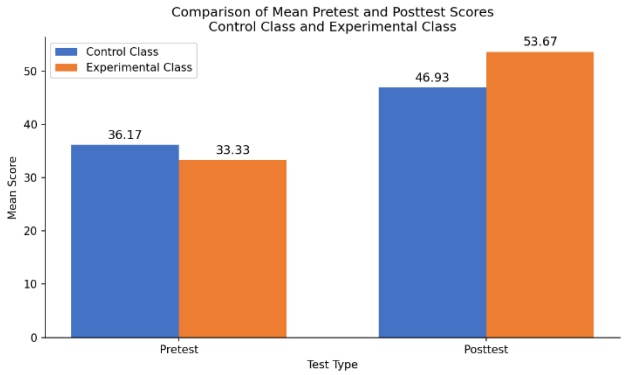


Figure 1 presents a visual comparison of the mean pretest and posttest scores of the experimental and control classes. As the graph shows, the control class started ahead at the pretest (a statistically significant difference), with a mean of 36.17 compared with 33.33 in the experimental class. However, after treatment, this situation reversed: the experimental class's posttest mean (53.67) came out above the control class's (46.93). This pattern shows that the rise in mathematical literacy skills was greater in the experimental class than in the control class, which is consistent with, but does not by itself demonstrate, a contribution from the TAI cooperative learning model integrated with Wayground to that rise.

With a general picture of both classes' pretest and posttest averages in hand, the analysis proceeded to prerequisite testing—namely normality and homogeneity checks—prior to hypothesis testing. These checks exist to confirm that the data meet the assumptions parametric statistical methods require, thereby lending statistical credibility to whatever results the analysis produces. Because the assumptions of regression concern the model residuals, these group-wise checks are preliminary, and the residual-based checks are reported together with the ANCOVA results.

Table 3. Shapiro-Wilk Normality Test Results

Data	Shapiro-Wilk Statistic	df	Sig.
Experimental Pretest	0.973	30	0.614
Control Pretest	0.959	30	0.296
Experimental Posttest	0.942	30	0.100
Control Posttest	0.975	30	0.678

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

The normality test on the mathematical literacy skills data was carried out using the Shapiro-Wilk test, with the results presented in Table 3. It was found that the significance value for the experimental class pretest reached 0.614, the control class pretest reached 0.296, the experimental class posttest reached 0.100, and the control class posttest reached 0.678. Because each of these values exceeded 0.05 (Sig. > 0.05), the data from all four groups can be considered normally distributed. Meeting this normality assumption is an important prerequisite before the data can be analyzed using parametric statistical techniques, such as the regression models used in this study. If the data were not normally distributed, parameter estimation and significance testing in the regression model would risk being biased, making conclusions about whether the two classes differ less statistically defensible. With this assumption satisfied, the analysis could proceed to the homogeneity test as the next prerequisite before hypothesis testing.

Table 4. Homogeneity Test Results (Levene's Test)

Test	Levene Statistic	df1	df2	Sig.
Pretest (Pre_Total)	1.091	1	58	0.301
Posttest (Post_Total)	0.095	1	58	0.759

The homogeneity test results for the mathematical literacy skills data are set out in Table 4. Levene's Test served as the method here, with 0.05 set as the threshold for significance. Referring to the “Based on Mean” results, the Levene Statistic obtained was 1.091 for the pretest data (=1, =58, Sig.=0.301) and 0.095 for the posttest data (=1, =58, Sig.=0.759). Since the significance values came out above 0.05, the variance across both groups can be regarded as homogeneous. This assumption of variance homogeneity is important to satisfy because the dummy-variable regression technique assumes that the variability of the data in both groups is relatively equal, so that the observed mean difference is not distorted by an unbalanced spread of data between groups. Should the two groups' variances turn out to be non-homogeneous, the standard error of the regression coefficient estimate would risk being inaccurate, which could affect the precision of the significance test. With this assumption satisfied, the analysis could proceed to hypothesis testing using regression with a dummy variable and the pretest as a covariate.

Because the regression assumptions concern the model residuals, the ANCOVA model met them: the standardized residuals were normally distributed (Shapiro-Wilk W = 0.979, p = 0.403); the error variances were equal across the two classes (Levene's test on the residuals, F = 0.089, p = 0.766; Breusch-Pagan test, p = 0.760); the pretest-posttest relationship was linear (quadratic term, p = 0.634); and the regression slopes were homogeneous (group × pretest interaction, p = 0.928). The Durbin-Watson statistic was 1.94, which indicates no serial correlation in the order of the data, but it cannot rule out dependence among students within the same class; no observation was influential (largest Cook's distance = 0.14).

ANCOVA results. The experimental class had a significantly lower pretest mean than the control class ($t(58) = -2.16$, $p = 0.035$), and the pretest correlated only weakly with the posttest ($r = -0.12$). With the pretest as covariate, the ANCOVA (Table 5) showed a significant group effect, $F(1, 57)$

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

= 27.72, $p < 0.001$, partial $\eta^2 = 0.33$. The adjusted posttest means were 53.73 (SE = 0.90) for the experimental class and 46.87 (SE = 0.90) for the control class, an adjusted difference of 6.85 points (95% CI [4.25, 9.46]; $f^2 = 0.49$). This is close to the unadjusted difference of 6.73 points because the pretest explained almost none of the variance in posttest scores (B = 0.04, p = 0.739).

Table 5. ANCOVA Results (Posttest Adjusted for Pretest)

Source	Sum of Squares	df	Mean Square	F	Sig.
Pretest	2.634	1	2.634	0.112	0.739
Group	651.699	1	651.699	27.724	<0.001
Error	1339.899	57	23.507	–	–
Total	2022.600	59	–	–	–

Table 6. Model Summary

Model	R	R Square	Adjusted R Square
1	0.580	0.336	0.325

Table 7. ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	680.067	1	680.067	29.380	0.000
Residual	1342.533	58	23.147	–	–
Total	2022.600	59	–	–	–

Table 8. Coefficients

Model	B	Std. Error	t	Sig.
(Constant)	46.933	0.878	53.431	0.000
Dummy	6.733	1.242	5.420	0.000

Tables 6 to 8 report a supplementary unadjusted simple regression of posttest scores on the group dummy, which does not control for pretest scores. The results in Table 7 show an F value of 29.380

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

at a significance level of $p < 0.001$, meaning that in this unadjusted model was rejected and accepted. This means that the posttest means of the two classes differ significantly in this unadjusted model. Based on Table 8, the regression equation = 46.933 + 6.733D was obtained, with the constant value of 46.933 representing the estimated mean of the control class () and the coefficient 6.733 representing the additional mean for the experimental class (), so that the estimated mean of the experimental class is 46.933 + 6.733 = 53.666, consistent with the data in Table 2. Based on Table 6, the R Square value of 0.336 indicates that the group variable accounts for 33.6% of the variance in posttest scores in this model ($f^2 = R^2/(1 - R^2) = 0.336/0.664 \approx 0.51$). Because this model does not control for the pretest scores, which differed between the classes at baseline, this value is limited to the variance in posttest scores associated with class membership and should not be read as the magnitude of a causal treatment effect (the pretest-adjusted ANCOVA is reported above in Table 5).

3.2.Discussion

From the data analysis, a clear picture emerged of how mathematical literacy skills compared between students taught with the Team Assisted Individualization (TAI) cooperative learning model integrated with Wayground and those taught through conventional methods. According to the descriptive statistics, the experimental class posted a mean posttest score of 53.67 for mathematical literacy skills, compared with just 46.93 in the control class, amounting to a 6.73-point difference. The group-wise preliminary tests indicated normality and homogeneity of variance, and the residual-based checks of the ANCOVA model, which are the relevant assumptions for regression, were also satisfied, so that the analysis could be carried out. In the pretest-adjusted ANCOVA the group effect was statistically significant ($F(1, 57) = 27.72$, $p < 0.001$; partial $\eta^2 = 0.33$), and the supplementary unadjusted model, in which the group variable accounted for 33.6% of the posttest variance ($R^2 = 0.336$), led to the same conclusion. Because of the quasi-experimental design, these results are interpreted below as a difference associated with the combined TAI–Wayground intervention rather than as a causal effect.

The pretest-adjusted ANCOVA gave a partial η^2 of 0.33, which corresponds to Cohen's $f^2 = \text{partial } \eta^2/(1 - \text{partial } \eta^2) \approx 0.49$ (adjusted difference = 6.85 points, 95% CI [4.25, 9.46]). Using Cohen's (1988) benchmarks for f^2 (small = 0.02, medium = 0.15, large = 0.35), this value falls in the large range. The supplementary unadjusted model gave a similar value ($R^2 = 0.336$; $f^2 = R^2/(1 - R^2) = 0.336/0.664 \approx 0.51$), but because that model ignores the baseline difference between the classes, it describes only the share of posttest variance associated with class membership and should not be read as the size of a causal treatment effect.

The higher posttest scores of the experimental class may be related to several features of the intervention, although this study did not measure the mechanisms involved and cannot test them. First, the TAI model positions students as active participants through the formation of small, heterogeneous groups, in which students can work together and help one another understand the material, while assessment remains individual. Whether this structure operated

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

through motivation, self-efficacy, or other processes was not measured here and remains a question for future research. In line with Slavin (2010), the TAI model combines the advantages of individual learning with team learning, so that each student holds responsibility for the group's success, which is theorized to support motivation and understanding. This is consistent with earlier findings reported by Tarim and Akdeniz (2008), which showed that the application of the TAI model had a positive effect on the mathematics achievement of elementary school students, as well as other studies Indriyani and Sujana (2021), Sabillah et al. (2019), and Siregar (2021), which likewise reported higher learning outcomes among students taught through cooperative models such as TAI.

Second, Wayground, employed as a learning medium, may also have supported students' mathematical literacy skills by presenting questions interactively and contextually, which gives students opportunities to practice understanding, interpreting, and communicating data. However, because the study had no TAI group without Wayground, the specific contribution of Wayground cannot be separated from that of TAI, and the findings should be read as the result of the combined intervention. Suryapuspitarini et al. (2018) made a similar point, arguing that interactive question presentation helps train critical thinking and problem-solving skills, which count among the key elements of mathematical literacy. In addition, the real-time learning-progress monitoring feature of Wayground can also make it easier for teachers to evaluate students' understanding of statistics material (Sari & Munir, 2024). In contrast, conventional learning in the control class tended to be more teacher-centered, which may have given students fewer opportunities to practice interpreting data and communicating the results of problem solving, although this explanation was not tested directly. Overall, the combined TAI–Wayground intervention was associated with higher posttest scores than conventional learning (and this difference remained after adjusting for the pretest), but this finding should not be taken as evidence of a causal effect or of the separate contribution of either component.

On the other hand, it should be acknowledged that the use of the Wayground application as a learning medium also carries a number of limitations that could potentially affect the research results. First, the effectiveness of this application is highly dependent on the availability of internet access and adequate digital devices, so that its implementation in schools with limited technological infrastructure may not produce similar results. Second, there is a possible novelty effect, in which the increase in motivation and learning outcomes among students in the early stages is caused by their interest in the new and unfamiliar medium, rather than solely by the advantages of the learning model itself. Third, the varying levels of digital literacy among students can affect the smoothness of using the application, so that students who are less accustomed to technology may experience technical obstacles that affect their performance. Therefore, the generalization of this study's findings needs to take into account the context of technology availability and students' familiarity with digital learning media.

In the unadjusted model, 66.4% of the variance in posttest scores was not accounted for by the group variable. The present data do not allow this residual variance to be attributed to

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

any specific source, because no other variables were measured or included in the model. Future studies could measure relevant covariates in advance and include them in the analysis, together with a pretest-adjusted comparison of the intervention and control groups, to examine how much additional variance they explain.

At baseline (pretest), the mean mathematical literacy score of the control class (36.17) was significantly higher than that of the experimental class (33.33; $t(58) = -2.16$, $p = 0.035$), as shown in Table 2. Because the groups were not equivalent at the pretest, baseline differences have to be controlled before the posttest difference can be interpreted. The unadjusted regression does not do this, and the larger raw gain of the experimental class (20.34 versus 10.76) may also reflect regression toward the mean or more room for improvement among students who started lower. The unadjusted analysis therefore cannot show that the increase in the experimental class was mainly due to the learning treatment. The appropriate test is the pretest-adjusted comparison from the ANCOVA (Table 5: adjusted means = 53.73 and 46.87, $F(1, 57) = 27.72$, $p < 0.001$, partial $\eta^2 = 0.33$). The difference therefore remained significant after controlling for the pretest, and because the pretest correlated only weakly with the posttest ($r = -0.12$), the adjusted difference (6.85 points) was close to the unadjusted one (6.73 points).

This finding has practical implications for mathematics teachers in designing learning for statistics material as well as other mathematics material that requires literacy skills. Teachers may consider combining cooperative learning models that emphasize social interaction and individual responsibility, such as TAI, with interactive technology-based learning media to create more meaningful learning experiences, bearing in mind that these implications rest on a single quasi-experimental study. In addition, immediate feedback provided through applications such as Wayground can be utilized by teachers as a formative assessment tool to continuously monitor the development of students' mathematical literacy skills, enabling teachers to promptly provide learning interventions tailored to students' needs.

4. Conclusion

The results of this quasi-experimental study show that students taught with the Team Assisted Individualization (TAI) cooperative learning model integrated with the Wayground application had a higher mean posttest score (53.67) than students taught with conventional learning (46.93), although the experimental class started with a lower mean pretest score (33.33 versus 36.17). An ANCOVA with the pretest as covariate showed a statistically significant difference after baseline adjustment (adjusted means 53.73 vs. 46.87; $F(1, 57) = 27.72$, $p < 0.001$; partial $\eta^2 = 0.33$), and a supplementary unadjusted dummy-variable regression gave a consistent result ($p < 0.001$). Because the study used intact classes and TAI and Wayground were implemented together, the findings indicate a difference associated with the combined intervention for grade VIII students of SMP IT Darul Muqomah Al-Khairiyyah and should not be interpreted as a causal effect of either component.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

The pretest-adjusted effect size from the ANCOVA was partial $\eta^2 = 0.33$ ($f^2 = 0.49$), a large effect by Cohen's benchmarks; the supplementary unadjusted model gave a similar value ($R^2 = 0.336$; $f^2 \approx 0.51$). With this caution, the TAI cooperative learning model integrated with Wayground can be considered a promising alternative approach to teaching mathematics, particularly statistics, for raising the mathematical literacy skills of grade VIII students at SMP IT Darul Muqomah Al-Khairiyyah, pending replication with stronger designs.

This study has several limitations that should be considered in future research. First, this study only involved a sample of 60 students drawn from only one school, limiting the reach of these findings when it comes to generalizing them across junior high school students as a whole. Second, this study focused only on statistics material, so it is not yet known whether similar differences would appear on other mathematics topics. Third, the relatively short duration of the study (five meetings of 120 minutes each) is also a limitation in observing whether the observed difference in students' mathematical literacy skills persists over the longer term. Fourth, the quasi-experimental design with intact classes, the baseline difference between the classes, and the combined delivery of TAI and Wayground limit causal inference and prevent the difference from being attributed to either component. Fifth, because students were nested in two intact classes, the independence assumption of the analysis may not be fully met. Sixth, the researcher both taught the two classes and scored the responses; no formal fidelity checklist or inter-rater agreement was available; and parental consent and student assent were obtained only informally, through group messages rather than signed forms, which may have introduced bias and ethical limitations. Seventh, the same instrument was used for the pretest and the posttest, the scorer was not blind to class membership, and the highest posttest score in the experimental class reached the maximum possible score (60), so that test-retest effects, scoring bias, and a ceiling effect cannot be ruled out; moreover, with only two intact classes, the reported p-values and effect sizes are likely to be optimistic. For this reason, subsequent research should aim to include a broader sample from a number of schools, a wider range of mathematics topics, a longer study duration, designs that separate the contribution of TAI from that of Wayground, independent teachers and scorers, documented implementation fidelity, parallel test forms, blinded scorers, formal consent procedures, and measured covariates such as learning motivation and student learning independence in order to obtain more comprehensive results.

5. Acknowledgments

The authors express their praise and thankfulness to God Almighty for the mercy and blessings He has bestowed, which made it possible to complete this article. The authors also wish to thank the supervising lecturer, the school principal, the mathematics teacher, and all students of SMP IT Darul Muqomah Al-Khairiyyah who helped and participated in the implementation of this research. The authors also extend sincere thanks to their parents for their prayers, love, support, and sacrifice, which have continually been a source of strength in completing this research.

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

6. References

Anwar, N. (2018). Peran kemampuan literasi matematis pada pembelajaran matematika abad-21. *PRISMA, Prosiding Seminar Nasional Matematika*, *1*, 364–370.

Asri, I. H., Lasmawan, I. W., & Suharta, I. G. P. (2023). Kompetensi abad 21 sebagai bekal menghadapi tantangan masa depan. *Kappa Journal*, *7*(1), 97–107. https://doi.org/10.29408/kpj.v7i1.12999

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Daulay, L. A., Zahari, C. L., Harahap, R., & Riani, D. (2025). JUCAMA learning model: A pathway to improved mathematical communication. *Axiom: Jurnal Pendidikan dan Matematika*, *14*(2), 230–244.

Endriani, R., & Rakhmawati, F. (2019). Perbedaan kemampuan pemecahan masalah dan berpikir kreatif matematika siswa yang diajarkan dengan model pembelajaran problem based learning dan pembelajaran kooperatif tipe two stay-two stray pada materi sistem persamaan linear tiga variabel kelas X SMA. *Axiom: Jurnal Pendidikan dan Matematika*, *8*(2). https://doi.org/10.30821/axiom.v8i2.6335

Fajartiani, T., Habibi, A., Rosalina, D., & Karsiwan, K. (2024). Peran pendidikan dalam meningkatkan kualitas sumber daya manusia. *JEALO Jurnal Administrasi Pendidikan*, *6*(1), 9–17.

Fauziah, N. S., Juariah, J., Nuraida, I., & A, T. T. W. (2021). Penalaran matematis melalui group to group exchange berbantuan Quizizz. *Jurnal Analisa*, *7*(2), 146–154. https://doi.org/10.15575/ja.v7i2.9601

Indriyani, K. W., & Sujana, I. N. (2021). Penerapan model pembelajaran kooperatif tipe Team Assisted Individualization (TAI) untuk meningkatkan aktivitas dan hasil belajar siswa pada mata pelajaran ekonomi di SMA Negeri 2 Denpasar. *Ekuitas: Jurnal Pendidikan Ekonomi*, *9*(2), 230–239. https://doi.org/10.23887/ekuitas.v9i2.30840

Islamiati, N., & Putra, I. S. (2025). Analisis kesulitan siswa SMP dalam menyelesaikan soal cerita matematika. *JIIBAS: Jurnal Ilmu Pendidikan, Bahasa, dan Sastra*, *1*(1), 9–15. https://doi.org/10.66371/jiibas.v1i1.20

Latifah, T., & Afriansyah, E. A. (2021). Kesulitan dalam kemampuan pemecahan masalah matematis siswa pada materi statistika. *JARME: Journal of Authentic Research on Mathematics Education*, *3*(2), 134–150. https://doi.org/10.37058/jarme.v3i2.3207

Mandasari, L. (2016). Peningkatan kemampuan berpikir kreatif matematis siswa SMA melalui problem based learning menggunakan software Autograph. *Jurnal As-Salam*, *1*(1), 143–152.

Maryati, I., Priatna, N., & Martadiputra, B. A. P. (2019). Analysis of madrasah Tsanawiyah students for statistical literacy skills. *Journal of Physics: Conference Series*, *1280*(4), 042016. https://doi.org/10.1088/1742-6596/1280/4/042016

Mashuri, S., Kurniawan, R., & Jahring. (2023). Kemampuan literasi matematis siswa dalam menyelesaikan soal berorientasi PISA konten quantity. *EDU-MAT: Jurnal Pendidikan Matematika*, *11*(1). https://doi.org/10.20527/edumat.v11i1.15417

OECD. (2023). *PISA 2022 assessment and analytical framework*. OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia

Oktaviranda, A., & Asmara, A. (2021). Kemampuan literasi melalui pengembangan soal pada materi statistika. *Jurnal Tarbiyah*, *30*(1), 15–26. https://doi.org/10.30829/tar.v30i1.2244
Putrawangsa, S., & Hasanah, U. (2022). Analisis capaian siswa Indonesia pada PISA dan urgensi kurikulum berorientasi literasi dan numerasi. *EDUPEDIKA*, *1*(1), 1–12. https://doi.org/10.60004/edupedika.v1i1.1
Rahma, F. L., & Reflina. (2023). Analisis kemampuan literasi numerasi dalam menyelesaikan soal Programme for International Student Assessment (PISA). *Jurnal Karya Pendidikan Matematika*, *10*(1), 11–20. https://doi.org/10.26714/jkpm.10.1.2023.11-20
Rambe, L. H., Malau, L., Manurung, M. E. R., Lubis, N. A., & Denti, W. K. (2023). Literasi matematika sebagai kompetensi utama di abad 21 dan implikasinya pada pendidikan di sekolah dasar. *Jurnal Pendidikan Guru Sekolah Dasar*, *1*(1), 10. https://doi.org/10.47134/pgsd.v1i1.109
Sabillah, F. S., Yensy, N. A., & Maizora, S. (2019). Penerapan model pembelajaran kooperatif tipe TAI (Team Assisted Individualized) berbantuan LKPD berbasis saintifik untuk meningkatkan hasil belajar pada siswa kelas VIII.2 SMPN 4 Kota Bengkulu. *Jurnal Penelitian Pembelajaran Matematika Sekolah*, *3*(1), 110–119.
Sari, A. P., & Munir, M. (2024). Pemanfaatan teknologi digital dalam inovasi pembelajaran untuk meningkatkan efektivitas kegiatan di kelas. *Digital Transformation Technology (Digitech)*, *4*(2), 977–983. https://doi.org/10.47709/digitech.v4i2.5127
Sari, I. M., Dewi, F. A., Fadila, N., & Rivadah, M. (2021). Analisis implementasi kebijakan pendidikan UU No. 20 Tahun 2003 terhadap pendidikan nasional di Indonesia. *Jurnal Soshum Insentif*, *4*(1), 98–103. https://doi.org/10.36787/jsi.v4i1.552
Singgi, A. N., Mangadi, R., Naomi, Mandacan, M., & Stefanus, R. (2026). Pemanfaatan aplikasi Quizizz sebagai media digital untuk meningkatkan minat belajar sekolah menengah atas (SMA). *Harmoni Pendidikan: Jurnal Ilmu Pendidikan*, *3*(1). https://doi.org/10.62383/hardik.v3i1.2658
Siregar, T. J. (2021). Peningkatan kemampuan pemecahan masalah matematis dan keterampilan sosial siswa SMP melalui pembelajaran kooperatif tipe STAD. *Axiom: Jurnal Pendidikan dan Matematika*, *10*(1), 97–109. https://doi.org/10.30821/axiom.v10i1.9265
Slavin, R. E. (2010). *Cooperative learning: Teori, riset dan praktik*. Nusa Media.
Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
Suryapuspitarini, B. K., Wardono, & Kartono. (2018). Analisis soal-soal matematika tipe higher order thinking skill (HOTS) pada Kurikulum 2013 untuk mendukung kemampuan literasi siswa. *PRISMA, Prosiding Seminar Nasional Matematika*, *1*, 876–884.
Tarim, K., & Akdeniz, F. (2008). The effects of cooperative learning on Turkish elementary students' mathematics achievement and attitude towards mathematics using TAI and STAD methods. *Educational Studies in Mathematics*, *67*(1), 77–91. https://doi.org/10.1007/s10649-007-9088-y

This is an open access article under the HYPERLINK "https://creativecommons.org/licenses/by/4.0/" *CC BY 4.0 license.*
Copyright ©2025 by Author. Published by Universitas Kristen Indonesia