



The Effectiveness of the EduReligious English Model in Improving English Teaching Competence among Raudhatul Athfal Teachers

Pangulu Abdul Karim¹, Ernita Daulay², Eka Yusnaldi³

^{1,2,3}Universitas Islam Negeri Sumatera Utara Medan

Article Info

Article history:

Received 05-08-2026

Revised 21-08-2026

Accepted 24-08-2026

Keywords:

EduReligious English;
effectiveness; English
teaching competence;
Raudhatul Athfal teachers;
quasi-experimental design

ABSTRACT

This study aimed to examine the effectiveness of the EduReligious English Model in improving the English teaching competence of Raudhatul Athfal (RA) teachers. A quantitative approach was employed using a nonequivalent pretest–posttest control group quasi-experimental design. The study involved 30 RA teachers affiliated with the Islamic Kindergarten Teachers Association (IGRA) of Deli Serdang Regency, with 15 teachers assigned to the experimental group and 15 to the control group. The experimental group participated in an integrated training and coaching program combining developmentally appropriate pedagogy, functional English instruction, and Islamic values, while the control group participated in regular professional development activities. Data were collected through performance-based assessments and analyzed using descriptive statistics, independent- and paired-samples t-tests, normalized gain, and analysis of covariance (ANCOVA). The experimental group showed a greater improvement in teaching competence than the control group, with mean scores increasing from 62.20 to 79.33 compared with an increase from 62.67 to 69.40 in the control group. After controlling for pretest scores, the ANCOVA indicated a significant group effect, $F(1, 27) = 48.17$, $p < .001$, partial $\eta^2 = .641$, with an adjusted mean difference of 10.35 points. The posttest Hedges' g of 1.45 indicated a large effect. The experimental group also achieved a higher normalized gain than the control group (0.463 vs. 0.181). These findings provide evidence of the effectiveness of the EduReligious English Model in improving English teaching competence among RA teachers within the study sample. The model offers a potentially useful professional development approach for integrating English language instruction, developmentally appropriate pedagogy, and Islamic values in Islamic early childhood education.

This is an open access article under the CC BY-SA license.



Corresponding Author:

Pangulu Abdul Karim

Dosen Pembelajaran Matematika di PGMI Universitas Islam Negeri Sumatera Utara Medan, Jl. William Iskandar No V Medan Estate

Email: panguluabdulkarim@uinsu.ac.id

INTRODUCTION

Raudhatul Athfal (RA) is an Islamic early childhood education institution that is expected to support children's development while simultaneously cultivating religious values and moral character. This dual function makes English teaching in RA different from English instruction at higher levels of education. English cannot simply be introduced as a collection of vocabulary and language structures. It needs to be presented through experiences that are meaningful, enjoyable, and appropriate for young children's developmental characteristics.

Play, movement, songs, storytelling, repetition, imitation, and interaction are therefore important parts of the learning process. In this setting, teachers play a central role in determining whether English becomes a meaningful classroom experience or merely an additional subject.

The need for appropriate English instruction has become increasingly relevant as young children are exposed to digital media, technology, and forms of global communication at an earlier age. Yet greater exposure to English does not automatically lead to meaningful learning. For young learners, the primary concern is not mastery of formal grammar. More fundamental outcomes include phonological awareness, simple vocabulary, willingness to respond in English, and positive experiences when using the language. Pinter (2017) emphasizes that young language learners benefit from meaningful contexts, active participation, repeated exposure, and appropriate instructional support. Garton and Copland (2019) similarly point out that teaching English to young learners needs to consider their age, interests, identities, cultural backgrounds, and learning environments. These principles are consistent with developmentally appropriate practice in early childhood education, which requires teachers to adapt instruction to children's developmental characteristics and lived experiences (Friedman et al., 2022; Alanís & Sturdivant, 2023).

The challenge, therefore, is not simply whether English should be introduced in RA, but whether teachers are adequately prepared to teach it in ways that young children can understand and enjoy. RA teachers need to use classroom English clearly, select appropriate vocabulary, pronounce words intelligibly, provide simple instructions, and create activities that encourage children to participate. They also need to manage the classroom, use suitable learning media, assess children's development, and respond to differences among learners. These demands require more than knowledge of English itself. They require an integration of language competence and pedagogical competence.

Research on early childhood education indicates that this integration cannot always be assumed. Rakhmania et al. (2023) reported that teachers' pedagogical competence is influenced by educational background and professional experience. Witarsa and Alim (2022) likewise emphasized subject-matter mastery, instructional planning, classroom implementation, and continuous professional development as important aspects of professional competence. The implication is straightforward: teachers may understand the general principles of early childhood education while still experiencing difficulties when those principles have to be translated into actual classroom practice.

This gap between knowing and doing becomes particularly important in English teaching. Harmanto et al. (2021) found that early childhood teachers identified instructional media and materials as major needs in developing their ability to teach English, followed by teaching techniques and other professional needs. Teachers consequently require opportunities to practice rather than merely receive information. They need to experience how a classroom expression is used, how a story is delivered, how a song or movement activity is managed, and how children respond to different forms of instruction. Feedback is equally important because teachers cannot always identify weaknesses in their own classroom performance.

Professional development, therefore, needs to move beyond the traditional model of short-term training. Silalahi and Sahara (2022) identified training, mentoring, workshops, lesson study, formal education, and experience-based learning as strategies for developing the

professional competence of early childhood educators. Ulfah et al. (2023) also found that tiered professional development can contribute to improvements in the professional competence of kindergarten and RA teachers. These findings suggest that professional learning becomes more meaningful when teachers are given opportunities to practice, receive feedback, and connect new knowledge with the realities of their classrooms.

There is another issue that cannot be separated from the RA context: Islamic values. English instruction in RA should not make religious identity an additional or unrelated component of learning. Instead, values can be embedded naturally in classroom activities. A lesson on My Body, for example, can be connected with gratitude and personal hygiene. A lesson about animals can encourage compassion toward living creatures. Expressions such as please, thank you, sorry, and may I help you? can simultaneously develop English and reinforce politeness, empathy, and prosocial behaviour. Abdullah et al. (2024) provide conceptual support for connecting religious values with active learning experiences.

This means that an effective professional development program for RA teachers has to address several needs at once. Teachers need English that is functional for classroom interaction. They need pedagogical strategies that are appropriate for young children. They also need to understand how Islamic values can be integrated into learning without making English instruction rigid or overly formal.

Technology adds another layer to this challenge (Rohman et al., 2023). Digital resources can support teachers in preparing instructional materials, providing audiovisual examples, documenting children's learning, and developing portfolios (Jamin & Rohman, 2026). However, technology itself does not guarantee effective learning. Sutiyono et al. (2022) found that early childhood educators generally showed positive attitudes toward technology integration but still faced difficulties in selecting appropriate digital tools and maintaining children's attention. Marsegi et al. (2023) similarly reported differences in teachers' competence in digital portfolio assessment. At the same time, Sukini et al. (2024) found positive effects of blended learning supported by digital platforms on language skills among early childhood education teachers. Taken together, these findings suggest that technology is most useful when it strengthens not replaces meaningful interaction between teachers and children.

The same principle applies to active learning. Young children learn through doing, responding, and interacting. Role-play, storytelling, picture books, songs, games, and movement activities can create opportunities for language use while also supporting social and character development (Sani et al., 2024; Harefa et al., 2025). Classroom management also matters because even a well-designed activity may fail when it does not match children's attention span, movement needs, or social characteristics. Kusumaningtyas and Aprianto (2025) identified play-based learning, flexible classroom organization, effective teacher–child communication, and consistent routines as important elements of classroom management in early childhood education.

Teachers' social-emotional competence further shapes this process. Young children are more willing to experiment with a new language when they feel safe, encouraged, and accepted (Natasya et al., 2026). Teachers need to respond to errors without creating anxiety and to build interactions that encourage participation. Oktarina and Kurniawati (2025) found associations between teachers' social-emotional competence and teacher–student relationships, classroom

management, and social-emotional learning. These findings reinforce an important point: English teaching competence in RA should not be reduced to teachers' knowledge of English. It is reflected in what teachers actually do in the classroom.

Against this background, the EduReligious English Model was developed as an integrated professional development program for RA teachers. The model brings together three areas that are often treated separately: developmentally appropriate early childhood pedagogy, functional English instruction, and Islamic values. Its implementation consists of needs assessment, language and values mapping, facilitator modelling, microteaching, classroom implementation, reflection, and individualized coaching. The underlying idea is practical. Teachers are not expected to master a new approach simply by listening to explanations. They are given opportunities to observe, practice, implement, reflect, and improve.

The model is therefore positioned as a response to the gap between professional knowledge and classroom performance. Previous studies have examined early childhood teacher competence, English teaching needs, professional development, technology integration, active learning, and religiously oriented learning. However, these aspects have generally been investigated separately. Limited empirical evidence is available on an integrated professional development model that simultaneously addresses functional English instruction, developmentally appropriate pedagogy, and Islamic value integration for RA teachers. More importantly, there is limited evidence regarding whether participation in such an integrated model can produce greater improvement in teachers' actual English teaching competence than regular professional development.

This gap is important because improvement cannot be established simply by asking teachers whether they feel more confident or whether they understand the training materials. A teacher may understand play-based learning in theory but still struggle to implement it. Likewise, a teacher may know many English words but have difficulty giving simple classroom instructions or responding naturally to children's communication. For this reason, the present study focuses on performance-based English teaching competence as the primary outcome.

The novelty of this study lies primarily in the substantive integration of functional English instruction, developmentally appropriate pedagogy, and Islamic value integration within a single professional development model for RA teachers. Rather than examining these elements independently, the study investigates them as interconnected components of teacher development. The study also examines whether teachers who participate in the EduReligious English Model show greater improvement in performance-based teaching competence than teachers who participate in regular professional development. The quasi-experimental design and statistical controls are used to strengthen the empirical examination of this substantive proposition, rather than being claimed as the novelty itself.

Accordingly, this study aims to: (1) describe the English teaching competence of RA teachers before and after participation in the EduReligious English Model; (2) examine differences in post-intervention teaching competence between the experimental and control groups after controlling for pretest scores; (3) determine the magnitude of the intervention effect; and (4) evaluate the implementation fidelity of the program. The findings are expected to provide empirical evidence that can inform RA institutions, professional teacher organizations, and teacher training providers in designing professional development that is

responsive to the linguistic, pedagogical, and religious characteristics of Islamic early childhood education.

METHOD

Research Approach and Design

This study employed a quantitative approach using a nonequivalent pretest–posttest control group quasi-experimental design. This design was selected because teachers were assigned to the research groups without individual randomization. Both groups completed pretest and posttest assessments. The experimental group received the EduReligious English Model, whereas the control group participated in the institution's regular professional development program. The inclusion of a pretest enabled the researchers to examine the baseline equivalence of the two groups while controlling for initial differences through analysis of covariance (ANCOVA).

Table 1. Quasi-Experimental Research Design

Group	Pretest	Intervention	Posttest
Experimental	O ₁	X: EduReligious English Model	O ₂
Control	O ₃	C: Regular professional development (no intervention)	O ₄

Research Setting, Population, and Sample

The study was conducted at the IGRA of Deli Serdang Regency over a six-week period during the second semester of the 2025/2026 academic year. The target population comprised Raudhatul Athfal (RA) teachers responsible for teaching children aged 5–6 years who were actively involved in introductory English instruction.

The study sample consisted of 30 teachers who met the predetermined inclusion criteria: actively teaching, willing to participate throughout the intervention, and possessing complete pretest and posttest data. Of these, 15 teachers were assigned to the experimental group and 15 to the control group.

A balanced sample composition was maintained to ensure that between-group comparisons were not influenced by unequal group sizes. All 30 participants completed both the pretest and posttest, resulting in a complete dataset with no participant attrition. The unit of analysis was the individual teacher. Given the relatively small sample size, statistical reporting extended beyond *p*-values to include confidence intervals, partial eta squared, and Hedges' *g* as measures of effect size.

Purposive sampling was employed to select participants whose characteristics aligned with the objectives of the professional development program and who completed all required activities. Group assignment followed the existing teacher activity groups rather than individual randomization, thereby producing a nonequivalent group design. Potential baseline imbalances were addressed through two procedures: comparing pretest scores using an independent-samples *t*-test and incorporating pretest scores as a covariate in the ANCOVA.

Research Data and Quality Control

The research dataset comprised 30 participant records and eight variables: teacher identification code, group code, group label, pretest score, posttest score, gain score, normalized gain (N-gain), and N-gain category. Participant identities were anonymized using the codes K01–K15 for the control group and E01–E15 for the experimental group.

Preliminary data screening confirmed the absence of duplicate participant codes and missing values. Gain scores were recalculated as the difference between posttest and pretest scores, whereas normalized gain (N-gain) was computed using the formula $(posttest - pretest) / (100 - pretest)$. The recalculated values were fully consistent with those recorded in the original dataset. No participants or data points were excluded from the statistical analyses.

Variables and Operational Definitions

The independent variable was the implementation of the EduReligious English Model. This model is operationally defined as a structured professional development program comprising needs assessment, English language and Islamic values mapping, instructional modelling, microteaching practice, classroom implementation in Raudhatul Athfal (RA) settings, and reflective coaching. The intervention was considered successfully implemented when participants attended the minimum required number of sessions, completed the instructional materials, conducted microteaching sessions, received facilitator feedback, and submitted evidence of classroom implementation. Implementation fidelity was evaluated using facilitator checklists and participant activity logs.

The dependent variable was the English teaching competence of Raudhatul Athfal (RA) teachers, defined as their ability to plan, implement, and evaluate communicative English instruction that is developmentally appropriate for young children while integrating Islamic values. Teaching competence encompassed five dimensions: functional English language competence, early childhood pedagogical competence, instructional material, media, and assessment development, integration of religious values, and social-emotional interaction and reflective practice. The primary competency score was derived from classroom performance observations rather than self-reported perceptions.

Table 2. Operationalization of English Teaching Competence among Raudhatul Athfal (RA) Teachers

Dimension	Key Indicators	Score Source
Language Competence	Classroom English, age-appropriate vocabulary, pronunciation clarity, fluency, and the ability to simplify classroom instructions.	Functional performance and microteaching observation
Pedagogical Competence	Developmentally appropriate learning objectives, play-based instruction, scaffolding, classroom management, differentiated instruction, and authentic assessment.	Instructional materials rubric and classroom observation

Instructional Materials and Media	Relevance of instructional themes, quality of stories, songs, and language games, use of concrete and digital learning media, and safe instructional practices.	Product evaluation and implementation portfolio
Religious Integration	Alignment between language instruction and Islamic values, accuracy of Islamic messages, teacher role modelling, and integration of values into classroom practices.	Expert evaluation rubric and classroom observation
Social-Emotional Competence	Warm teacher–child interactions, emotional support, responses to children's errors, child participation, and teacher reflective practice.	Classroom observation and reflective journals

EduReligious English Model Intervention

The intervention was structured into eight core training sessions followed by one classroom implementation cycle. The first session introduced the program, conducted a needs assessment, and familiarized participants with the principles of developmentally appropriate practice. The second session focused on classroom English expressions and pronunciation. The third emphasized songs, movement-based activities, and Total Physical Response (TPR). The fourth explored storytelling techniques and picture-book interaction. The fifth concentrated on role-play and language games, while the sixth addressed the integration of English language instruction with Islamic values. The seventh session focused on instructional media and digital portfolio development. The eighth and final training session was devoted to microteaching, structured feedback, and revision of instructional materials.

Following the core sessions, teachers in the experimental group implemented one EduReligious English module in their respective Raudhatul Athfal (RA) classrooms. Classroom implementation was observed by facilitators or trained observers using a standardized observation rubric. Participants subsequently attended coaching sessions to reflect on language clarity, children's responses, time management, the appropriateness of Islamic value integration, and strategies for improving instructional activities. Each session lasted 120 minutes, and the entire intervention was completed over a six-week period. All instructional materials, learning activities, facilitator roles, and implementation schedules were documented in an intervention manual to ensure replicability and facilitate the assessment of implementation fidelity.

Teachers in the control group participated in eight regular professional development sessions, each lasting 120 minutes. The training covered introductory RA curriculum content, English vocabulary development, instructional administration, and general educational discussions. Unlike the experimental condition, the control program did not include structured microteaching, classroom implementation observations, or individualized coaching. Equal instructional time was maintained across both groups to minimize the possibility that observed differences resulted solely from variations in training duration. Upon completion of the posttest, participants in the control group were offered the EduReligious English module and coaching sessions as an ethical measure to ensure equitable access to the intervention.

Research Instruments

The primary instrument was the RA Teachers' English Teaching Competency Rubric. The rubric comprised 30 performance indicators evaluated using a four-level performance scale: *not demonstrated*, *emerging*, *developing*, and *consistent/effective*. Each performance level was accompanied by observable behavioral descriptors to enhance scoring objectivity and minimize subjective judgment. Two independent observers evaluated either recorded or live microteaching sessions. Dimension scores were aggregated to generate an overall competency score after scoring direction and weighting had been predetermined prior to data analysis.

The second instrument was a functional English performance assessment. This assessment measured teachers' ability to comprehend classroom instructions, deliver instructional directions, read short stories aloud, pronounce thematic vocabulary accurately, and respond appropriately to classroom situations. Performance was evaluated using criteria related to clarity, accuracy, fluency, vocabulary appropriateness, and the ability to simplify language for young learners. The assessment was specifically designed to measure the practical English proficiency required for Raudhatul Athfal (RA) teachers rather than general academic language competence.

The third instrument was the EduReligious English instructional materials rubric. The rubric evaluated the alignment between language objectives and child development outcomes, the relevance of Islamic values, the appropriateness of play-based learning activities, the quality of instructional media, lesson sequencing, authentic assessment practices, inclusivity, and classroom safety. Additional instruments included an implementation fidelity checklist, reflective journals, participant background forms, and attendance records. Although a teacher self-efficacy questionnaire could be administered as a secondary outcome measure, it did not replace performance-based assessments.

All raw scores were transformed to a standardized scale ranging from 0 to 100. The composite teaching competency score was calculated by weighting microteaching observations (40%), the functional English performance assessment (25%), the instructional materials rubric (25%), and the classroom implementation portfolio (10%). These weightings were established before statistical analysis to prevent post hoc adjustments aimed at inflating intervention effects. Classroom observation received the greatest weight because the dependent variable was operationally defined as teaching performance. Dimension-specific analyses were reported as secondary outcomes and did not replace the composite competency score as the primary outcome measure.

Instrument Validity and Reliability

Content validity was established through expert evaluation involving specialists in English language education, early childhood education, and Islamic education. Expert reviewers assessed the relevance of each indicator, the clarity of the performance descriptors, developmental appropriateness, and the adequacy of Islamic value integration. The primary dataset used for statistical analysis contained only composite scores rather than item-level responses. Consequently, item-level validity coefficients were not recalculated using data from the 30 principal participants and should instead be reported based on the findings of the separate instrument validation study.

The instruments were piloted with teachers who shared characteristics similar to those of the target participants but were not included in the main study sample. For the observational rubric, reliability assessment ideally included inter-rater agreement, whereas instruments consisting of multiple indicators were evaluated through internal consistency analysis. Reliability analyses were not repeated using the primary dataset because the available data file did not contain item-level responses or separate scores from individual raters. This limitation is explicitly acknowledged to avoid reporting unsupported instrument reliability coefficients.

Data Collection Procedures

The study commenced with institutional approval, participant recruitment, the acquisition of written informed consent, and the collection of baseline participant characteristics. Pretests were administered to both groups within the same time frame. Whenever feasible, performance recordings were anonymized through coded identifiers so that observers remained blind to participants' group assignments. Following the pretest, the experimental group received the EduReligious English intervention, whereas the control group participated in the comparison condition. Implementation fidelity was documented throughout every training session. Posttests were administered immediately after completion of the intervention using identical procedures, duration, instruments, and testing conditions for both groups.

Data Analysis

Data were analyzed using IBM SPSS Statistics or equivalent statistical software. Preliminary analyses included verification of participant identification codes, screening for missing values, examination of score ranges, and confirmation of the consistency of derived variables. Descriptive statistics were reported as sample size, mean, standard deviation, minimum, and maximum values. Within-group changes were analyzed using paired-samples *t*-tests. Baseline equivalence, posttest differences, gain scores, and normalized gain (N-gain) scores were examined using independent-samples *t*-tests following Levene's test for homogeneity of variances.

The primary analysis employed analysis of covariance (ANCOVA), with posttest scores specified as the dependent variable, group as the fixed factor, and pretest scores as the covariate. Assumption testing included the independence of observations, linearity of the relationship between pretest and posttest scores, normality of residuals, homogeneity of variances, the absence of influential outliers, and homogeneity of regression slopes. The interaction between group membership and pretest scores was examined to assess the homogeneity of regression slopes assumption. If participants were clustered within educational institutions and the intraclass correlation coefficient (ICC) was non-negligible, multilevel modelling or cluster-adjusted standard errors would be more appropriate than conventional ANCOVA.

Statistical significance was established at $\alpha = 0.05$. ANCOVA results were reported using *F* values, degrees of freedom, *p*-values, partial eta squared, adjusted means, adjusted mean differences, and 95% confidence intervals. Hedges' *g* was calculated to estimate the standardized between-group posttest effect size with correction for small sample bias. Confidence intervals for Hedges' *g* were obtained through 10,000 bootstrap replications using

a fixed random seed. Normalized gain (N-gain) scores were categorized as low ($g < 0.30$), moderate ($0.30 \leq g < 0.70$), and high ($g \geq 0.70$).

FINDINGS

Participant Characteristics and Data Flow

The analysis included 30 teachers, equally distributed between the experimental group ($n = 15$) and the control group ($n = 15$). All participants completed both the pretest and posttest assessments, resulting in a complete dataset with no participant exclusions. Each participant was assigned a unique identification code. Pretest scores ranged from 50 to 73, whereas posttest scores ranged from 61 to 89. Data completeness by study group is presented in Table 3.

Table 3. Distribution of Participants and Data Completeness by Study Group

Characteristic	Experimental Group	Control Group	Total
Number of teachers	15	15	30
Participant codes	E01–E15	K01–K15	30 unique codes
Complete pretest data	15 (100%)	15 (100%)	30 (100%)
Complete posttest data	15 (100%)	15 (100%)	30 (100%)
Cases included in the analysis	15	15	30

Data Integrity Assessment

Table 4. Summary of Dataset Integrity Assessment

Component	Assessment Criteria	Result	Decision
Teacher identification codes	No missing or duplicate codes	30 unique codes	Meets the criterion
Group assignment	0 = Control; 1 = Experimental	15 : 15	Balanced
Gain score	Posttest – Pretest	30/30 verified	Consistent
Normalized gain (N-gain)	(Posttest – Pretest) / (100 – Pretest)	30/30 verified	Consistent

Descriptive Statistics of Teaching Competence

Table 5. Descriptive Statistics of Teaching Competence Scores

Group	Pretest (M ± SD)	Posttest (M ± SD)	Mean Change (95% CI)	N-gain (M ± SD)
Experimental	62.20 ± 6.46	79.33 ± 7.58	17.13 [14.55, 19.72]	0.463 ± 0.135
Control	62.67 ± 5.56	69.40 ± 5.63	6.73 [4.88, 8.59]	0.181 ± 0.094

The mean teaching competency score of the experimental group increased from 62.20 ($SD = 6.46$) at pretest to 79.33 ($SD = 7.58$) at posttest. This represented a mean improvement of 17.13 points (95% CI [14.55, 19.72]). In contrast, the control group improved from 62.67 ($SD = 5.56$) to 69.40 ($SD = 5.63$), corresponding to a mean increase of 6.73 points (95% CI [4.88, 8.59]).

No statistically significant difference was observed between the two groups at baseline, $t(28) = 0.21$, $p = .834$, indicating comparable initial teaching competency. Following the

intervention, however, the experimental group achieved significantly higher posttest scores than the control group, $t(28) = 4.08, p < .001$.

Assumption Testing

Table 6. Summary of ANCOVA Assumption Testing

Assumption	Test/Indicator	Result	Decision
Normality of residuals	Shapiro–Wilk test	$W = 0.988; p = .974$	Assumption satisfied
Homogeneity of variances	Levene's test (posttest)	$F = 2.415; p = .131$	Assumption satisfied
Linearity	Quadratic component of pretest	$F = 1.316; p = .262$	Assumption satisfied
Homogeneity of regression slopes	Group $\times$ Pretest interaction	$F = 0.131; p = .721$	Assumption satisfied
Influential outliers	Studentized residuals / Cook's D	$ r = 2.253; D = 0.141$	No influential outliers detected

The residuals from the ANCOVA model were normally distributed, $W = 0.988, p = .974$. Levene's test indicated homogeneity of variances for posttest scores, $F(1, 28) = 2.42, p = .131$, confirming that the assumption of equal variances was satisfied. The quadratic component of the pretest was not statistically significant, $F(1, 26) = 1.32, p = .262$, supporting the assumption of linearity. Likewise, the interaction between group and pretest scores was not significant, $F(1, 26) = 0.13, p = .721$, indicating that the assumption of homogeneity of regression slopes was met. In addition, the Breusch–Pagan test did not detect heteroscedasticity, $F(2, 27) = 1.19, p = .318$. The maximum absolute studentized residual (2.25) and Cook's D value (0.141) further indicated that no individual case exerted undue influence on the fitted model.

Hypothesis Testing and Effect Size

Table 7. ANCOVA Results for English Teaching Competence

Source	SS	df	MS	F	p	Partial η^2
Pretest	797.656	1	797.656	47.936	$< .001$	0.640
Group	801.585	1	801.585	48.172	$< .001$	0.641
Error	449.277	27	16.640			

Table 8. Adjusted Means and Effect Size

Group	Adjusted Mean	SE	95% CI	Adjusted Mean Difference	Hedges' g (95% CI)
Experimental	79.54	1.05	[77.38, 81.70]	10.35	1.45 [0.73, 2.69]
Control	69.19	1.05	[67.03, 71.36]	-	-

After controlling for pretest scores, a statistically significant difference in posttest scores was observed between the experimental and control groups, $F(1, 27) = 48.17, p < .001$, partial $\eta^2 = .641$. The adjusted mean posttest score for the experimental group was 79.54 (SE

= 1.05, 95% CI [77.38, 81.70]), whereas the control group obtained an adjusted mean of 69.19 ($SE = 1.05$, 95% CI [67.03, 71.36]). The adjusted mean difference between the two groups was 10.35 points (95% CI [7.29, 13.41]). The ANCOVA model, including both pretest scores and group membership, explained 77.4% of the variance in posttest scores ($R^2 = .774$). Pretest performance also emerged as a significant predictor of posttest scores, $F(1, 27) = 47.94$, $p < .001$.

The between-group posttest difference yielded a Hedges' g of 1.45 with a 95% bootstrap confidence interval of [0.73, 2.69], indicating a large effect size. Consistent findings were obtained from the gain score analysis: the experimental group demonstrated a mean gain that was 10.40 points greater than that of the control group, $t(28) = 7.01$, $p < .001$, Hedges' $g = 2.49$. Although the confidence intervals surrounding the effect sizes were relatively wide because of the modest sample size (15 teachers per group), the lower bound of the confidence interval for the posttest Hedges' g still indicated a practically meaningful intervention effect.

N-Gain Distribution and Consistency of Improvement

The experimental group achieved a mean normalized gain (N-gain) of 0.463 ($SD = 0.135$), which falls within the moderate category, whereas the control group obtained a mean N-gain of 0.181 ($SD = 0.094$), corresponding to the low category. The between-group difference in N-gain (0.282) was statistically significant, $t(28) = 6.65$, $p < .001$.

Among the teachers in the experimental group, 12 of the 15 participants achieved a moderate N-gain, while the remaining three teachers were classified in the low category. In contrast, 13 teachers in the control group fell within the low category, and only two teachers achieved a moderate N-gain. All participants demonstrated positive gain scores, indicating consistent improvement across both groups; however, the magnitude of improvement was substantially greater among teachers who participated in the EduReligious English Model intervention.

The findings demonstrate that the EduReligious English Model produced significantly higher English teaching competency scores than the regular professional development program after controlling for participants' baseline performance. The adjusted mean difference of 10.35 points was not only statistically significant but also educationally meaningful. Furthermore, the large effect size, reflected by a partial η^2 of .641, indicates that group membership accounted for a substantial proportion of the variance in posttest scores after adjustment for pretest differences. This interpretation is reinforced by the Hedges' g value of 1.45, suggesting that the superiority of the intervention cannot be attributed merely to sample size or statistical significance testing.

The descriptive results further support this conclusion. Teachers in the experimental group improved by an average of 17.13 points, whereas those in the control group improved by 6.73 points. The improvement observed in the control group suggests that routine professional development, accumulated teaching experience, and repeated testing may contribute to enhanced teaching competence over time. Nevertheless, the 10.40-point difference in gain scores, together with the significantly higher normalized gain (N-gain) achieved by the experimental group, indicates that the EduReligious English Model provided benefits beyond those associated with conventional professional development.

The practical significance of these findings should be interpreted in relation to the construct measured. The composite competency score represented teachers' performance across multiple domains, including functional English proficiency, pedagogical competence, instructional planning, the integration of Islamic values, and classroom implementation. The adjusted mean score of 79.54 obtained by the experimental group, compared with 69.19 for the control group, reflects a substantial difference in overall teaching quality. However, because the available dataset contained only composite scores, it was not possible to determine which competency dimension contributed most strongly to the observed improvement.

The effectiveness of the EduReligious English Model may be explained by its close alignment with teachers' professional needs. Harmanto et al. (2021) identified instructional media, teaching materials, and classroom techniques as priority areas for teachers' professional development. The present intervention directly addressed these needs through instructional modelling, microteaching, lesson preparation, classroom implementation, and individualized coaching. This integrated design enabled participants to translate theoretical knowledge into authentic classroom practice rather than limiting professional learning to conceptual understanding.

The present findings are also consistent with those reported by Silalahi and Sahara (2022) and Ulfah et al. (2023), who emphasized that professional competence develops through systematic training, mentoring, teaching experience, and sustained follow-up support. The EduReligious English Model extends these principles by explicitly linking training sessions with classroom implementation. The high level of implementation fidelity, together with the superior performance of the experimental group, is consistent with the proposition that sustained coaching facilitates the transfer of professional learning into instructional practice. Nevertheless, this interpretation remains inferential rather than causal, as demonstrating coaching as a mediating mechanism would require direct process measures and a substantially larger sample.

The observed improvements in instructional language use may also be interpreted in light of previous studies highlighting the effectiveness of role-playing (Sani et al., 2024) and picture-book-based instruction (Harefa et al., 2025). Songs, storytelling, language games, and role-play provide meaningful communicative contexts in which teachers can practice authentic classroom English. It is important to note, however, that the present study assessed teachers' instructional competence rather than children's English achievement. Consequently, no conclusions should be drawn regarding improvements in children's English proficiency unless such outcomes are measured using valid and reliable assessment instruments.

Several mechanisms may explain the effectiveness of the intervention. First, the structured training enhanced both pedagogical knowledge and functional English proficiency by providing teachers with a systematic framework for selecting classroom language, formulating learning objectives, and integrating Islamic values into instructional activities. Second, observational learning occurred through facilitator modelling, allowing teachers to observe effective use of classroom English, movement-based instruction, visual materials, storytelling, and appropriate responses to children's errors. Third, deliberate practice was promoted through repeated microteaching sessions accompanied by structured, criterion-based feedback, enabling participants to refine their instructional performance progressively.

Compared with conventional professional development, the principal strength of the EduReligious English Model appears to lie in the integration of its instructional components. Demonstration alone provides examples without ensuring mastery, whereas practice without explicit performance criteria limits the effectiveness of feedback. Similarly, classroom implementation without follow-up coaching may allow ineffective practices to persist. The EduReligious English Model integrates instructional modelling, guided practice, classroom implementation, and reflective coaching into a coherent developmental cycle. Moreover, Islamic values are embedded throughout learning objectives, stories, classroom routines, language choices, and teacher–child interactions rather than being presented as isolated theoretical content. This integrated structure theoretically distinguishes the intervention from conventional workshops of comparable duration.

These findings have important practical implications for IGRA and administrators of Raudhatul Athfal (RA) institutions. Teacher professional development should extend beyond the delivery of theoretical knowledge by incorporating classroom demonstrations, guided practice, systematic observation, constructive feedback, instructional material revision, and implementation mentoring. Training content should be grounded in authentic classroom situations routinely encountered by teachers, including giving concise instructions, opening learning activities, managing classroom transitions, reading stories aloud, facilitating educational games, and concluding lessons through reflection and prayer. Likewise, the integration of Islamic values should be evaluated through observable instructional practices rather than merely the frequency of religious vocabulary.

IGRA may also strengthen professional learning by establishing communities of practice in which teachers regularly review microteaching videos, exchange instructional resources, and discuss challenges encountered when using English in early childhood classrooms. Educational administrators should ensure that teachers receive sufficient instructional time, access to appropriate teaching materials, and institutional support from school leaders so that competencies developed during training are maintained beyond the intervention period. Future evaluations should employ the same competency rubric across multiple assessment occasions to monitor sustained professional growth. In addition, storing dimension-level data would enable facilitators to identify whether teachers require further support in language competence, pedagogy, instructional media, religious integration, or social-emotional classroom management. Such evidence would facilitate more accurate, measurable, and sustainable professional development planning. Finally, future reporting should clearly distinguish statistical significance, practical significance, and the limitations regarding the generalizability of the findings.

Several limitations should be acknowledged. First, the study involved only 30 teachers from a single regional organization, limiting the generalizability of the findings to the broader population of RA teachers. Second, participants were assigned to groups without individual randomization; therefore, although ANCOVA reduced baseline differences, it could not completely eliminate the possibility of selection bias. Third, competency was assessed only immediately after the intervention, leaving the long-term sustainability of the observed improvements unknown. Fourth, the primary dataset contained only composite scores, preventing independent replication of analyses related to instrument validity, reliability, and changes across individual competency dimensions. Finally, the confidence intervals

surrounding the effect sizes remained relatively wide because of the modest number of participants in each group.

Future studies should employ cluster randomization across a larger number of Raudhatul Athfal (RA) institutions, include follow-up assessments conducted four to eight weeks after the intervention, and utilize multilevel analytical approaches where appropriate. Once improvements in teachers' professional competence have been firmly established, subsequent research may investigate their impact on classroom interaction quality and children's English learning outcomes. Mediation analyses could further examine whether teacher self-efficacy, instructional material quality, or implementation fidelity explain the observed intervention effects. Replication across different geographical regions and participant populations with varying levels of initial competence is also needed to evaluate the robustness, scalability, and equity of the EduReligious English Model.

CONCLUSION

The findings indicate that the EduReligious English Model was effective in improving the English teaching competence of Raudhatul Athfal (RA) teachers within the sample drawn from the IGRA of Deli Serdang Regency. After controlling for pretest scores, the experimental group achieved significantly higher posttest scores than the control group, $F(1, 27) = 48.17, p < .001$, partial $\eta^2 = .641$. The adjusted mean difference reached 10.35 points, while a Hedges' g value of 1.45 indicated a large intervention effect. Furthermore, the experimental group achieved a moderate mean normalized gain (N-gain), which was substantially higher than that of the control group, whose average N-gain fell within the low category. Collectively, these findings support the potential of the EduReligious English Model as a professional development program that integrates functional English instruction, early childhood pedagogy, and Islamic values.

From a conceptual perspective, the model represents a promising framework for professional development by combining functional English, early childhood pedagogical principles, and Islamic values within a coherent instructional approach. Its implementation, however, requires facilitators with interdisciplinary expertise, sufficient time for microteaching and coaching activities, developmentally appropriate instructional materials, and systematic performance evaluation. Broader implementation should be recommended only after further empirical evidence demonstrates consistent effectiveness across diverse contexts, confirms the psychometric quality of the assessment instruments, and establishes that the educational benefits justify the resources required for implementation. Ongoing evaluation will also be essential to ensure the model's long-term effectiveness and sustainability.

REFERENCE

- Abdullah, A., Wijayanti, A., Suryono, W., Joesidawati, M. I., Sarnoto, A. Z., Hiola, S. F., Nugraha, A. R., Sari, W. D., Sukini, S., & Santosa, T. A. (2024). Qualitative study: Comparison of implementation of the effectiveness of the ethno-religious-based SAVI model in improving problem-solving skills in PAI learning. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 8(5), 1245-1256. <https://doi.org/10.31004/obsesi.v8i5.6192>

- Alanís, I., & Sturdivant, T. D. (Eds.). (2023). Focus on developmentally appropriate practice: Equitable and joyful learning in preschool. National Association for the Education of Young Children.
- Cohen, L., Manion, L., & Morrison, K. (2018). Research methods in education (8th ed.). Routledge.
- Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications.
- Field, A. (2018). Discovering statistics using IBM SPSS Statistics (5th ed.). SAGE Publications.
- Friedman, S., Wright, B. L., & Masterson, M. L. (Eds.). (2022). Developmentally appropriate practice in early childhood programs serving children from birth through age 8 (4th ed.). National Association for the Education of Young Children.
- Garton, S., & Copland, F. (Eds.). (2019). The Routledge handbook of teaching English to young learners. Routledge.
- Harefa, S. M., Mustapa, N., Kurniati, E., & Wardhani, W. D. L. (2025). The effect of reading one picture storybook each day children's language skills and character growth. PAUDIA: Jurnal Penelitian dalam Bidang Pendidikan Anak Usia Dini, 14(4), 775-790. <https://doi.org/10.26877/paudia.v14i4.2268>
- Harmanto, B., Wijayanto, H., Rahmawati, I. Y., & Fadlillah, M. (2021). The early childhood education teachers' needs of being able to teach English well. Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini, 6(3), 1382-1392. <https://doi.org/10.31004/obsesi.v6i3.1256>
- Jamin, H., & Rohman, N. (2026). TRANSFORMASI DIGITAL DI MADRASAH IBTIDAIYAH DI KOTA BANDA ACEH DAN KABUPATEN ACEH BARAT. *BIDAYAH: STUDI ILMU-ILMU KEISLAMAN*, 246-256.
- Kusumaningtyas, N., & Aprianto, R. (2025). Literature analysis on classroom management in early childhood education in Indonesia. Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini, 9(1), 350-358. <https://doi.org/10.31004/obsesi.v9i1.6786>
- Marsegi, S. M., Nurhayati, S., Ansori, A., & Hendriana, H. (2023). Digital-based portfolio assessment competence of early childhood educators. Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini, 7(1), 251-259. <https://doi.org/10.31004/obsesi.v7i1.3360>
- Natasya, A. S., Rohman, N., & Andhika, M. R. (2026). THE INFLUENCE OF THE PICTURE AND PICTURE MODEL ON THE LEARNING OUTCOMES OF SBDP MADRASAH IBTIDAIYAH STUDENTS. *JURNAL TARBIYAH*, 33(1), 60-75. <https://doi.org/10.30829/tar.v33i1.5506>
- Oktarina, N., & Kurniawati, F. (2025). Assessing teachers' social emotional competence: A systematic review. Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini, 9(1), 285-294. <https://doi.org/10.31004/obsesi.v9i1.6792>
- Pinter, A. (2017). Teaching young language learners (2nd ed.). Oxford University Press.
- Rakhmania, R., Purwanti, M., & Riyanti, B. P. D. (2023). Gambaran kompetensi pedagogik guru PAUD dalam memahami teori dan praktik pendidikan untuk anak usia dini. Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini, 7(6), 6591-6608. <https://doi.org/10.31004/obsesi.v7i6.5340>
- Rohman, N., Hardiyati, M., Rizqia, M., Simangunsong, N., & Wulandari, D. R. (2023). Edu-Technopreneurship: The Concept Of Educational Business Development In The Digital Era. *Proceedings: Dirundeng International Conference On Islamic Studies*, 35-48. <https://www.ejournal.staindirundeng.ac.id/index.php/dicis/article/view/2573>
- Rozikan, M., & Adisti, A. R. (2025). A qualitative analysis of innovative teaching and counselling techniques to foster early childhood growth and resilience. Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini, 9(3), 722-734. <https://doi.org/10.31004/obsesi.v9i3.6928>

- Sani, S., Syamsi, K., Wahid, A., Solissa, E. M., Wantu, H. M., Panggabean, T. E., Kuniawan, A., & Santosa, T. A. (2024). The effectiveness of ethno-pedagogy-based role playing models in improving language and reading skills in early childhood. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 8(6), 1568-1576. <https://doi.org/10.31004/obsesi.v8i6.6273>
- Silalahi, R. Y. B., & Sahara. (2022). Upaya pengembangan SDM guru PAUD berbasis kompetensi profesional. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 6(6), 6478-6491. <https://doi.org/10.31004/obsesi.v6i6.2691>
- Simanjuntak, M. B. (2024). Exploring the nexus between English language competency in listening skills and emotional social development. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 8(3), 473-481. <https://doi.org/10.31004/obsesi.v8i3.5401>
- Sukini, S., Uluk, E., Masruchiyah, N., Nurhayati, R., Safar, M., Agustina, I., Sari, W. D., & Santosa, T. A. (2024). Effectiveness of blended learning model assisted by Scholoogy to improve language skills of early childhood education teachers. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 8(6), 1363-1374. <https://doi.org/10.31004/obsesi.v8i6.6226>
- Sutiyono, A., Hastomo, T., & Tanod, M. J. (2022). Educators' perception towards early childhood education in technology integration: A case study. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 6(6), 7323-7333. <https://doi.org/10.31004/obsesi.v6i6.3837>
- Ulfah, S. M., Anisah, A., Anisia, I. W., Ridwan, R., & Bangsawan, I. (2023). Efektivitas diklat berjenjang tingkat dasar untuk meningkatkan kompetensi profesional guru PAUD. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 7(5), 6415-6427. <https://doi.org/10.31004/obsesi.v7i5.4311>
- Witarsa, R., & Alim, M. L. (2022). Kompetensi profesional guru pada lembaga pendidikan anak usia dini. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 6(6), 5799-5807. <https://doi.org/10.31004/obsesi.v6i6.3258>