

Integrating Tradition and Pedagogical Innovation: A Training of Trainers (ToT) Model for *Sharaf* Instruction in Ma'had Aly

Muhammad Fadhil Al Ghifari¹, Abdurrahman², Luqman Hakim³

¹ Universitas Al-Qolam Malang, Malang, Indonesia; fadhilalghifari23@pasca.alqolam.ac.id

² Universitas Al-Qolam Malang, Malang, Indonesia; Abdurrahman@alqolam.ac.id

³ Universitas Al-Qolam Malang, Malang, Indonesia; Luqmanhakim@alqolam.ac.id

ARTICLE INFO

Keywords:

Training of Trainers (ToT);
Arabic morphology instruction;
pesantren education;
active learning strategies;
formative assessment

Article history:

Received 2025-09-26

Revised 2025-11-26

Accepted 2026-03-16

ABSTRACT

Arabic morphology (*Sharaf*) forms the foundation of Arabic literacy in Indonesian Islamic boarding schools (*pesantren*). However, limited pedagogical preparation among *Sharaf* teachers often leads to uneven learning outcomes and weak formative monitoring. This study develops a contextual Training of Trainers (ToT) model that integrates the classical text *Amsilatu al-Taṣrīfiyyah*, the traditional *bandongan-sorogan* instructional system, and active learning strategies within a structured formative assessment framework. A qualitative case study was conducted at Ma'had Aly An-Nur II Al-Murtadlo from July 2024 to June 2025. Data were collected through classroom observations, semi-structured interviews, and document analysis involving 16 teacher-trainees, 32 classes (≈1,255 students), and several institutional informants. The ToT program followed a four-stage cyclical model: short workshops, micro-teaching, peer observation and feedback, and journal-based reflection. Process indicators (teaching rubric scores and reflection quality) and outcome indicators (student *Sharaf* scores and hamzah-writing accuracy) were monitored quarterly. Data were analyzed using the Miles–Huberman–Saldaña framework. The findings show substantial improvements in pedagogical practice and student achievement. The average teaching rubric score increased from 2.8 to 3.6, while reflection quality rose from 2.5 to 3.7. Student learning outcomes also improved, with average *Sharaf* scores rising from 60–64 to 72–77 and hamzah-writing accuracy reaching 75–80%. These trends indicate a clear relationship between strengthened teacher pedagogical competence and improved student learning outcomes. The ToT–*Sharaf* model demonstrates that *pesantren* traditions can function as catalysts for pedagogical innovation when combined with active learning and formative assessment. The model offers a replicable framework for teacher development and sustainable improvement of Arabic morphology instruction in *pesantren* contexts.

This is an open access article under the [CC BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.



Corresponding Author:

Muhammad Fadhil Al Ghifari

Universitas Al-Qolam Malang, Malang, Indonesia; fadhilalghifari23@pasca.alqolam.ac.id

1. INTRODUCTION

Sharaf (Arabic morphology) is the foundation that determines the accuracy of reading, writing, and understanding Arabic texts, both classical and modern, and thus serves as a pillar in the learning ecosystem in Islamic boarding schools (*pesantren*). However, in many Islamic schools (*madrasah diniyah*), the availability of pedagogically competent *Sharaf* teachers is inadequate; as a result, learning outcomes tend to be uneven across classes/levels and are difficult to measure consistently. Several recent studies emphasize that strengthening *Sharaf* competency in Islamic boarding schools requires reconciling the strength of the traditional yellow book literacy tradition with modern, evidence-based pedagogical strategies (Aliyah, 2018; Raswan et al., 2022). With a large and heterogeneous student population, the context of An-Nur II Al-Murtadlo demands a teacher development model that is contextual to the Islamic boarding school, structured, and measurable in its impact—not simply increasing teaching hours or varying material (Aliyah, 2018; Krisnawilujeng, 2023).

The *bandongan-sorogan* tradition has proven to be a robust vehicle for transmitting knowledge in Islamic boarding schools (*pesantren*); it enables the internalization of textual discipline and learning etiquette. However, in large classes, this tradition faces challenges: uneven active participation, difficult early detection of misconceptions, and often delayed feedback. Studies in Islamic boarding schools (*pesantren*) demonstrate the need to restructure classroom interactions so that oral-text interaction remains primary, but is supported by drills, individual guidance, and analogy exercises that facilitate *tashrif* patterns (Ansyah, 2020; Raswan et al., 2022). In theory, the Pedagogical Content Knowledge (PCK) framework emphasizes that teaching effectiveness depends on the integration of *Sharaf* content knowledge with pedagogical strategies appropriate to the learner's context and objectives (Shulman, 1986, 1987). Therefore, strengthening *Sharaf* cannot simply stem from “mastery of rules” but must be integrated with the design of learning experiences that provide space for targeted noticing, scaffolding, and deliberate practice.

First, training/professional development for Arabic language teachers in Indonesia has proven beneficial, but is largely general in nature and has not been specifically tailored to the *Sharaf* subdiscipline and the needs of Islamic boarding school classrooms (Makruf, 2020; Aliyah, 2018). Second, the integration of local traditions—*bandongan/sorogan*—with modern pedagogy is still often reported descriptively, rarely accompanied by measurable performance indicators such as grade trends per *cawu* or writing accuracy indicators (e.g., *hamzah*), even though both are necessary for ongoing quality audits (Krisnawilujeng, 2023; Raswan et al., 2022). Third, the research corpus tends to evaluate one aspect—teacher competence or student achievement—so that the relationship between training → practice → learning outcomes is less fully depicted in the Islamic boarding school context (Aliyah, 2018; Krisnawilujeng, 2023). These threads point to the need for a design that integrates tradition, active strategies, and formative assessment into an operational and measurable architecture.

Based on the preceding synthesis, several gaps can be identified in the existing literature. First, there is no specific Training of Trainers (ToT) model for *Sharaf* that systematically maps instructional outcomes according to the sequence of *Amsilatu al-Taşrifīyyah*. Consequently, the alignment between textbook progression and the stages of pedagogical training has not been adequately documented (Aliyah, 2018). Second, although the integration of the *bandongan-sorogan* tradition with active learning strategies and formative assessment has been discussed conceptually, it is rarely evaluated using reliable and repeatable indicators across instructional periods (Krisnawilujeng, 2023; Zhang et al., 2024). Third, two-tier impact evaluations—covering both the pedagogical competence of prospective teachers and student learning outcomes—remain largely descriptive, relying primarily on narrative reports rather than on linked process–outcome indicators (Makruf, 2020; Schellekens et al., 2021).

Accordingly, a contextual, operational, and measurable *Sharaf* ToT model is required so that its quality can be replicated and systematically audited within the Islamic boarding school ecosystem. This study therefore aims to: (1) design a *Sharaf*-ToT model that aligns the sequence of *Amsilatu al-Taşrifīyyah* with the *bandongan-sorogan* tradition and active learning strategies; (2) map process and outcome indicators that enable systematic quarterly monitoring; and (3) evaluate the model's two-layer impact—

on teacher pedagogical competence and student learning outcomes—within the context of An-Nur II, which features large and heterogeneous classes. The study contributes by proposing a measurable Sharaf-specific ToT architecture, including process indicators (e.g., peer observation and reflective journals) and outcome indicators (e.g., quarterly grade trends and writing accuracy), enabling the empirical tracing of relationships between training, classroom practice, and learning outcomes. The theoretical framework draws on Pedagogical Content Knowledge (PCK) (Shulman, 1986, 1987) and formative assessment in language learning (Schellekens et al., 2021; Zhang et al., 2024).

2. METHODS

This study uses a qualitative case study approach (single case) to examine in depth the construction and implementation of the Training of Trainers (ToT) model in the field of Sharaf in the context of Ma'had Aly An-Nur II Al-Murtadlo. The choice of this design is based on the need to understand the process, actors, and institutional context holistically, while also enabling explanation building and pattern matching between the designed process-outcome indicators and field findings (Creswell & Poth, 2018; Yin, 2018). The single case focus is relevant because the ToT-Sharaf program is a bounded system with a clear locus, period, and participants, so that thick description can be achieved without losing the depth of evidence.

The research was conducted in the An-Nur II Islamic boarding school ecosystem, which houses a madrasah diniyah (Islamic religious school) with a large and heterogeneous student population. The ToT (ToT) design was conducted in June 2024 to map learning outcomes, develop an observation rubric, interview guide, and a reflection journal sheet that tracks the sequence of chapters in Amsilatu al-Taṣrīfiyyah. The ToT and teaching practice took place from July 2024 to June 2025 in the form of a teaching-practice cycle and classroom observations. Impact evaluations were conducted quarterly (in line with the cawu calendar) to compile evidence of the process and results periodically. The researcher acted as an observer-as-participant in the ToT and classroom observations, as well as as a data collector (interviews and document studies), while maintaining critical distance through reflexive memoing to manage researcher bias (Merriam & Tisdell, 2016).

The research participants consisted of three layers. The first layer consisted of 24 prospective teachers (final semester students) selected through the Sharaf content test and micro-teaching screening. 16 students who met the inclusion criteria (content test passing score $\geq 70/100$, commitment to attending all stages of the ToT, and willingness to teach) became the ToT teachers and served as the main unit of process analysis.

The second layer consisted of students in 32 classes ($\approx 1,255$ students) taught by 16 teachers (an average of two classes per teacher). To capture the learning experience representatively, one student per class ($n = 32$) was selected using representative purposive sampling, taking into account level (junior high school/senior high school), achievement (high–medium–low), and attendance. The third layer consisted of key informants (madrasah leaders, supervising lecturers, and pesantren administrators) to enrich the policy perspective and program sustainability. The sample size and composition were directed at sufficient information to observe signs of theme saturation in interviews and observations (Guest et al., 2020), while also supporting cross-role source triangulation (Miles et al., 2014).

Table 1. Framework Participants & Sampling Techniques

Layers	Unit/Role	N	Criteria Inclusion	Reason for Selection	Notes
1	Prospective teachers (students) end	24	Passed the Sharaf content test & micro-teaching screening	Source candidate instructor class	16 passed as teacher
2	Teacher ToT (process unit)	16	Commitment follow all over cycle ToT; load ± 2 classes	Evaluate change practice	Code M-01...M-16

3	Students (result units)	≈1,255 (32 classes)	All over students in 32 classes taught teacher ToT	Catch achievements Study	Interview sample : 1/ class (n=32)
4	Informant key	3-5	Madrasah leaders, lecturers mentor, caregiver	Perspective policies & support	Interview deep

Intervention ToT designed four unifying stage tradition bandongan – sorogan with modern pedagogical strategies and assessment formative. Workshop -oriented stage achievements guide teacher formulate objective operational per sub- chapter of Sharaf (e.g. *wazan, bina ' mu' tal, i' lāl*), compiling worked examples, and planning exercise tiered (easy→medium→difficult), including exit tickets and quick checks as part assessment formative. The micro-teaching stage (10–15 minutes per teacher) provides room simulation evaluated learning with rubric bait back (scale 1–4) at four dimension: clarity goals & flow, management classes & scaffolding, active strategies, and assessment formative fast. The peer observation & feedback stage facilitates observation cross between teachers in class real use same rubric, followed by a 5–10 minute debrief containing appreciation and two specific action points. Stage reflection-based journal, direct teacher analyzes achievements, misconceptions that arise (especially in *mu' tal / i' lāl*), effective / ineffective strategies, and plan improvements for the next cycle. Fourth stage. This walk repeats (cycles) according to the calendar, Cawu, and progress material, *Amsilatu al- Taṣrīfiyyah*.

Table 2. Stage Flow Session ToT (duration 60 minutes; weekly)

No	Minute	Stage	Main Activities	Role of Participants /Lecturers	Product /Evidence
1	0-10	Short Workshop	Engineering modeling lecturer teaching weekly Sharaf topics; agreeing on learning outcomes & indicators formative	Lecturers & all student	Point sheet mini lesson plan key / sketch
2	10-30	Micro-teaching	1-2 students to practice mini lesson (7-12 minutes)	Student practitioner	Video/ sheet scenario; notes observer
3	35-55	Peer observation & Feedback come back	Peer fill in sheet observation, lecturer summarize 2-3 action points	Colleagues & lecturers	Observation sheet; summary of observer actions
4	55-56	Planning repair	Practice write plan repair short	Student practitioner	Journal reflection; mini lesson plan

Data sources include observation classroom, semi- structured interviews, and documentation. Observation carried out at least three times per teacher per cycle (at least ±48 observations/cycle), using 4- dimensional rubric (scale 1–4) for evaluate clarity goals & flow, management class & scaffolding, active strategies (question and answer) guided, practice structured, worked examples), and assessment formative fast (exit ticket, mini-quiz, short error analysis). Interview covers teacher (e.g., sub- chapter linkages with achievements cawu, handling misconceptions, *mu' tal / i' lāl*, usage assessment formative, and its impact on improvement learning), students (experience learning, usefulness, and the need to come back fast), and key informants (support policy, eligibility replication, and the need for infrastructure). Documentation consists of on mini RPP, journal reflection teacher, peer-feedback sheet, recap cross -Sharf value cawu for 32 classes, example product study (sheet work) for analysis errors, as well as minutes program coordination.

Table 3. Data Sources, Instruments, and Output

Data source	Instrument	Focus/ Variable	Frequency	Output/Format
Observation class	dimensional rubric (1–4)	Goals & flow; management & <i>scaffolding</i> ; active strategies; assessment formative	≥3×/ teacher / cycle	Score/ dimension + notes qualitative
Interview	structured guide (per role)	Practice teaching; misconception; action carry on <i>feedback</i> ; support policy	End of cycle & quarter	Transcript summary + <i>memo</i>
Documentation	Mini lesson plan; journal reflection; <i>peer-feedback</i>	Planning; reflection; action points	Weekly / every session	archive / sheet stuffing
Achievements students	Recap value per <i>cawu</i> ; <i>anchor items</i> accuracy Hamzah	Average value; % accuracy; remedial	Cawu 1–3	Summary table <i>cawu</i> (ready graphics)

The indicator study is operationalized in two interconnected domains. In the process realm, the main indicator is score rubrics (1–4) per dimension, notes misconceptions that arise and how they are handled, as well as quality reflection (depth analytical, evidence supporters, specificity plan action continue). In the realm of results, indicators covers trend average cross -Sharf value *cawu* (1→2→3) per class, accuracy writing (eg. *hamzah*) on anchor items is equivalent every *cawu*, and gradual remedial notes For *bina' mu' tal / i' lāl* (target, strategy, progress). Indicator results are positioned as an amplifier for evaluating direction changes and verifying process findings, instead of as base inference statistics (Miles et al., 2014).

Table 4. Process & Outcome Indicators (definition operational & calculation)

Realm	Indicator	Definition Operational	How to Calculate	Source
Process	Rubric / dimension scores	Teacher performance in 4 dimensions (1–4)	Average per dimension per cycle	Observation sheet
Process	Quality reflection	Depth analytical & action next (1–4)	Rubric score reflection	Journal reflection
Results	Average value <i>cawu</i>	Sharaf's achievements each class per <i>cawu</i>	Class average → averaged across 32 classes	Recap mark
Results	Accuracy Hamzah	% correct on <i>anchor items</i> equivalent	(# correct / total) × 100%	Assignment / exam sheets
Results	IQR (distribution)	Width of spread value per <i>cawu</i>	Q3 – Q1	Recap mark

Data collection follows channel timing. At this stage, pre-implementation (June 2024) was carried out, including compilation of curriculum ToT, rubric blueprint observation, interview guide, journal format reflection, and peer-feedback SOP; all instruments were piloted on one class for perfect language and anchor rubric. At stage implementation (July 2024–June 2025), each sub- chapter taught through workshop sequence → micro-teaching → observation class → debrief → reflection → lesson plan adjustments; held every two weeks 30–45 minute peer-coaching clinic For dissecting teaching clips or notes key observation. Recording field done consistently throughout observation with the format context – event – interpretation early. In the evaluation quarterly, process indicators (scores rubric, action points, quality reflection) and results (average value, accuracy) were compiled To assess the trajectory of change, then discussed in a meeting, with a 60–90 minute reflection to validate findings, and agreed on a repair cycle next.

Data analysis follows Miles–Huberman–Saldaña framework through three stage overlapping: data condensation, data display, and conclusion drawing/verification (Miles et al., 2014). In condensation, open coding was carried out on the transcripts, interviews, notes observation, journal reflections, and artifacts class; example code beginning includes goal-setting clarity, scaffolding move,

worked-example use, formative-check, misconception handling, peer-feedback uptake, and reflection depth. Short figures (average value and percentage accuracy) extracted as a quantitative memo amplifier. On the display, the results served in trajectory table per teacher (average score rubric / dimension \times cycle + action points), matrix Triangulation (Teacher – Student – Leader \times Theme: design, active strategy, assessment formative, reflection, constraints mu' tal / i' lāl), as well as map theme (theme parent \rightarrow sub-theme \rightarrow quote evidence). The verification stage combines pattern matching (e.g., correlation narrative between score strengthening and trending rubrics improved value / accuracy) with explanation building (mechanism) of how worked examples + exit ticket lower misconception specific), accompanied by brief member checking to represent each role. For clarification.

Data validity is maintained through triangulation sources and methods, member checking on the summary core findings, and peer debriefing with a lecturer mentor for measuring fairness interpretation. Dependability is guaranteed with an audit trail (version instruments, dictionary code, decision collapse / expand themes, dated analytic memos), while confirmability is enforced through a chain of evidence tracking from claim to evidence and traces of reflexive memoing. Transferability is sought via the thick description above context (structure ToT, calendar academic, profile class, infrastructure) so that readers can evaluate the suitability of the context with the setting (Lincoln & Guba, 1985; Nowell et al., 2017). Management ethics includes written informed consent, right to withdraw, anonymization (M-01...M-16; Class-01...Class-32), encrypted data storage with control access tiered, and adherence to academic integrity. The potential for role bias and the Hawthorne effect was minimized through observation, repetition, blending roles of researchers, and normalizing the existence of observers in class.

Limitations of the methodological approach are stated in a transparent way. Single site (an Islamic boarding school) limits generalization, so that focus on transferability instead of generalization statistics. Indicators results nature descriptive comparative, not inferential; this is in harmony with the target study to evaluate direction changes and the mechanisms behind them, not to estimate the effect on the population. Although the design is cyclical, process– outcome related indicators, and a robust audit trail allow replication methodologically in other Islamic boarding school settings.

3. FINDINGS AND DISCUSSION

3.1 ToT Model Design - Sharaf

The research findings confirm that the design of the Training of Trainers (ToT) program in Sharaf at Ma'had Aly An-Nur II Al-Mutadlo is not merely a short-term training initiative, but rather an integrated and contextually grounded model of pedagogical capacity development. This model combines the classical grammatical framework of Amsilatu al-Taşrīfiyyah—a foundational reference in the teaching of Arabic morphology in Islamic boarding schools—with the traditional bandongan-sorogan approach, which emphasizes oral-textual interaction. In addition, it incorporates active learning strategies and formative assessment practices that characterize contemporary pedagogy. The model is structured around four core phases—short workshops, micro-teaching, peer observation and feedback, and reflective practice—which collectively form a cyclical learning process. This cycle facilitates the progressive development of teacher competence, moving from foundational understanding to sustained reflective practice.

The success of this design is closely linked to the pivotal role of supervising lecturers and institutional leaders at Ma'had Aly, who emphasized a co-design principle from the outset. This collaborative approach integrates academic perspectives with the practical needs of madrasah diniyah. Such a principle aligns with Lewis's (2002) conception of lesson study, wherein educators collaboratively design, observe, and systematically reflect on instructional practices. Furthermore, Darling-Hammond et al. (2017) argue that effective teacher development programs must be grounded in iterative practice and continuous improvement cycles. Within the context of the ToT-Sharaf program, these principles are operationalized through a recurring cycle of workshops, micro-teaching, peer

observation, and reflection. This cycle not only strengthens mastery of morphological content but also cultivates a reflective habitus among prospective teachers.

At the national level, these findings complement the study by Nurhidayati et al. (2022), which demonstrates that micro-teaching based on peer feedback significantly enhances novice teachers' pedagogical competence. However, the ToT-Sharaf model introduces a novel dimension through the full integration of the bandongan–sorogan tradition into a modern teacher training framework. This indicates that pedagogical capacity building in this context is not merely a transplantation of Western methodologies, but rather a creative adaptation that preserves the essence of the Islamic boarding school tradition while incorporating evidence-based practices. Consequently, this study reinforces the notion that local traditions can not only coexist with contemporary pedagogical innovations but can also enrich the learning process (Raswan et al., 2022).

Five achievements formulated learning the core of the success of this model : (1) mastery Sharaf's core material from pattern *tsulātsī* until *mazid* with accuracy linguistics high; (2) ability combine method traditional with technique active like quiz fast and discussion group; (3) skills management adaptive class to size class large and heterogeneous; (4) application assessment formative fast, for example exit tickets and quick checks, which enable teachers to detect misconception in real time; and (5) skills reflection self through journal post-teaching that leads to improvement sustainable. Integration of dimensions, content, and pedagogy. This is fully in line with the concept of Pedagogical Content Knowledge (PCK) initiated by Shulman (1987), namely uniting knowledge content and method delivery that allows teachers not only to know "what" to teach, but also "how" to teach it in an effective and contextual way.

Table 5. ToT Process and Result Indicators - Sharaf

Main Components	Design Description	Specific Goals	Products/Evidence Produced
Curriculum Foundation	The book <i>Amsilatu al-Taṣrifīyyah</i> was chosen as the core reference for studying Arabic morphology.	Ensuring integration between Sharaf's classical content and modern pedagogical objectives.	Map of the material sequentially from <i>tsulātsī</i> to <i>mazīd</i> , including the chapters <i>mu'tal</i> and <i>i'lāl</i> .
Islamic Boarding School Traditions	The bandongan–sorogan method is maintained as the main interaction format between teachers and students.	Maintaining the unique character of Islamic boarding schools and ensuring internalization of textual discipline.	Practice bandongan & sorogan in each session, sorogan notes.
Active Learning Strategies	Integration of <i>active learning</i> (discussion, <i>exit ticket quizzes</i> , <i>worked examples</i>).	Increase student participation, strengthen understanding of morphological concepts.	Weekly quiz questions, student worksheets.
Stage 1 – Short Workshop	The lecturer models the teaching technique for the weekly Sharaf topic (first 10 minutes).	Provide practical examples and agree on <i>learning outcomes</i> and formative indicators.	Key point sheet, mini lesson plan sketch.
Stage 2 – Micro-teaching	1–2 students practice a 7–12 minute mini-lesson.	Practice teaching skills on a small scale with quick feedback.	Video/scenario sheets, observer notes.
Stage 3 – Peer Observation & Feedback	Colleagues observe teaching practice, lecturers summarize 2–3 <i>action points</i> .	Sharpen your ability to reflect and accept constructive feedback.	Observation sheet, summary of <i>action points</i> .
Stage 4 – Journal Reflection	The student intern wrote a short improvement plan.	Cultivating <i>reflective practice</i> and continuous improvement.	Reflection journal, updated mini lesson plan.

Quarterly Evaluation Mechanism	Monitoring the achievement of the process and results of each cawu (student exam scores, hamzah accuracy, teaching rubric scores).	Ensuring the sustainability and accountability of improving pedagogical competence and student learning outcomes.	Recap of cawu values, observation reports and reflections.
---------------------------------------	--	---	--

3.2 Components and Indicators Process and Result Achievements

Evaluation ToT -Sharaf is implemented in a way quarterly (cawu) to ensure that development competence is not only natural, but can be monitored in a sustainable way. Two realms indicator mapped in a way explicit: process and results. The domain of the assessment process dynamics teachings of the students through score rubric four dimensions (clarity goals and flow, management classes and scaffolding, active strategies, and assessment formative), as well as quality reflection self in journal post-teaching. Meanwhile, the results evaluate achievements, understanding students through trend mark, Sharaf test, every cawu, level accuracy writing hamzah, and remedial notes for complex topics like bina ' mu' tal and i' lāl.

Table 6. ToT Process and Result Indicators - Sharaf

Realm	Indicator	Operational Definition	Data source
Process	4-dimensional rubric score	Clarity of goals & flow, classroom management, active strategies, formative assessment (scale 1–4)	Observation sheet
Process	Reflection quality	Depth of analysis & follow-up plan (scale 1–4)	Reflection journal
Results	Average cawu value	Average Sharaf score across 32 classes per quarter	Madrasah grade recap
Results	Hamzah's accuracy	Percentage correct on equivalent items for each cawu	Exam/assignment sheets

The data indicate a consistent improvement across all process indicators. The average teaching rubric score increased from 2.8 in the first cawu to 3.6 in the third cawu, while the quality of reflective practice, initially averaging 2.5, rose to 3.7. These gains suggest a substantial internalization of reflective practices and an enhanced capacity among students to critically evaluate their own learning processes. Schön (1983) underscores that reflection-in-action and reflection-on-action constitute the core of professional competence in educators. The present findings affirm the relevance of this theoretical framework within the context of Islamic boarding school education.

In terms of learning outcomes, the average scores of Sharaf students and their accuracy in writing hamzah demonstrate a parallel upward trend, corresponding with the improvement in teacher competence. This positive trajectory indicates a strong and plausible causal relationship between the quality of instructional processes and student learning outcomes. These results are consistent with Hafidah et al. (2024), who found that active learning strategies and formative assessment significantly enhance students' understanding of Arabic morphology.

3.3 Impact: Teachers and Students

3.3.1 Improvement Competence Pedagogy Student

Deep Interview with 16 students, teacher revealed surge trust self, mastery methods and dexterity management class post-ToT. Before following the program, pattern teaching tend monotonous and lecture - cantered; afterward, the teacher start blend *bandongan* – *sorogan* with exercise *imlā* ', quiz fast, and worked examples that invite students active practice pattern *tasrīf*. Darling-Hammond et al. (2008) showed that the most effective teacher training program is when participants have their own knowledge, beginning with adequate knowledge, because training can focus on developing pedagogy. This is confirmed here: all participants have passed the selection, beginning with a minimum score of 70/100 on the Sharaf content test, so that ToT plays a role in strengthening skills methodologically, not just remedial content.

Weekly micro-teaching and peer feedback play an important role in balancing differences in background and experience in teaching. Some of the students have already experienced teaching jurisprudence or grammar, whereas other only once lead informal sorogan. This difference potentially causes resistance to the new method, but with peer coaching guidance, participants, in a way, gradually internalize this innovative approach. This finding is strengthened by Nahar, Zulheddi, & Ghifari (2022), who confirm that micro-teaching in Islamic boarding schools effectively equips senior students as teacher cadres with skills reflective.

3.3.2 Achievements Student Learning

The impact of the ToT program on students' understanding of Sharaf was recorded clearly in quantitative data, exam, and precision writing, Hamzah.

Table 7. Recap Achievement of Understanding of Madrasah Diniyah Students Post- ToT

Period (2024–2025)	Main Material (Amsilatu al-Taṣrīfiyyah)	Percentage of Students Understanding the Material	Hamzah's Accuracy in Writing
Aug–Sep 2024	Basic Wazan & Mauzun	60–68%	70%
Oct 2024	Bina' tsulātsī & hamzah rules	70–78%	75%
Nov 2024	Rubā'ī mujarrod & fa'idah patterns	80–85%	78%
Dec 2024–Feb 2025	Bina' mu'tal & i'lāl	65–72%	76%
Mar–Jun 2025	Reinforcement of all patterns & final remedial	75–82%	80%

Ascension mark exam from 60–64 on first cawu to 72–77 on third cawu, accompanied by accurate writing, Hamzah's stable score in the range of 75–80%, confirms that the quality of teaching has an impact on understanding morphology santri. Eventhough thus, the decline achievement in the material bina ' mu' tal and i' lāl indicates that topics end Amsilatu al- Taṣrīfiyyah demanding more teaching strategies clinical and gradual, for example with development of thematic item banks and worked examples as suggested by Darling-Hammond et al. (2017).

The students' perspective also emphasized change positive. A students 1st grade high school said that quizzes help them remember patterns and tasks more, whereas the Head of Madrasah Diniyah looks at the presence of students in ToT as a "strategic step" that adds discipline to students and guards the quality of teaching, although it means limited. This testimony indicates that the impact of ToT goes beyond improving the competence of individual teachers and provides a contribution to the institutional reality for madrasahs.

3.3.3 Analysis, Critical and Discussion Theoretical

From a theoretical perspective, the ToT–Sharaf model reinforces the relevance of Pedagogical Content Knowledge (PCK) within the context of Islamic boarding school education. Teachers are not only expected to understand the rules of Arabic morphology but also to communicate them effectively using adaptive, data-informed, and evidence-based strategies. The integration of the bandongan–sorogan methods with formative assessment demonstrates that local traditions do not hinder innovation; rather, they can serve as a foundation for meaningful pedagogical transformation (Raswan et al., 2022).

This study also contributes to the literature on Arabic language teacher training, which has largely focused on speaking skills or the use of digital technology (Rachmawati et al., 2024). By specifically addressing morphology (Sharaf), this research fills a gap identified by Aliyah (2018) and Krisnawilujeng (2023) regarding the limited emphasis on Sharaf as a subdiscipline in systematic teacher training. At the same time, the findings confirm that tradition-based training models can serve as best practices for similar educational contexts in Indonesia.

Moreover, the workshop–micro-teaching–peer observation–reflection cycle operationalizes the concept of reflective practice (Schön, 1983), where novice teachers learn through direct experience, evaluation, and continuous improvement. The use of post-teaching reflective journals encourages teachers to critically reassess their instructional decisions and refine their teaching strategies. Supported by strong institutional leadership within the pesantren, the ToT–Sharaf model successfully cultivates a reflective professional habitus, which is essential for sustaining teacher professionalism (Whitehead & Peckham, 2022).

3.3.4 Implications

The implications of this study can be examined from several interconnected perspectives, including Islamic boarding school education policy, Arabic language teaching practices, teacher capacity development, and inter-institutional collaboration. First, from a policy perspective, the findings confirm that the Training of Trainers (ToT)–Sharaf model—integrating the *bandongan–sorogan* tradition with active learning strategies and formative assessment—can serve as a robust reference for improving the quality of Arabic instruction in pesantren. Institutions facing shortages of qualified Sharaf teachers may adopt this model as a structured pre-service or in-service training program, particularly for senior students preparing to become teachers. With strong institutional support, such as internal pesantren regulations or alignment with the Ministry of Religious Affairs, the ToT program can function as a sustainable and systemic mechanism for developing teachers who not only master Arabic morphology but also possess modern, assessment-driven pedagogical competencies, consistent with recommendations by Linda Darling-Hammond and colleagues on the importance of structured professional development policies.

Second, the study highlights the importance of strengthening formative assessment and continuous monitoring. The ToT–Sharaf model demonstrates that systematic formative practices—such as exit tickets, short quizzes, and peer feedback—can be effectively embedded within traditional pesantren learning without compromising its identity. Quarterly monitoring using both process and outcome indicators enhances institutional accountability and supports data-driven decision-making. Leaders can utilize performance recaps and indicators such as students' accuracy in hamzah writing to evaluate both student progress and instructional effectiveness. This approach fosters a sustainable culture of quality improvement, aligning with insights from Catherine Lewis on the role of continuous monitoring in strengthening collective teacher professionalism.

Third, the findings suggest the need for the development of more clinical and thematic instructional materials. While overall improvements in student performance indicate the effectiveness of the model, challenges in mastering complex topics such as *bina' mu'tal* and *i'lāl* point to the necessity of scaffolded learning resources. Pesantren can develop item banks, worked examples, and case-based worksheets to support gradual and in-depth understanding. Such materials align with evidence-based recommendations that complex learning requires structured scaffolding to guide students from basic to advanced conceptual understanding.

Fourth, the model presents strong potential for replication, scalability, and inter-pesantren collaboration. The success of the ToT–Sharaf framework provides a transferable model that can be adapted to different institutional contexts. Collaborative implementation—through regional pesantren forums or Ma'had Aly networks—can expand its impact and address broader shortages of pedagogically competent Sharaf teachers. Furthermore, such collaboration can foster communities of practice among Arabic language educators, as emphasized by Etienne Wenger, where professional learning is enhanced through shared experience and collective reflection.

Fifth, at the national level, this study contributes to the development of Arabic language teacher capacity in Indonesia. By integrating Pedagogical Content Knowledge (PCK), as conceptualized by Lee Shulman, with reflective practice, as proposed by Donald Schön, the ToT–Sharaf model demonstrates that teacher development can be rooted in local traditions while remaining aligned with global pedagogical principles. This model offers a valuable reference for higher education institutions,

policymakers, and professional associations in designing standardized training frameworks that balance content mastery and pedagogical expertise.

Overall, the implications of this study extend beyond a single institution, highlighting the strategic value of the ToT–Sharaf model for strengthening Arabic language education at policy, practice, and capacity-development levels. It illustrates how traditional educational practices can be revitalized through modern pedagogical approaches, creating a productive synergy that enhances the quality and sustainability of Arabic instruction in Indonesian pesantren.

3.3.5 Limitations and Recommendations for Future Research

This study has several limitations. Practical constraints, such as limited instructional resources (e.g., damaged whiteboards and insufficient markers), occasionally affected the learning process. Moreover, although supported by quantitative data, the study primarily adopts a qualitative design, which limits statistical generalization; therefore, findings should be interpreted analytically rather than universally. Future research is recommended to examine the effectiveness of the ToT–Sharaf model in diverse pesantren contexts and to explore the integration of digital technologies as a complement to the *bandongan–sorogan* tradition, particularly in addressing the needs of digitally oriented learners.

4. CONCLUSION

This study demonstrates that the Training of Trainers (ToT)–Sharaf model developed and implemented at Ma'had Aly An-Nur II Al-Murtadlo effectively addresses long-standing challenges in teaching Arabic morphology within the pesantren context. By identifying key issues—such as limited teacher preparation in Sharaf, the absence of structured performance indicators, and the misalignment between teacher evaluation and student learning outcomes—the study provides a strong foundation for meaningful pedagogical reform. Using a qualitative case study approach with triangulated data from observations, interviews, and documentation, the research introduces a four-stage cyclical ToT framework consisting of workshops, micro-teaching, peer observation and feedback, and reflective journaling. This model successfully integrates the classical text *Amsilatu al-Taṣrīfiyyah*, traditional *bandongan–sorogan* methods, active learning strategies, and systematic formative assessment. The findings indicate significant improvements in teaching quality, reflective practice among teacher-trainees, and students' mastery of Sharaf, including better accuracy in *hamzah* writing, confirming a strong link between enhanced pedagogical competence and improved student outcomes. Notably, the study highlights that pesantren traditions can support, rather than hinder, educational innovation when combined with reflective and evidence-based approaches. While teacher-trainees show stronger instructional and classroom management skills, students achieve more consistent understanding of *tashrif* patterns, although complex topics such as *bina' mu'tal and i'lāl* still require more focused instructional support. At the institutional level, the ToT–Sharaf model proves feasible for broader adoption, supported by standardized procedures, unified assessment rubrics, and regular reflective evaluations. Although the single-site design limits generalization, the study offers a clear and transferable framework that bridges traditional Islamic education with modern pedagogical practices, ultimately enhancing teacher professionalism and improving the quality of Arabic morphology instruction in pesantren.

REFERENCES

- Aliyah, A. (2018). *Pesantren tradisional sebagai basis pembelajaran nahwu dan sharaf dengan menggunakan kitab kuning*. Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab, 6(1), 1–25. <https://doi.org/10.23971/altarib.v6i1.966>
- Ansyah, S., Ritonga, M., & Alrasi, F. (2020). Sistem kaji duduak sebagai strategi pembelajaran maharah al-qirā'ah di madrasah. Arabi: Journal of Arabic Studies, 5(2), 191–200. <https://doi.org/10.24865/ajas.v5i2.257>

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Brookhart, S. M. (2017). *How to give effective feedback to your students* (2nd ed.). ASCD.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://doi.org/10.54300/122.311>
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915–945. <https://doi.org/10.3102/00028312038004915>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). ASCD.
- Krisnawilujeng, K., Fatoni, A., & Sarimov, R. (2023). The relationship between mastery of nahwu–sharf and the ability to translate Arabic texts into Indonesian. *Arabi: Journal of Arabic Studies*, 8(2), 183–192. <https://doi.org/10.24865/ajas.v8i2.502>
- Lewis, C. (2002). *Lesson study: A handbook of teacher-led instructional change*. Research for Better Schools.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.
- Maxwell, J. A. (2010). Using numbers in qualitative research. *Qualitative Inquiry*, 16(6), 475–482. <https://doi.org/10.1177/1077800410364740>
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.
- Moton, B. A., & Daniels, C. (2022). Exploring application of the Training of Trainers (ToT) model on faculty professional development and teaching practices in a summer learning community. In *Experiences and research on enhanced professional development through faculty learning communities* (pp. 1–35). IGI Global. <https://doi.org/10.4018/978-1-6684-5332-2.ch001>
- Muhlis, M., Sumbulah, U., Maimun, A., & Barizi, A. (2023). Implementation of the andragogy approach to improve students' Arabic understanding. *Ijaz Arabi: Journal of Arabic Learning*, 6(2), 361–367. <https://doi.org/10.18860/ijazarabi.v6i2.20636>
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16, 1–13. <https://doi.org/10.1177/1609406917733847>
- Paas, F., & Sweller, J. (2012). An evolutionary update of cognitive load theory. *Educational Psychology Review*, 24, 27–45. <https://doi.org/10.1007/s10648-011-9179-2>
- Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE.
- Rachmawati, U. A., Yahya, A., Pratiwi, P. S., Utami, S. P., Hasan, F. U., & Rama Abirami, K. (2024). Takallam: An Android application to enhance literacy of Arabic expression. *Izdihar: Journal of Arabic Language Teaching, Linguistics, and Literature*, 7(3), 375–392. <https://doi.org/10.22219/jiz.v7i3.34762>
- Raswan, R., Abdul Wahab, M., & Hakki, S. (2022). Simplifikasi morfologi Arab (sharf) dengan pendekatan konstruktivisme dan analogi. *Arabi: Journal of Arabic Studies*, 7(1), 25–37. <https://doi.org/10.24865/ajas.v7i1.485>
- Schön, D. A. (2017). *The reflective practitioner: How professionals think in action*. Routledge. <https://doi.org/10.4324/9781315237473>

- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Tersta, F. W., Sekonda, F. A., Arief, H., Vahlepi, S., & Firman, F. (2022). Pelatihan pengajaran bahasa Arab bagi guru MI Quhas Primary School Dar Al-Masaleh Jambi. *DEDIKASI: Jurnal Pengabdian Masyarakat*, 4(1), 71–75. <https://doi.org/10.32332/d.v4i1.3724>
- Whitehead, J., & Peckham, M. (2022). The reflective practitioner. In *Network leadership: Navigating and shaping our interconnected world* (Chap. 18). Routledge. <https://doi.org/10.4324/9781003092582-18>
- Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.