Educational Journal of Learning Technology
Vol. 4, No. 2, pp. 402-413, January 2027
https://10.58230/edutech.v4i1.156
Received October 25, 2026; Revised November 06, 2026; Accepted November 11, 2026

Change in Qur'anic Reading Literacy Following the Combined Application of the Tahsin and Talaqqi Methods: A One-Group Pretest-Posttest Study at Ma'had al-Jami'ah

Nur Hatia[1], Sukirman[2], Muhammad Zuljalal Al-Hamdany[3]
[1,2,3]Universitas Islam Negeri Palopo
Agatis Street, Balandai Subdistrict, Bara District, Palopo City, South Sulawesi, Indonesia
nurhatia2004@gmail.com

Abstract. This study examined changes in Quranic reading literacy among UIN Palopo students who joined the Student Pesantren Program at Ma'had Al-Jami'ah and estimated the magnitude of that change. A quantitative, pre-experimental one-group pretest–posttest design was used. Twenty-one students were selected from 126 program participants through purposive sampling, and the target sample size was calculated using the Slovin formula (e = 20%). We measured reading literacy with an oral Qur'an reading test scored on four aspects (makharijul huruf, tajwid, fluency, and fashahah), each rated 1–4 and converted to a 100-point scale. Data were analyzed using descriptive statistics, the Shapiro–Wilk test, the Wilcoxon signed-rank test, and the effect size r. The mean score rose from 59.82 (SD = 7.01) at pretest to 82.44 (SD = 6.13) at posttest, a mean gain of 22.62 points, and all 21 students improved. The Wilcoxon test showed a significant difference ($Z = -4.01$, $p < .001$) with a large effect size ($r = .88$), and students in the poor and adequate categories moved into the good and very good categories. Limitations include the lack of a control group and the use of different pretest/posttest passages. The findings indicate a strong association between the combined methods and improved reading literacy, though causation cannot be established.

Keywords: Tahsin Method; Talaqqi Method; Quranic Reading Literacy; University Students; Ma'had Al-Jami'ah.

1. Introduction

Reading the Qur'an correctly is a fundamental religious competence for every Muslim and the primary gateway to understanding its teachings. In Islamic education, this competence transcends the mere decoding of written letters; it encompasses four essential aspects: accurate pronunciation of letters (makharijul huruf), correct application of recitation rules (tajwid), fluency, and clarity of delivery, including proper stopping and starting (fashahah)(Sastera et al., 2026; Tambusai et al., 2023). In this article, Quranic reading literacy refers to the combined mastery of these four aspects. This literacy is rooted in the revelation itself, with the divine command to read (Q.S. Al-'Alaq/96:1–5; *Qur'an Kemenag*, n.d.) and to recite the Qur'an deliberately and distinctly (tartil; Q.S. Al-Muzzammil/73:4; *Qur'an Kemenag*, n.d.-b).

Despite this emphasis, reading ability among university students remains uneven. At UIN Palopo, empirical findings make this issue evident. A study by Abdillah (2025) on the Student Pesantren Program at UIN Palopo found that during the initial placement test, many students still read haltingly and frequently violated tajwid rules. Specifically, of the 55 sampled students, none reached the "Very Good" category; 38.1% were classified as "Poor" and 50.9% as "Fair" (Abdillah, 2025). This is a significant concern for an Islamic university, where basic Qur'an reading is expected as part of students' religious competence. Differences in educational background, limited habituation, less effective learning methods, and discontinuous practice are plausible contributors.

©Authors
This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

In response, UIN Palopo runs the Student Pesantren Program through its Ma'had Al-Jami'ah technical unit, which operates in a residential environment to support religious habituation (Muldiah & Wulandari, 2025). However, it must still address the wide range of students' initial reading abilities (Nurzykila, 2026).

To address this problem, two instructional methods are commonly used: Tahsin and Talaqqi. Clarifying their distinct functions is crucial to understanding their combined efficacy. Tahsin functions primarily as a quality-refinement process; it sets the criteria for correct recitation—accurate makharijul huruf, applied tajwid, fluency, and fashahah—and provides structured drills through which the learner's production is brought toward those criteria (Siagian & Zailani, 2021). Talaqqi, on the other hand, serves as the pedagogical mechanism through which this refinement is delivered: the teacher models the recitation, the learner listens and imitates, and then recites before the teacher to receive immediate, individualized correction (Akil & Syafi'i, 2025; Muktafi & Umam, 2022).

The combination of Tahsin and Talaqqi is considered more appropriate than using either method in isolation because they are complementary rather than substitutable. Tahsin supplies the content and quality standard of recitation, but on its own, it can proceed without individualized verification of each learner's production. Talaqqi provides the pedagogical mechanism—direct modeling, listening, and immediate correction—but without a structured quality framework, it may not systematically address tajwid and makharijul huruf. Combining them allows Talaqqi's face-to-face, corrective loop to deliver Tahsin's quality standards, ensuring each learner's errors are diagnosed against explicit criteria rather than corrected incidentally. The rationale for this combined approach can be explained mechanistically by two learning theories. From a behaviorist perspective, the repeated drill–feedback cycles function as stimulus–response–reinforcement sequences: the teacher's model is the stimulus, the learner's imitation is the response, and immediate correction reinforces accurate articulation over successive trials (Darmayanti et al., 2024). From a social constructivist perspective, Talaqqi operationalizes Vygotsky's zone of proximal development: the teacher provides scaffolded support just beyond the learner's independent performance, and gradually withdraws it as the learner's recitation becomes accurate (Puntambekar, 2022; Tohari & Rahman, 2024; Vygotsky, 1978).

To align this theoretical intervention with the measured outcomes, the intervention stages were explicitly linked to the four aspects of Quranic reading literacy. The Tahsin stage specifically targeted makharijul huruf (articulation drills with teacher modeling) and tajwid (rule identification and applied practice). Meanwhile, the Talaqqi stage targeted fluency (repeated listening–imitation–recitation cycles) and fashahah (guided attention to pausing and resuming, with immediate correction). Because these four aspects were practiced and corrected within the same instructional cycle, the posttest measures change across all four components rather than in a single dimension.

Previous studies highlight the benefits of each method. Tahsin has been linked to better reading quality among primary school pupils (Ismail et al., 2023) and madrasah students (Fitri et al., 2024)while Talaqqi has been associated with better tahsin ability among school students (Fauziah & Komariah, 2023) and examined in qualitative work on talaqqi sama' (face-to-face listening and recitation) (Taringan, 2026). Studies on intensive teaching of makharijul huruf similarly report gains in reading accuracy (Azifah et al., 2026; Ilahi & Fatihah, 2026). However, much of this evidence concerns a single method, school-level learners, or qualitative designs.

This study addresses this gap by examining the combination of Tahsin and Talaqqi rather than either method in isolation. Furthermore, it focuses on university students in a residential Ma'had Al-Jami'ah, a population and setting rarely addressed in the quantitative literature. In addition, it reports an effect size (using the Wilcoxon Signed-Rank Test), allowing the magnitude of change to be compared across studies rather than only its statistical significance. Finally, it links the intervention stages explicitly to the four measured components of Quranic reading literacy. The novelty lies in combining two methods and measuring effect size among university students in a *ma'had* (Islamic boarding institution) setting. Together, these features address a gap that previous studies have not yet filled.

This study therefore examined whether students' Quranic reading literacy in the Student Pesantren Program at Ma'had Al-Jami'ah UIN Palopo changed significantly after the combined application of Tahsin and Talaqqi. It specifically addressed two research questions: whether students' reading scores differed significantly before and after the intervention, and what the effect size was. Because the study employs a one-group pretest–posttest design without a control group, the word *effect* is used in a limited sense, referring to the change in scores after the treatment, and not as proof of causation.

2. Method

This study used a quantitative approach with a pre-experimental one-group pretest–posttest design (O1 X O2), in which O1 is the pretest, X is the combined tahsin and talaqqi treatment, and O2 is the posttest (Rahmatika & Zainul, 2021; Waruwu, 2024). The design compares scores before and after treatment but, without a control group, cannot rule out other explanations for score changes. The study took place at the Technical Implementation Unit (UPT) of Ma'had Al-Jami'ah, Universitas Islam Negeri (UIN) Palopo, South Sulawesi, Indonesia. The pretest was held on 15 July 2026, the treatment ran from 17 July to 3 August 2026, and the posttest was held on 5 August 2026. The treatment spanned 18 calendar days and included 12 structured sessions, each lasting 90 minutes (18 hours of direct instructional contact). As detailed in Section 2.4, the pretest and posttest used different passages (Q.S. Al-Qiyamah and Q.S. Al-Fatihah), introducing a potential threat to internal validity related to material equivalence, which is addressed in the limitations.

The population comprised all 126 students of the Student Pesantren Program across four faculties. Because the study prioritized feasibility and the depth of paired oral observation over broad population generalization, the researchers used a 20% error margin due to resource constraints and the study's exploratory nature. The Slovin formula was applied with this margin, yielding a target sample size of $n = 126 / (1 + 126 \times 0.20^2) = 20.86 \approx 21$. The Slovin formula was used only to determine the target sample size (n = 21); participants were then purposively selected based on specific criteria—namely, enrollment in the program, completion of both pretest and posttest, full attendance during the treatment, and complete paired data. Although the target was identified before the intervention, the final analyzed sample was confirmed only after the treatment concluded, when attendance and pretest–posttest completeness were verified. Excluding students who did not complete the full treatment or lacked paired data introduces a potential selection bias; completers may differ systematically in motivation or initial ability from non-completers. Because data on non-completers were not collected, the direction of this bias cannot be estimated. The findings thus describe the changes among completers, not the effect on all enrollees.

Tahsin and talaqqi were delivered as a single integrated treatment across three sequential stages over 18 calendar days with 12 structured sessions of 90 minutes each. Instruction was delivered by two certified *muwajjih* (resident mentors) holding bachelor's degrees in Islamic education and internal tahsin certification. To ensure consistency, both instructors used a shared session protocol—specifying the modeling–imitation–correction sequence, target verses, and rubric descriptors—and stayed aligned through weekly coordination before the posttest. A designated observer monitored treatment fidelity using observation checklists completed during each session, documenting adherence to the protocol and the modeling–imitation–correction sequence. Each stage followed the same cycle: the instructor modeled the reading, students listened and imitated, recited before the instructor, received immediate correction, and repeated until accurate. Within this cycle, the two methods were nested rather than sequential: tahsin provided the content and quality criteria for correct recitation, while talaqqi provided the interactional mechanism of direct modeling, listening, and immediate individualized correction.

Initially, the treatment targeted *makharijul huruf*, with tahsin involving the instructor demonstrating articulation points and leading letter-by-letter drills. In contrast, talaqqi involved the student reciting letters and words before the instructor for immediate *makhraj* correction. Subsequently, the focus shifted to *tajwid*, where tahsin involved the instructor explaining and modeling rules (*nun sakinah*, *tanwin*, *mad*, *qalqalah*) in target verses. In contrast, talaqqi involved the

student reciting those verses as the instructor identified errors and prompted rule recall in real time. In the final phase, the treatment targeted fluency and *fashahah*; here, tahsin consisted of modeling paced reading and *waqf–ibtida'* patterns, while talaqqi involved the student reading whole verses with the instructor immediately correcting pauses, starts, and pacing. Across all stages, tahsin modeling preceded talaqqi recitation, and talaqqi correction fed back into tahsin drilling, forming a continuous corrective loop. Program documents and an interview with the *muwajjih* provided only contextual information and were not used to test the hypotheses.

Reading literacy was measured using an oral Qur'an reading test, in line with rubric-based assessment of Qur'an reading (Bahri & Nurmawati, 2023; Maita et al., 2026). Each student read a set passage aloud and was scored on four aspects using the operational rubric presented in Table 1. The sum of the four aspect scores (maximum 16) was converted to a 100-point scale using the formula: final score = (sum of scores / 16) × 100, and classified as very good (80–100), good (70–79), adequate (60–69), poor (50–59), or very poor (0–49) (Sugiyono, 2023).

Table 1. Operational Rubric for the Four Aspects of Qur'anic Reading Literacy (1–4 Scale)

Aspect	Score 1 (Poor)	Score 2 (Adequate)	Score 3 (Good)	Score 4 (Very Good)
Makharijul huruf	More than 5 letters misarticulated; errors distort meaning	3–5 letters misarticulated; noticeable errors	1–2 minor articulation errors; meaning intact	All letters articulated correctly according to *makhraj*
Tajwid	More than 5 rule violations (*nun sakinah*, *tanwin*, *mad*, *qalqalah*)	3–5 rule violations; inconsistent application	1–2 minor violations; rules mostly applied	All rules applied correctly and consistently
Fluency	Frequent hesitation (>5 stops), word repetition, or fragmented reading	3–5 hesitations; halting but comprehensible	1–2 brief hesitations; generally smooth	Smooth, steady reading without hesitation or repetition
Fashahah	Unclear pronunciation; more than 3 *waqf/ibtida'* errors	Some clarity issues; 2–3 *waqf/ibtida'* errors	Minor clarity issues; 1 *waqf/ibtida'* error	Clear pronunciation; correct *waqf* and *ibtida'* throughout

Content validity was assessed by one expert validator, Dr. Makmur, S.Pd.I., M.Pd.I., a senior lecturer in Qur'anic studies with over 10 years of experience teaching tajwid and serving as a Ma'had assessment coordinator. He evaluated 11 criteria across four domains: (a) alignment of indicators with the four measured aspects, (b) clarity of rubric descriptors, (c) appropriateness of passage difficulty, and (d) practicality of administration. The total score was 42/44 (95.45%), classified as very valid (76–100%; Susilawati et al., 2022), based on a document review and face-to-face consultation on 18 June 2026. A key limitation is that only one validator was involved; thus, content validity rests on a single expert's judgment and does not reflect validity across independent reviewers.

Regarding reliability, although two muwajjih delivered the instructional treatment, a single examiner (the first author) scored both the pretest and posttest. No independent second rater was involved in the scoring process. As a result, inter-rater reliability could not be calculated. This recognized limitation restricts the generalizability of scoring consistency; we retained and acknowledged it in the limitations section. To partially mitigate this, the examiner conducted a self-calibration review of the rubric descriptors and pilot-scored two recorded samples prior to the pretest to ensure internal consistency in applying the criteria. Future studies should include an independent second rater and report inter-rater agreement (Cohen, 2013).

The pretest used Q.S. Al-Qiyamah/75:1–15 and the posttest used Q.S. Al-Fatihah/1:1–7. These passages differ in length (15 vs. 7 verses) and familiarity; Al-Fatihah is recited 17 times daily during

prayers and is therefore far more familiar to students. This familiarity may have inflated posttest scores, independent of the treatment effect. The passages were selected because both contain a comparable range of tajwid rules and Al-Fatihah is the program's standard exit-assessment passage; however, no empirical difficulty index was established to demonstrate equivalence. Consequently, the findings are interpreted cautiously as a change in performance across two non-equivalent passages, not as a pure treatment effect.

Descriptive statistics (minimum, maximum, mean, and standard deviation) were computed first. Because $n < 50$, we applied the Shapiro–Wilk test to the gain scores (posttest − pretest) to select the test of difference, and to the pretest and posttest scores as a complement. Normal gain scores called for the paired-samples t-test; otherwise, we used the Wilcoxon signed-rank test at a significance level of .05 to test H0 (no significant difference between pretest and posttest scores) against H1 (a significant difference). Effect size was computed as $r = |Z| / \sqrt{N}$ and interpreted with Cohen's (1988) benchmarks: .10 small, .30 medium, .50 large. All analyses were conducted on the 21 students who completed both measurements. Given the non-equivalent test materials and the absence of a control group, the statistical results are interpreted as descriptive of within-group change rather than as proof of causal effect.

Several methodological limitations qualify the findings. The absence of a control group means that factors other than the treatment may have contributed to score changes. The treatment lasted only 2.5 weeks (17 July–3 August 2026); this short duration may limit learning depth and retention of gains, as no delayed posttest was administered. The use of non-equivalent passages (Al-Qiyamah vs. Al-Fatihah), with Al-Fatihah being far more familiar to students, represents a serious threat to internal validity that may have inflated posttest scores. Inter-rater reliability could not be calculated because only one examiner scored both measurements. The sample was small (n = 21), purposively selected, and sized with a 20% error margin, limiting generalizability.

3. Result

Descriptive Statistics

Table 2 summarizes the 21 students' scores. At pretest, scores ranged from 50.00 to 75.00 (M = 59.82, SD = 7.01), and the median (56.25) was lower than the mean, which shows that initial reading ability was below the optimal level and that a few higher-scoring students pulled the average up. At posttest, scores ranged from 75.00 to 93.75 (M = 82.44, SD = 6.13; median = 81.25). The mean score therefore rose by 22.62 points. The lowest posttest score (75.00) equaled the highest pretest score, and the standard deviation decreased from 7.01 to 6.13, so posttest scores were somewhat less dispersed.

Individual-level data visualization is not presented because the raw per-student scores were not retained in a disaggregated form; the analysis therefore relies on aggregate descriptive statistics and category distributions.

Table 2. Descriptive Statistics of Pretest, Posttest, and Gain Scores (N = 21)

Statistic	Pretest	Posttest	Gain
Minimum	50.00	75.00	18.75
Maximum	75.00	93.75	25.00
Mean	59.82	82.44	22.62
Median	56.25	81.25	25.00
Standard deviation	7.01	6.13	3.11

Note. Gain = posttest − pretest. Scores are on a 0–100 scale.

Aspect-level analysis**.** A disaggregated analysis of the four assessed aspects (*makharijul huruf*, *tajwid*, *fluency*, and *fashahah*) could not be performed because the original dataset retained only composite total scores, not aspect-level scores. This constitutes a recognized limitation: although the composite score improved substantially, the available data do not permit identification of which aspect contributed most to the overall gain. Consequently, this study does not examine the aspect-specific effects of the combined tahsin–talaqqi treatment. Future research should retain disaggregated aspect scores to enable such analysis.

Achievement Categories

Table 3 shows the shift in achievement categories. At pretest, 12 students (57.1%) were in the poor category, eight (38.1%) were adequate, one (4.8%) was good, and none were very good; 20 of the 21 students were therefore poor or adequate. At posttest, 16 students (76.2%) were very good, and five (23.8%) were good, and no student remained in the poor or adequate category. All 12 students who were initially in the poor category moved to the good or very good category.

Table 3. Distribution of Achievement Categories at Pretest and Posttest (N = 21)

Category	Score range	Pretest, *f* (%)	Posttest, *f* (%)
Very good	80–100	0 (0.0)	16 (76.2)
Good	70–79	1 (4.8)	5 (23.8)
Adequate	60–69	8 (38.1)	0 (0.0)
Poor	50–59	12 (57.1)	0 (0.0)
Very poor	0–49	0 (0.0)	0 (0.0)

Note. Ranges are inclusive of the lower bound and exclusive of the upper bound (e.g., 70.00–79.99). Percentages are of the 21 students and may not sum to exactly 100 because of rounding.

Gain Patterns

A notable feature of the data is that gains followed only two discrete patterns: 18.75 points (equivalent to three steps on the 16-point rubric) for eight students, and 25.00 points (four steps) for thirteen students. No intermediate gain values were observed. This uniformity is unlikely to reflect a naturally continuous distribution of improvement; rather, it is probably an artifact of the four-level rubric, which permits only whole-step increments (e.g., 1→2, 1→3, 1→4) and thus coarsens change measurement. Because each aspect is scored on a 1–4 scale, the smallest possible gain per aspect is 1 point (6.25 points on the 100-point scale), and the maximum is 3 points (18.75 points). The observed gains of 18.75 and 25.00 therefore represent three-step and four-step total increases across the four aspects, respectively. This coarseness limits the precision of the gain estimates and should be considered when interpreting the magnitude of improvement. Future studies should consider using a finer-grained rubric (e.g., 1–6 or 1–10) to capture gradual change.

Gains were also similar across faculties: mean gains were 21.88 points for Tarbiyah and Teacher Training (n= 6), 22.50 for Sharia (n = 5), 22.50 for Islamic Economics and Business (n = 5), and 23.75 for Ushuluddin, Adab and Dakwah (n = 5). These subgroups are too small for inferential comparison and are reported descriptively only.

Normality and Hypothesis Testing

The Shapiro–Wilk test indicated that the gain scores were not normally distributed ($W = .620$, $p < .001$), and neither were the pretest scores ($W = .893$, $p = .025$) or the posttest scores ($W = .844$, $p = .003$) (Table 4). Because the normality assumption was violated, the Wilcoxon signed-rank test was used to compare pretest and posttest scores, as it is based on ranks and does not require normally distributed differences. The test yielded $Z = -4.01$, $p < .001$, so H_0 was rejected, and H1 was accepted: reading scores after the treatment were significantly higher than before. All 21 pairs showed a

positive difference (sum of positive ranks = 231, mean rank = 11.00); there were no negative differences and no ties.

Table 4. Results of the Normality Test, Wilcoxon Signed-Rank Test, and Effect Size

Analysis	Statistic	*p*	Interpretation
Shapiro–Wilk, pretest	$W = .893$	.025	Not normal
Shapiro–Wilk, posttest	$W = .844$	.003	Not normal
Shapiro–Wilk, gain	$W = .620$	< .001	Not normal
Wilcoxon signed-rank	$Z = -4.01$	< .001	Significant difference
Effect size	$r = .88$	–	Large

Note. Significance level = .05; $N = 21$ pairs.

Effect Size

The effect size was $r = |-4.01| / \sqrt{21} = .88$, which is well above the .50 benchmark and is classified as large. The observed improvement following the treatment was therefore not only statistically significant but also large in magnitude: the mean score shifted from the poor category (59.82) to the very good category (82.44), a gain of about 22.6 points, and 16 of the 21 students reached the very good category. However, because the design lacked a control group and used non-equivalent passages, this improvement cannot be attributed solely to the treatment. The findings should therefore be read as a change in performance following the intervention, not as definitive proof of the method's effectiveness.

4. Discussion

Main Finding and Its Interpretation

The central finding is a large and consistent improvement in Quranic reading scores after the combined tahsin and talaqqi treatment. Every student improved, and the group shifted from mostly poor and adequate categories to good and very good ones. This confirms that reading scores were significantly higher after the treatment and that the magnitude of change was substantial. The pretest results also validate the problem that motivated the program: twenty of the 21 students began in the poor or adequate categories despite studying at an Islamic university, which is consistent with reports that Qur'an reading among UIN Palopo students is varied and often below the desired level (Abdillah, 2025). This baseline justifies a structured, corrective program and gives the observed gain clear practical meaning.

However, the magnitude of this improvement must be interpreted with caution. The posttest passage (Q.S. Al-Fatihah) is far more familiar to students than the pretest passage (Q.S. Al-Qiyamah) because it is recited in every unit of the ritual prayer. This familiarity may have inflated posttest scores independently of the treatment. Consequently, the large effect size cannot be taken as a pure estimate of the treatment's effect; part of the gain may reflect familiarity with the passage rather than improved reading ability.

Furthermore, the lower dispersion at posttest suggests that scores became more even, but it does not show that students became equal in ability. The uniformity of gains every student gained exactly 18.75 or 25.00 points may reflect steady progress. However, it may also reflect the coarseness of a four-level rubric applied by a single examiner. A finer rubric and additional raters would show whether students actually progressed at different rates.

Theoretical Explanation

This improvement aligns with the behaviorist view that skills are built through repeated stimulus–response practice and reinforcement (Darmayanti et al., 2024). The instructor's model recitation and corrections served as stimuli; students' imitation and repeated reading were the

responses; immediate correction reinforced accurate responses, while inaccurate ones were repaired through repetition. Higher posttest scores can thus be understood as a product of drill and habituation, core features of tahsin (Siagian & Zailani, 2021).

The findings also align with Vygotsky's social constructivism, in which students who could not yet read accurately received direct, structured help from a more capable instructor. This help can be seen as scaffolding within the zone of proximal development, gradually withdrawn as learners become able to read on their own (Tohari & Rahman, 2024; Vygotsky, 1978). In this view, the teacher acts as a facilitator and mediator rather than a mere transmitter of information (Masgumelar & Mustafa, 2021)and interaction through communication, direction, and reinforcement has likewise been described as central to learning at the Palopo institution (Sukirman & Makmur, 2021).

Crucially, tahsin and talaqqi are not merely two methods delivered side by side; they are functionally complementary. Tahsin supplies the content and quality criteria of correct recitation what the articulation, tajwid rules, fluency, and fashahah should sound like and provides the structured drills through which the learner's production is brought toward those criteria (Siagian & Zailani, 2021). Talaqqi supplies the interactional mechanism through which that refinement is delivered: the teacher models the recitation, the learner listens and imitates, then recites before the teacher and receives immediate, individualized correction (Muktafi & Umam, 2022).

This complementarity was visible throughout the intervention. When targeting makharijul huruf, tahsin provided the articulation drills and criteria for correct makhraj, while talaqqi provided the face-to-face correction that diagnosed each learner's specific errors. For tajwid, tahsin explained rules and modeled applications, while talaqqi let the instructor identify rule violations in the learner's recitation and prompt immediate recall. When addressing fluency and fashahah, tahsin modeled paced reading and waqf–ibtida' patterns, while talaqqi corrected the learner's pauses, starts, and pacing in real time.

In each phase, tahsin set the standard, and talaqqi enforced it through individualized corrective feedback, operating as a continuous corrective loop rather than as sequential phases. The three-stage design may also have contributed to the observed improvement: the progression from makharijul huruf to tajwid to fluency and fashahah follows a component-to-integrated sequence and aligns with the four aspects scored on the rubric. This alignment between what was taught and what was measured plausibly explains the magnitude of the gain. However, because the analysis used only the total score, the study cannot determine which aspect improved most, nor can it claim that all four aspects improved equally.

Taken together, the results suggest that behaviorist and constructivist accounts need not compete in explaining Qur'an reading instruction. A cycle of modeling, imitation, recitation before the instructor, immediate correction, and guided repetition appears to draw on both: repetition and reinforcement strengthen accurate responses. In contrast, the instructor's adaptive support lets students with different starting points progress. This is offered as a working proposition grounded in the present data, not as a tested theory.

Comparison with Previous Studies

The findings align with earlier studies, but the comparison requires critical qualification. Tarigan and Setiawan (2026) reported that a tahsin program improved reading quality among primary school pupils in Thailand, and Fauziah and Komariah (2023) found that talaqqi was associated with better tahsin ability among school students. Studies on teaching makharijul huruf and tajwid (Azifah et al., 2026) likewise indicate that systematic instruction with practice and correction improves reading accuracy, while related work has examined tahsin among madrasah ibtidaiyah students (Maulida et al., 2026) and talaqqi in Qur'an memorization (Sastera et al., 2026).

These studies differ from the present one in context (school vs. residential ma'had), participants (children vs. university students), design (qualitative or descriptive-correlational vs. quantitative pretest–posttest), and measurement (teacher ratings or interviews vs. rubric-based oral assessment with effect size). Because of these differences, we cannot directly compare the present

effect size with theirs. The consistent direction is informative, but it does not mean the magnitude of improvement is comparable.

The present study's contribution lies in combining the two methods, using university-level participants, and reporting a quantitative effect size feature earlier studies did not address. At the same time, the absence of a control group and the non-equivalent passages mean that the present effect size cannot be treated as a definitive estimate of the methods' effectiveness. The similar gains across faculties suggest that the combined approach worked for students from different academic backgrounds, which is relevant given the wide range of starting abilities reported in the program (Nurzykila, 2026). However, subgroups of five or six students allow only a tentative reading.

The Role of the Ma'had Context

The Ma'had Al-Jami'ah environment is a potential contextual factor that may have contributed to the observed improvement, but its independent effect was not measured in this study. The ma'had offers a structured residential setting in which Qur'an practice is part of daily religious life (Muldiah & Wulandari, 2025). Students live in proximity to instructors and peers, participate in communal prayers, attend religious lectures, and are immersed in an environment that encourages habitual Qur'an reading. This immersion may have reinforced the treatment by providing continuous practice opportunities and social reinforcement.

However, because the intensity of these activities was not measured or controlled, the ma'had environment should be described as a plausible contextual contributor, not as an empirically demonstrated cause. The observed improvement may reflect a combination of the treatment and the broader ma'had environment, and the two cannot be separated in the present design. Future studies should measure the frequency and intensity of ma'had activities outside the intervention to estimate their independent contribution.

Limitations

Several limitations qualify these conclusions. The absence of a control group is the most important, as factors other than the treatment may have contributed to the observed improvement. The intervention involved 12 sessions over 18 days, but students' independent practice outside these sessions was not measured; students who practiced more may have improved more, and this cannot be separated from the treatment effect.

The habitual practice of reading the Qur'an in the residential ma'had, where it is part of daily religious life (Muldiah & Wulandari, 2025)may also have reinforced the treatment, but its independent contribution was not measured. Participation in other religious activities, such as character education, religious lectures, and communal prayers, may likewise have influenced students' motivation and reading practice. Finally, familiarity with the test material remains a serious concern: the posttest used Q.S. Al-Fatihah, which is recited in every unit of the ritual prayer and is therefore likely more familiar to students than the pretest passage. This familiarity may have inflated posttest scores independently of the treatment. Because none of these factors were measured or controlled, the observed improvement cannot be attributed solely to the combined tahsin–talaqqi treatment.

Measurement also carries limits. The non-equivalence of the pretest and posttest passages is the most consequential measurement threat. Because no empirical difficulty index was established to demonstrate equivalence, the magnitude of the improvement cannot be attributed solely to the treatment; the findings should therefore be read as a change in performance across two non-equivalent passages, not as a pure treatment effect. Scores came from a single examiner, so inter-rater reliability is unknown. The sample was small, purposively selected, restricted to students with complete data, and sized with a generous 20% error margin; there was also no delayed posttest, so retention of gains is unknown.

The large effect size does not eliminate these threats to internal validity; it describes the magnitude of the difference within the studied sample, not its cause. The findings should therefore be interpreted as a strong association between the combined methods and improved reading literacy, not as proof that the methods caused the improvement. Because tahsin and talaqqi were delivered together, their separate contributions cannot be identified.

Contributions and Implications

The study contributes in three ways. It provides quantitative evidence, with an effect size, on the combined use of two widely practiced Qur'an-teaching methods among university students, a setting less examined than school-level instruction. It links the observed improvement to two complementary learning theories through a concrete teaching cycle, offered as a proposition to be tested. It reports the weaknesses of its design openly, so that later studies can target them directly.

Practically, Ma'had managers can adopt the staged sequence used here: articulation, then tajwid, then fluency and fashahah, with a consistent cycle of modeling, imitation, recitation, correction, and repetition. Instructors should keep correction immediate and individual, because students begin at very different levels. For assessment, programs would benefit from a rubric with more levels and from at least two examiners. Future studies should use a pretest–posttest control group or non-equivalent control group design (Sugiyono, 2023), counterbalanced parallel passages, multiple raters with agreement statistics, a delayed posttest, aspect-level analysis, and separate tahsin-only and talaqqi-only groups to isolate each method's contribution.

5. Conclusion

This study examined the change in Quranic reading literacy among 21 students at UIN Palopo after the combined application of tahsin and talaqqi at Ma'had Al-Jami'ah. In answer to the first research question, the results show a significant difference between pretest and posttest scores, with reading scores significantly higher after the treatment. In answer to the second research question, the effect size was large. The observed scores increased substantially, and all students improved. At the posttest, every student was in the good or very good category, including those who had started in the poor category. These results address both objectives: reading scores were significantly higher after the treatment, and the difference was large. However, the findings should be interpreted as an improvement following the implementation of the combined methods, not as definitive proof of a causal relationship.

The study's design limits the conclusions that can be drawn. The absence of a control group, the use of different passages at pretest and posttest, the reliance on a single examiner, and the small purposive sample mean that the results indicate a strong association between the combined tahsin–talaqqi treatment and improved reading literacy. They do not establish causation. The observed improvement may also reflect familiarity with the passage, practice intensity, the ma'had environment, or other religious activities that were not measured. Consequently, the findings should be read as changes within the 21 students studied in the specific context of Ma'had Al-Jami'ah UIN Palopo, not as evidence that the method is guaranteed to be effective for all students or all Islamic educational institutions.

Three recommendations follow from these findings. Ma'had Al-Jami'ah managers should continue the staged tahsin–talaqqi sequence and record each student's progress on the four rubric aspects to enable evaluation at the aspect level. Instructors should give immediate, individual correction, with extra time for students in the lowest initial categories. Researchers should replicate the study with a control group, a larger sample drawn with a more conservative error margin, parallel test passages of verified equivalent difficulty, at least two examiners with reported agreement statistics, and a delayed posttest. Future studies should also compare tahsin-only, talaqqi-only, and combined conditions, and examine changes across each specific aspect of Quranic reading proficiency to determine which aspect contributes most to overall improvement.

6. References

Abdillah, F. (2025). Efektivitas Program Pesantren Mahasiswa Upt Ma’had Al-Jami’ah dalam Meningkatkan Kemampuan Membaca Al-Qur’an Mahasiswa Institut Agama Islam Negeri Palopo [Diploma, IAIN Palopo]. https://repository.uinpalopo.ac.id/id/eprint/10825

Akil, A. A., & Syafi'i, A. (2025). Meningkatkan Kemampuan Membaca Al-Qur'an Menggunakan Metode Talaqqi pada Mahasiswa HI 1 di UIN Alauddin Makassar. Jurnal Ilmiah Widya Borneo, 8(2), 121–134. https://doi.org/10.56266/widyaborneo.v8i2.363

Azifah, A., Ulfa, M., & Mujahidah, A. S. (2026). Strengthening Tajweed and Makharijul Letters in Improvement Fashahah Reading for Students at TPQ Fastabiqul Khairat Nunukan Mosque. Salam International Journal of Islamic Education, 5(1), 36–48. https://doi.org/10.22219/sinjie.v5i1.42947

Bahri, S., & Nurmawati. (2023). Pengembangan Penilaian Literasi Baca Al-Qur'An (Libaqur) Di Sdn Muka Sungai Kuruk, Aceh Tamiang. Pendas : Jurnal Ilmiah Pendidikan Dasar, 8(1), 333–343. https://doi.org/10.23969/jp.v8i1.7128

Cohen, J. (2013). Statistical Power Analysis for the Behavioral Sciences. Routledge.

Darmayanti, R., Nuryami, Sukriyah, Y., Sahara, N., Suprayitno, K., Susetyarini, R. E., & Baiduri. (2024). Behaviorisme dalam Pendidikan: Pembelajaran Berbasis Stimulus-Respon. Penerbit Adab.

Fauziah, H., & Komariah, R. (2023). Pengaruh Penerapan Metode Talaqqi Terhadap Kemampuan Tahsin Siswa Pada Mata Pelajaran Baca Tulis Qur'an. MASAGI : Jurnal Pendidikan Agama Islam, 2(1), 274–278. https://doi.org/10.37968/masagi.v2i1.509

Fitri, A., Dahlan, M., Muhammadun, M., Usman, & D, M. A. (2024). Improving the Ability Read the Qur'an Through the Tahsin Method Based on the Learn Quran Tajwid Application for Students at MTs Putra DDI Mangkoso. International Journal of Health, Economics, and Social Sciences (IJHESS), 6(3), 648~656-648~656. https://doi.org/10.56338/ijhess.v6i3.4715

Ilahi, A. N., & Fatihah, A. D. A. (2026). Implementation of the Qiraati Method as an Effort to Improve Qur'an Reading Skills among Students of a Qur'anic Education Center. Quranic Sciences for Community Service, 1(1), 23–32. https://doi.org/10.69526/qscs.v1i1.460

Ismail, A., Nurrohim, A., Izzatunnisa, Saib, A., & Elbanna, M. (2023). Tahsin Learning Strategy And Method To Improve The Al-Qur'an Reading Quality For The Majlis Tafsir Al-Qur'an Community In Surakarta. Profetika: Jurnal Studi Islam, 24(02), 393–420. https://doi.org/10.23917/profetika.v24i02.2844

Maita, I., Nelvawita, & Saputra, M. N. (2026). Augmented Reality and Neural Network-Based Speech Recognition for Tahsin Digital AR. Jurnal Nasional Teknik Elektro Dan Teknologi Informasi, 15(3), 212–219. https://doi.org/10.22146/jnteti.v15i3.18758

Masgumelar, N. K., & Mustafa, P. S. (2021). Teori Belajar Konstruktivisme dan Implikasinya dalam Pendidikan dan Pembelajaran. GHAITSA : Islamic Education Journal, 2(1), 49–57. https://doi.org/10.62159/ghaitsa.v2i1.188

Maulida, U. S., Rahmi, N., & Hidayat, R. (2026). Internalization of Islamic Values Among Students Through Daily School Activities at Madrasah Ibtidaiyah Ash Shabirin. Integrated Education Journal, 3(2), 176–197. https://doi.org/10.67899/qwnv7k31

Muktafi, A., & Umam, K. (2022). Implementasi Metode Talaqqi dalam Menghafal Al-Qur'an di Pondok Pesantren. Dirasat: Jurnal Manajemen dan Pendidikan Islam, 8(2), 194–205. https://doi.org/10.26594/dirasat.v8i2.3070

Muldiah, S., & Wulandari, D. (2025). Sinergi Pendidikan Agama Islam Dan Manajemen Pendidikan Islam Dalam Pembentukan Karakter Mahasiswa Berbasis Nilai Islami. Aksioma Ad Diniyah : The Indonesian Journal Of Islamic Studies, 13(2). https://doi.org/10.55171/jad.v13i2.1488

Nurzykila, N. (2026). Evaluasi Program Ma'had Al-Jami'ah Terhadap Penguatan Literasi Al-Qur'an Dan Kecerdasan Emosional Mahasiswa Uin Palopo [Masters, Universitas Islam Negeri Palopo]. https://repository.uinpalopo.ac.id/id/eprint/12333/

Puntambekar, S. (2022). Distributed Scaffolding: Scaffolding Students in Classroom Environments. Educational Psychology Review, 34(1), 451–472. https://doi.org/10.1007/s10648-021-09636-3

Rahmatika, S., & Zainul, R. (2021). Pretest-Posttest Control Group Design of Discovery Learning-Based Content Learning System in The Materials of Atomic Structure and Periodic Elements System of Class X Vocational Schools To Improve High-Level Thinking Ability. EKSAKTA: Berkala Ilmiah Bidang MIPA, 22(4). https://doi.org/10.24036/eksakta/vol22-iss4/268

Sastera, A., Rahman, A., & Masudi, M. (2026). Persepsi Mahasantri Terhadap Penerapan Metode Talaqqi Dalam Proses Hafalan Tahsin Wal Tahfiz di Mahad Al-Jamiah IAIN Curup [Undergraduate, Institut Agama Islam Negeri Curup]. https://e-theses.iaincurup.ac.id/10596/

Siagian, E., & Zailani, Z. (2021). Metode Tahsin dalam Memperbaiki Bacaan Al-Quran. AR-RASYID : Jurnal Pendidikan Agama Islam, 1(2), 114–120. https://doi.org/10.30596/10.30596/arrasyid.v1i2.8743

Sugiyono. (2023). Metode Penelitian Kuantitatif, Kualitatif, Dan R & D. ALFABETA, cv.

Sukirman, S., & Makmur, M. (2021). The contextualization of educational values: Symbolic forms in the interaction of learning process in IAIN Palopo. Bahasa Dan Seni: Jurnal Bahasa, Sastra, Seni, Dan Pengajarannya, 49(1), 108–121. https://doi.org/10.17977/um015v49i12021p108

Susilawati, S., Doyan, A., & Muliyadi, L. (2022). Effectiveness of Guided Inquiry Learning Tools to Improve Understanding Concepts of Students on Momentum and Impulse Materials. Jurnal Penelitian Pendidikan IPA, 8(3), 1548–1552. https://doi.org/10.29303/jppipa.v8i3.1919

Tambusai, J., Muchtar, M., & Wiguna, S. (2023). Pengaruh Metode Tahsin Terhadap Kemampuan Membaca Al Qur'an Siswa Kelas VIII MTS AL-Ikhwan Kecamatan Padang Tualang Langkat. Journal of Student Research, 1(1), 340–349. https://doi.org/10.55606/jsr.v1i1.1035

Taringan, A. A. (2026). Penerapan Metode Talaqqi Sama' Dalam Meningkatkan Kemampuan Membaca Alquran Pada Ma'had Tahfidz Alabdia Thailand. Pendas : Jurnal Ilmiah Pendidikan Dasar, 11(01), 82–90. https://doi.org/10.23969/jp.v11i01.41865

Tohari, B., & Rahman, A. (2024). Konstruktivisme Lev Semonovich Vygotsky dan Jerome Bruner: Model Pembelajaran Aktif dalam Pengembangan Kemampuan Kognitif Anak. Nusantara: Jurnal Pendidikan Indonesia, 4(1), 209–228. https://doi.org/10.14421/njpi.2024.v4i1-13

Vygotsky. (1978). Educational Psychology.

Waruwu, M. (2024). Metode Penelitian dan Pengembangan (R&D): Konsep, Jenis, Tahapan dan Kelebihan. Jurnal Ilmiah Profesi Pendidikan, 9(2), 1220–1230. https://doi.org/10.29303/jipp.v9i2.2141