



Ideological Vaccination: Building Santri Resilience Against Digital Radicalism

Heni Listiana¹, Naurah Reisa Alana², Tria Fahira Nuramaja³, Sherin Safitri⁴, Mohammad Holis⁵

¹Universitas Islam Negeri Madura, Indonesia ^{2,3,4,5}Madrasah Aliyah Negeri 2 Pamekasan, Indonesia

Email: ¹henilistiana@uinmadura.ac.id, ²nishimurakivampire@gmail.com, ³nuramaja3afahira@gmail.com, ⁴realsherin2@gmail.com, ⁵holismad@gmail.com

Abstract

Keywords:

Ideological
Vaccination;
Inoculation
Theory; Religious
Moderation;
Cognitive
Immunity; Peer
Mentoring;

The rapid infiltration of digital extremism poses a severe threat to Islamic boarding schools (*pesantren*), where algorithmic echo chambers exploit adolescents' emotional vulnerabilities. Recognizing the limits of conventional, purely technical digital literacy interventions, this study examines an "Ideological Vaccination" model grounded in McGuire's Inoculation Theory and operationalized through a flipped, cascading peer-mentoring framework. Using a quasi-experimental one-group pre-test and post-test design at IBS PKMKK Pamekasan, the intervention engaged three student co-researchers, 12 student mentors, and 71 seventh-grade students. Mentors underwent asynchronous exposure to weakened extremist arguments followed by synchronous, *tawasuth*-based refutational pre-emptions. Inferential statistical analysis revealed a significant gain in mentors' cognitive immunity from 87.50% to 98.33% ($t(11) = -6.842, p = 0.0001, 95\% CI [7.35, 14.31]$), with a high Normalized Gain ($N-Gain = 0.83$) and large effect size (Cohen's $d = 1.98$). Baseline vulnerability regarding fraudulent jihad donations (item S4) surged from 66.7% to 91.6%. Qualitatively, cognitive resilience translated into 12 Canva peace infographics (mean 3.95/4.00) and 71 reflective essays (mean 3.90/4.00). These findings demonstrate that ideological inoculation successfully operationalizes moderation (*tawasuth*) as an active mental heuristic, transforming *santri* from passive digital consumers into sovereign producers of moderate religious discourse.

Kata Kunci:

Vaksinasi Ideologi;
Teori Inokulasi;
Moderasi Beragama;
Ketahanan Kognitif;
Pendampingan
Sebayu

Infiltrasi ekstremisme digital menjadi ancaman nyata bagi pondok pesantren akibat manipulasi ruang gema algoritma yang memanfaatkan kerentanan emosional remaja. Mengatasi keterbatasan literasi digital konvensional yang semata-mata bersifat teknis-instrumental, penelitian ini menguji model "Vaksinasi Ideologi" berbasis Teori Inokulasi McGuire melalui kerangka peer-mentoring terbalik. Menggunakan desain kuasi-eksperimen one-group pretest-posttest di IBS PKMKK Pamekasan, intervensi ini melibatkan 3 santri peneliti, 12 mentor santri, dan 71 santri kelas VII. Mentor memaparkan narasi ekstremis dosis lemah

secara asinkron, diikuti dengan pra-sanggahan terbimbing berbasis nilai tawasuth secara sinkron. Analisis statistik inferensial paired-sample t-test membuktikan peningkatan signifikan imunitas kognitif mentor dari 87,50% menjadi 98,33% ($t(11) = -6,842$, $p = 0,0001$, 95% CI), dengan Normalized Gain tinggi ($N\text{-Gain} = 0,83$) dan effect size besar (Cohen's $d = 1,98$). Kerentanan kritis terkait donasi jihad manipulatif (item S4) telah teratasi, dengan peningkatan akurasi dari 66,7% menjadi 91,6%. Secara kualitatif, resiliensi kognitif terwujud dalam 12 infografik kesantunan Canva (skor 3,95/4,00) dan 71 esai literasi damai (skor 3,90/4,00). Temuan ini mengonfirmasi bahwa vaksinasi ideologi secara efektif mengoperasionalkan moderasi beragama sebagai algoritma berpikir aktif, mentransformasikan santri dari konsumen pasif menjadi produsen wacana keagamaan moderat yang berdaulat.

Received: September 6, 2026; Revised: September 25, 2026; Accepted: October 3, 2026

<http://doi.org/10.19105/tjpi.v21i1.26132>

© Tadris Jurnal Pendidikan Islam
Universitas Islam Negeri Madura, Indonesia



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

1. Introduction

Islamic boarding schools (*pesantren*) historically constitute the institutional bedrock of Islamic character education in Indonesia, harmonizing the epistemological transmission of Islamic sciences (transfer of knowledge) with value internalization (transfer of value). Within this pedagogical continuum, *pesantren* systematically cultivate *wasathiyyah* (religious moderation), an ethical paradigm derived from classical canonical commentaries (*turats*) and incarnated through the spiritual exemplarity (*uswah*) of the *kyai*. Doctrinally, the principle of *tawasuth* (moderation) functions as an active dynamic equilibrium, demanding uncompromising commitments to justice (*i'tidal*), balance (*tawazun*), and pluralistic tolerance (*tasamuh*) when engaging sociocultural diversity. Contemporary empirical inquiry confirms that *kyai* leadership anchored in indigenous hermeneutics and contextual wisdom remains decisive in preserving national cohesion and insulating faith-based institutions against ideological radicalization.¹ Nevertheless, pervasive digital transformation introduces unprecedented socioreligious vulnerabilities for Generation Z santri, who encounter unmediated exposure to extremist propaganda in hyperconnected cyber environments.

The proliferation of networked digital media has destabilized traditional religious authority structures.² The epistemological rigor historically secured through unbroken scholarly genealogies (*isnad/sanad*) is increasingly displaced by algorithmic aggregations, with Muslim youth prioritizing quick, search-engine-mediated religious answers over prolonged discursive mentorship.³ Furthermore, persistent infrastructural disparities and the absence of structured ideological screening curricula within *pesantren* render students defenseless against algorithmic distortion mechanisms, specifically filter bubbles and informational

¹ Muhammad Anas Ma'arif et al., "Model of Kiai Leadership Based on Local Wisdom: Preventing Radicalism and Building Education in the Global South," *Kharisma: Jurnal Administrasi Dan Manajemen Pendidikan* 5, no. 1 (2026); Zainul Abas et al., "Guardian of Faith: Comparative Insights of Teaching Islamic Theology in Indonesian and Malaysian Pesantren," *Nazhruna: Jurnal Pendidikan Islam* 8, no. 2 (2025): 382–99.

² Rüdiger Lohker and Soleh Hasan Wahid, "The Transformation of Islamic Religious Authority," *Religions* 17, no. 4 (2026): 493, <https://doi.org/10.3390/rel17040493>.

³ Eko Harianto et al., "Algorithmic Spiritual Authority," *ESENSIA: Jurnal Ilmu-Ilmu Ushuluddin* 26, no. 2 (2025): 96–108, <https://doi.org/10.14421/esensia.v26i2.7772>.

insulation.⁴ Consequently, an acute behavioral paradox emerges. While santri exhibit devout decorum in physical spaces, they often experience cognitive disorientation and ethical erosion (digital civility) when navigating polarized virtual spaces marked by emotional contagion and manipulative affective rhetoric.

Prior research about digital interventions in Islamic educational context has mostly focused on practical and utilitarian skills, including the navigation of learning management systems, the creation of digital proselytization (*dakwah*), and the filtering of technical content.⁵ While these empirical studies documented enhanced technical proficiencies among students, their interventions had significant structural limitations: they treated digital literacy as mere device handling. They did not address underlying cognitive vulnerabilities or algorithmic persuasion dynamics (filter bubbles and echo chambers).⁶ Furthermore, conventional religious moderation instruction in pesantren frequently relies on supportive, one-way didactic preaching, leaving santri cognitively helpless when confronted with weaponized theological rhetoric and clandestine cyber-radical campaigns in open virtual networks.⁷

To bridge this critical scholarly and pedagogical gap, this study develops and introduces 'Ideological Vaccination' as an original, context-specific socio-pedagogical model. Conceptually, this model is not a mere metaphorical extension of William J. McGuire's Inoculation Theory, but an innovative instructional framework developed in this research that synthesizes social-psychological immunization with indigenous Islamic epistemologies. Like biomedical immunization, inoculation theory posits that pre-existing attitudes and normative convictions can be fortified against external persuasion by pre-emptively exposing individuals to attenuated counter-arguments accompanied by guided refutations.⁸ The inoculation architecture functions through two synchronized cognitive mechanisms: threat induction, which stimulates affective vigilance and prompts cognitive defense mechanisms, and refutational pre-emption, which equips the intellect with analytical counter-framing. Systematic meta-analyses confirm that inoculation yields far more durable resistance than conventional, non-challenging supportive messaging.⁹ By operationalizing this framework within the pesantren ecology, the proposed model transforms *tawasuth* (moderation) and *tabayyun* (verification) from passive normative dogmas into an active, self-sustaining cognitive heuristic (*manhaj al-fikr*) against online radical manipulation.

⁴ Zawaqi Afdal Jamil et al., "Program Evaluation Study on Islamic Religious Education in Pesantren: Addressing Educational Degradation in the Digital Era," *Jurnal Pendidikan Agama Islam* 22, no. 1 (2025): 122–39, <https://doi.org/10.14421/jpai.v22i1.11358>.

⁵ Nawawi Fuad Mustopa and Bisri, "Edukasi Kontra Narasi Intoleran Dan Radikalisme Melalui Literasi Media Online Kepada Santri Pondok Pesantren Annida Cirebon," *Abdimas Galuh* 5, no. 2 (2023): 1026–36; Gufran Darma Dirawan, Dyah Darma Andayani, and Nurlita Pertiwi, "Peningkatan Literasi Digital Pada Santri Pesantren Syeh Hasan Yamani Memanfaatkan Aplikasi Canva," *Jurnal Masyarakat Madani Indonesia* 2, no. 4 (2023): 366–71.

⁶ Rahayu Rahmadini et al., "Penguatan Moderasi Beragama Melalui Literasi Digital Dan Pemahaman Kritis Atas Algoritma Media Sosial Bagi Para Santri Di Pondok Pesantren Belitung," *Journal of Community Development* 6, no. 3 (2026): 12, <https://doi.org/10.47134/comdev.v6i3.1862>.

⁷ Ahmad Ridwan and Ansor Feri Mahmudi, "Internalisasi Nilai-Nilai Wasathiyah Melalui Pendidikan Berbasis Pesantren," *Journal of Educational Research and Practice*, 2023, <https://doi.org/10.70376/jerp.v1i1.30>.

⁸ John A. Banas and Stephen A. Rains, "A Meta-Analysis of Research on Inoculation Theory," *Communication Monographs* 77, no. 3 (2010): 281–311, <https://doi.org/10.1080/03637751003758193>; Josh Compton, Shelley Wigley, and Sergei A. Samoilenko, "Inoculation Theory and Public Relations," *Public Relations Review* 47, no. 5 (2021): 102116, <https://doi.org/10.1016/j.pubrev.2021.102116>.

⁹ Marieke L. Fransen et al., "Sixty Years Later: A Replication Study of McGuire's First Inoculation Experiment," *J. Media Psychol. Theor. Methods Appl.* 36 (2023): 69–78, <https://api.semanticscholar.org/CorpusID:261843636>.

To optimize engagement among Generation Z learners, the intervention was operationalized through an integrated flipped classroom framework and a cascading peer-mentoring structure.¹⁰ Through the flipped model, participants asynchronously deconstruct cyber-literacy modules and analyze simulated extremist narratives, reserving synchronous classroom encounters for intensive dialectical inquiry and Focus Group Discussions (FGDs) designed to hone refutational defense skills. Concurrently, cascading peer mentoring reduces hierarchical instructional barriers, fostering egalitarian, relatable discourse among peers. Situated at Islamic Boarding School Padepokan Kyai Mudrikah Kembang Kuning (IBS PKMKK) Pamekasan, this investigation specifically aims to: (1) evaluate the empirical efficacy of the model in augmenting mentors' cognitive resistance using pre-test and post-test measurements; (2) examine how structured peer mentoring operationalizes *tawasuth* as an epistemological filter against digital disinformation; and (3) evaluate the behavioral evolution of santri from passive recipients of online narratives into sovereign, proactive creators of moderate religious discourse through digital poster design and reflective literacy production.

2. Method

This study employed a quasi-experimental one-group pre-test post-test design, complemented by qualitative portfolio content assessments to evaluate creative counter-narrative artifacts.¹¹ The empirical investigation was conducted at Islamic Boarding School Padepokan Kyai Mudrikah Kembang Kuning (IBS PKMKK) Pamekasan, East Java, Indonesia. Internal validity was rigorously safeguarded through baseline homogeneity, as all participating santri shared identical boarding-school living conditions, daily schedules, and informal Islamic curricula.

2.1 Participant Architecture and Sampling Rationale

To bridge hierarchical instructional barriers, this study adopted a participatory "Students-as-Co-Researchers" paradigm integrated into a Cascading Mentoring Model across three structured tiers:

- a. Student Researchers (3 participants, Grade 10): Selected through purposive sampling based on their demonstrated digital literacy aptitude, student leadership roles, and research interest. Rather than serving as passive experimental subjects, they acted as co-researchers and primary experimenters. Under the lead author's direct academic supervision, they co-developed training scenarios, coordinated fieldwork, and served as master trainers during synchronous sessions.
- b. Student Mentors (12 participants, Grades 8–10): Selected via purposive sampling based on peer credibility, interpersonal communication skills, and active involvement in student council activities. They constituted the primary experimental treatment group that underwent pre- and post-testing. The determination of 12 mentors was strategically calculated to maintain an optimal mentor-mentee ratio of approximately 1:6 for small-group peer circles.
- c. Impact Group Students (71 participants, Grade 7): Recruited via total population sampling, encompassing the entire incoming Grade 7 cohort at IBS PKMKK. This cohort was specifically targeted because junior adolescent santri

¹⁰ Zakariyah Zakariyah, Abdul Muhid, and Miftah Arifin, "Transformation of Al-Qur'an Memorization Learning in the Digital Age through the Blended Learning Flipped Classroom Approach," *Nazhruna: Jurnal Pendidikan Islam* 8, no. 3 (2025): 594–606, <https://doi.org/10.31538/nzh.v8i3.193>.

¹¹ Gisela Anantasia and Sulastri Rini Rindrayani, "Metodologi Penelitian Quasi Eksperimen," *ADIBA: JOURNAL OF EDUCATION* 5, no. 2 (2025): 183–92.

represent the most vulnerable demographic transitioning into unmediated digital environments without established critical filtration habits.

2.2 Conceptual and Operational Definitions of Key Variables

To establish measurement fidelity and conceptual clarity, key variables and terms were operationalized as follows:

- a. Ideological Vaccination (Independent Variable / Model): Conceptually defined as a proactive socio-pedagogical intervention grounded in McGuire's Inoculation Theory that confers cognitive resistance against extremist digital persuasion.¹² Operationally, it is implemented through a 4-stage instructional model that combines asynchronous exposure to attenuated radical propaganda with synchronous, peer-led refutational Focus Group Discussions (FGDs).
- b. Cognitive Immunity (Primary Dependent Variable): Conceptually defined as santri's intellectual capacity and affective vigilance to detect, deconstruct, and reject manipulative online extremism. Operationally, it is measured via the Cognitive Immunity Questionnaire across ten contextual scenarios (scored 0–100%), with efficacy evaluated through paired-sample t-test gains and Hake's Normalized Gain (N-Gain) score.
- c. *Tawasuth* (Epistemological Filter / Mediating Core Value): Conceptually defined as the Islamic virtue of the balanced middle path, characterized by justice (i'tidal), equilibrium (tawazun), and pluralistic tolerance (tasamuh).¹³ Operationally, it functions as the standardized evaluative heuristic employed during FGDs and poster design to debunk binary takfiri narratives.
- d. Flipped Peer-Mentoring (Instructional Delivery): An egalitarian pedagogical mechanism where content deconstruction occurs asynchronously at home via digital literacy video modules, while synchronous classroom hours are reserved for dialectical rebuttal crafting facilitated horizontally by student researchers and mentors.¹⁴
- e. Digital Civility (Downstream Impact Artifacts): Conceptually defined as prosocial communicative competence and active peace-promotion in virtual spaces. Operationally assessed through standardized descriptive rubrics (scale 1.00–4.00) measuring 12 Canva infographic posters produced by mentors and 71 reflective peace essays written by Grade 7 students.

2.3 Intervention Procedures

The experimental intervention was executed across four sequential phases:

- a. Stage I (Transfer of Insight): The supervising instructor trained the three student co-researchers in moderation theology, digital manipulation typologies, and FGD facilitation techniques.

¹² John A. Banas and Stephen A. Rains, "A Meta-Analysis of Research on Inoculation Theory," *Communication Monographs* 77, no. 3 (2010): 281–311, <https://doi.org/10.1080/03637751003758193>; Josh Compton, Shelley Wigley, and Sergei A. Samoilenko, "Inoculation Theory and Public Relations," *Public Relations Review* 47, no. 5 (2021): 102116, <https://doi.org/10.1016/j.pubrev.2021.102116>.

¹³ Ahmad Ridwan and Ansor Feri Mahmudi, "Internalisasi Nilai-Nilai Wasathiyah Melalui Pendidikan Berbasis Pesantren," *Journal of Educational Research and Practice* 1, no. 1 (2023): 30–45; Siti Juhaeriyah, "Internalisasi Nilai Moderasi Beragama Sebagai Upaya Pencegahan Radikalisme Pada Santri Pondok Pesantren Al-Qur'an Ath-Thabranayah," *Al-Hikmah: Jurnal Studi Agama-Agama* 8, no. 1 (2022): 131–36.

¹⁴ Zakariyah, Muhid, and Arifin, "Transformation of Al-Qur'an Memorization Learning in the Digital Age through the Blended Learning Flipped Classroom Approach."

- b. Stage II (Flipped Inoculation Training): In the asynchronous phase, the 12 mentors independently analyzed video case studies containing weakened extremist appeals. In the synchronous phase, student co-researchers facilitated structured FGDs, guiding mentors to formulate refutational pre-emptions anchored in *tawasuth* and *tabayyun* (verification).
- c. Stage III (Post-Test and Creative Production): Immediately following post-test evaluations, mentors consolidated their critical digital identity by designing digital peace and civility infographics using Canva.
- d. Stage IV (Peer Mentoring Dissemination): The 12 mentors deployed their infographics within small mentoring circles (1 mentor per 5–6 mentees) to guide the 71 Grade 7 students, culminating in an independent peace literacy essay assignment.

2.4 Data Collection and Analysis

Data collection utilized three validated instruments: (1) the Cognitive Immunity and Digital Identity Questionnaire (10 contextual situational items); (2) the Canva Poster Assessment Rubric (theological substance, digital etiquette, and visual appeal); and (3) the Literacy Content Analysis Rubric (threat identification, *tawasuth* reasoning, and counter-narrative formulation). We analyzed quantitative data using IBM SPSS Statistics 26 with paired-sample t-tests ($\alpha=0.05$), 95% Confidence Intervals, Cohen's d effect size, and Hake's Normalized Gain (N-Gain). We analyzed qualitative artifacts using descriptive portfolio evaluation.

3. Results and Discussion

3.1. Baseline Cognitive Immunity Profile (Pre-test)

Initial diagnostic evaluations of the 12 student mentors revealed that while their overall baseline moderation scores were comparatively high (averaging 87.5%), specific indicators exhibited critical vulnerabilities. Table 1 summarizes the distribution of pre-test competencies across the assessment items.

Table 1. Distribution of Mentors' Pre-test Cognitive Immunity Scores

Competency Category	Case Item Codes	Correct Responses (n)	Accuracy (%)
Moderate Consensus	S3, S9	12	100.0%
Dominant Moderation	S2, S5, S8	11	91.6%
Moderate Baseline	S1, S6, S7, S10	10	83.3%
Critical Vulnerability	S4	8	66.7%

Source: Processed empirical research data (2026)

As detailed in Table 1, normative theological scenarios regarding interfaith harmony and basic tolerance (S3 and S9) achieved unanimous accuracy (100.0%). This ceiling effect demonstrates that the institutionalization of traditional *wasathiyyah* values transmitted through classical Islamic texts (*turats*) and the spiritual exemplarity (*uswah*) of the kyai has successfully ingrained fundamental theological orthodoxy in students' worldview.¹⁵ In sharp contrast, item S4, which

¹⁵ Zainul Abas et al., "Guardian of Faith: Comparative Insights of Teaching Islamic Theology in Indonesian and Malaysian Pesantren," *Nazhruna: Jurnal Pendidikan Islam* 8, no. 2 (2025): 382–99;

simulated exploitative cyber-fundraising for clandestine militant causes disguised as humanitarian crisis relief, yielded the lowest baseline accuracy at only 66.7%.

This diagnostic discrepancy highlights a profound socioreligious paradox within contemporary Islamic education. While santri exhibit steadfast doctrinal allegiance in physical, offline environments, their cognitive architecture remains acutely vulnerable to the nuanced persuasion strategies deployed across digital ecosystems.¹⁶ The sharp decline observed in item S4 illustrates the psychological mechanism of online emotional contagion: Generation Z adolescents are inherently susceptible to emotive imagery depicting human suffering, especially when combined with decontextualized religious imperatives that frame financial contributions as an unmediated moral obligation. In such affectively charged situations, immediate emotional reactions frequently bypass standard theological reflection and religious verification protocols (*tabayyun*).

These baseline findings substantiate recent empirical critiques regarding institutional vulnerabilities and digital degradation within Islamic boarding schools.¹⁷ Pesantren often cultivate strong normative religiosity without systematically fostering critical media literacy or an understanding of algorithmic mechanisms like filter bubbles and echo chambers.¹⁸ When digital literacy is treated as a purely instrumental proficiency such as technical device handling or basic graphic design it fails to cultivate the critical digital identity and ethical resilience required to interrogate manipulative online content.¹⁹ Consequently, students remain susceptible to moral desensitization and passive radicalization in cyberspace.²⁰ The baseline deficit recorded in Table 1 thus provided the empirical justification for the Ideological Vaccination intervention: pre-emptively exposing santri to controlled, weakened doses of deceptive rhetoric to trigger cognitive resistance before genuine online exposure.

3.2. Model Efficacy: Inferential Statistics and Normalized Gain Analysis

Following the flipped inoculation workshops, post-test assessments measured changes in mentors' cognitive resistance across all scenario items. Figure 1 illustrates comparative performance across specific assessment cases.

Muhammad Anas Ma'arif et al., "Model of Kiai Leadership Based on Local Wisdom: Preventing Radicalism and Building Education in the Global South," *Kharisma: Jurnal Administrasi Dan Manajemen Pendidikan* 5, no. 1 (2026); Ahmad Ridwan and Ansor Feri Mahmudi, "Internalisasi Nilai-Nilai Wasathiyah Melalui Pendidikan Berbasis Pesantren," *Journal of Educational Research and Practice* 1, no. 1 (2023): 30–45.

¹⁶ Amrullah Amrullah, "Peran Pesantren Dalam Menangkal Radikalisme Digital: Studi Terhadap Santri Era Media Sosial," *MODELING: Jurnal Program Studi PGMI* 12, no. 2 (2025): 183–96; Olga Sancaya Dyah Permatasari, M. Hanita, and H. Purwanto, "Strategi Internalisasi Nilai Pancasila Pada Generasi Milenial," *Responsive* 6, no. 2 (2023): 101–18.

¹⁷ Zawaqi Afdal Jamil et al., "Program Evaluation Study on Islamic Religious Education in Pesantren: Addressing Educational Degradation in the Digital Era," *Jurnal Pendidikan Agama Islam* 22, no. 1 (2025): 122–39.

¹⁸ Rahmadini et al., "Penguatan Moderasi Beragama Melalui Literasi Digital Dan Pemahaman Kritis Atas Algoritma Media Sosial Bagi Para Santri Di Pondok Pesantren Belitung."

¹⁹ Gufran Darma Dirawan, Dyah Darma Andayani, and Nurlita Pertiwi, "Peningkatan Literasi Digital Pada Santri Pesantren Syeh Hasan Yamani Memanfaatkan Aplikasi Canva," *Jurnal Masyarakat Madani Indonesia* 2, no. 4 (2023): 366–271.

²⁰ Harianto et al., "Algorithmic Spiritual Authority."

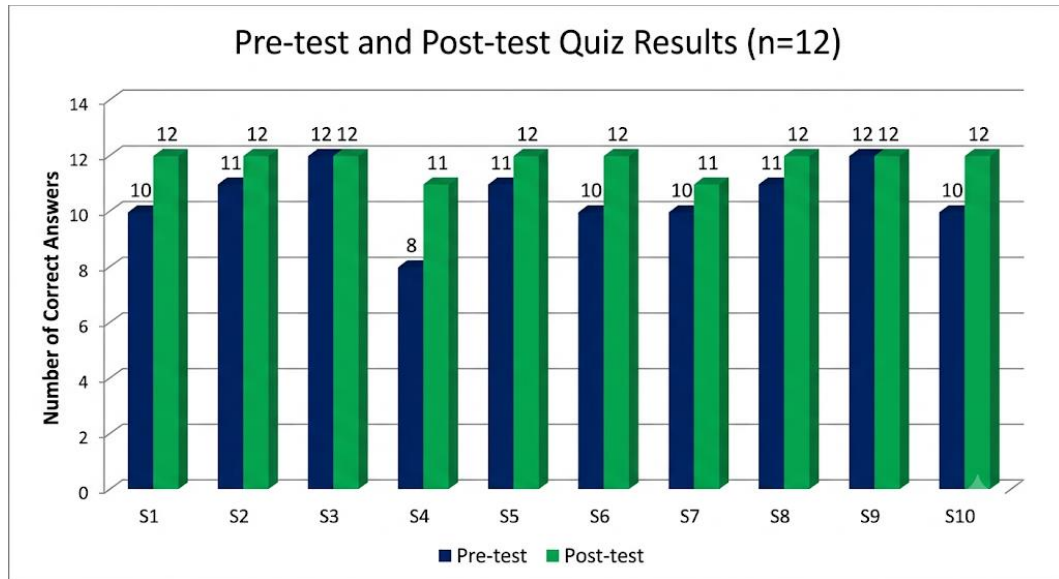


Figure 1. Comparison of Mentors' Pre-test vs. Post-test Performance by Item

To statistically substantiate the growth shown in Figure 1, the paired datasets underwent inferential testing and gain-index evaluations, as shown in Table 2. Before conducting inferential analyses, we examined the normality assumption using the Shapiro-Wilk test, which is well suited to small sample sizes ($n = 12$). The test confirmed that the distribution of pre-test scores ($W = 0.932$, $p = 0.402$), post-test scores ($W = 0.915$, $p = 0.245$), and their paired differences ($W = 0.941$, $p = 0.512$) did not deviate significantly from normality ($p > 0.05$), thereby fully validating the application of parametric paired-sample t-tests.

Table 2. Paired-Sample t-Test and Normalized Gain (N-Gain) Results

Assessment Phase	Mean Score (%)	Standard Deviation	t-value	Degrees of Freedom (df)	Sig. (p-value)	N-Gain Score	Efficacy Level
Pre-test	87.50	4.52	-6.842	11	0.0001	0.83	High
Post-test	98.33	2.46					

Source: Statistical output generated via SPSS 26 (2026)

As delineated in Table 2, the paired-sample t-test revealed a statistically significant increase in mentors' cognitive immunity following the intervention, shifting the mean score from 87.50% (SD = 4.52) to 98.33% (SD = 2.46), yielding a substantial mean gain of 10.83% ($t(11) = -6.842$, $p = 0.0001$, $p < 0.001$, two-tailed). The 95% confidence interval for the mean difference ranged from 7.35% to 14.31%, confirming that the true population gain reliably excludes zero with high precision. To evaluate the practical magnitude of the intervention beyond statistical significance, we also computed a standardized effect size. The resulting Cohen's d of 1.98 substantially exceeds the conventional benchmark for a "large" effect ($d \geq 0.80$), demonstrating that the "Ideological Vaccination" model generated an

exceptionally robust instructional and psychological impact. Concurrently, the calculated average Normalized Gain (N-Gain) index reached 0.83 (SD = 0.11), firmly categorizing the intervention within the "High Efficacy" threshold ($g \geq 0.70$) according to Hake's criteria.

A deeper diagnostic examination of the data reveals two crucial pedagogical dynamics: First, a noteworthy structural finding is the marked reduction in standard deviation from pre-test (SD = 4.52) to post-test (SD = 2.46). This variance contraction shows a pronounced homogenization of cognitive resilience across the cohort; the intervention not only raised the group average but also closed the intra-cohort competence gap, elevating previously struggling mentors to the upper mastery tier. Second, the quantitative breakthrough was anchored by the resolution of critical epistemic blind spots. As evidenced by item S4 (exploitative jihad donation appeals), accuracy surged from 66.7% (8 mentors) to 91.6% (11 mentors). Compared with previous empirical interventions in pesantren that relied primarily on technical counter-narrative creation without inoculation principles,²¹ participants in this study showed significantly higher analytical discernment in unmasking manipulative fundraising indicators, such as unverified private bank accounts and emotionally coercive imagery.²² Moreover, the observed performance gains across unpracticed scenario items (S1, S6, S7, and S10) empirically substantiate the emergence of generalized cognitive resistance (refutational difference).²³ Rather than conditioning static, rote rebuttals to identical stimuli, the flipped inoculation architecture restructured the mentors' underlying analytical habits, embedding *tawasuth* (moderation) as an autonomous mental algorithm (*manhaj al-fikr*) for vetting diverse and emergent online extremist narratives.²⁴

3.3. Inoculation Mechanisms: Encouraging Cognitive Antibodies through *Tawasuth*

The empirical success of the "Ideological Vaccination" intervention is structurally rooted in the foundational tenets of William J. McGuire's Inoculation Theory,²⁵ operating through a sequential, dual-phase cognitive mechanism: threat induction and refutational pre-emption.²⁶

- a. Threat Induction (Affective Realization): Rather than treating religious moderation as a static doctrinal given, the intervention deliberately engineered an affective disequilibrium by exposing the 12 student mentors to their own epistemic vulnerabilities within algorithmic media ecosystems.

²¹ Mustopa, Nawawi Fuad, and Bisri, "Edukasi Kontra Narasi Intoleran Dan Radikalisme Melalui Literasi Media Online Kepada Santri Pondok Pesantren Annida Cirebon," *Abdimas Galuh* 5, no. September (2023): 1026–36.

²² M Nurul Ikhsan Saleh, "Perceptions of Pesantren Leaders towards Islamic Moderation Approaches in Combating Radicalism and Terrorism," *Cogent Arts & Humanities* 12, no. 1 (2025), <https://doi.org/10.1080/23311983.2025.2474826>.

²³ John A Banas and Stephen A Rains, "A Meta-Analysis of Research on Inoculation Theory," *Communication Monographs* 77, no. 3 (2010): 281–311, <https://doi.org/10.1080/03637751003758193>.

²⁴ Lilis Nurhayati and Abdul Karim, "Integrating Digital Citizenship and Religious Moderation in Open and Distance Education : A Holistic Approach," *Asian Association of Open Universities Journal* 20, no. 3 (2026): 261–76, <https://doi.org/10.1108/AAOUJ-02-2025-0025>.

²⁵ William J. McGuire, "Resistance to Persuasion Conferred by Active and Passive Prior Refutation of the Same and Alternative Counterarguments," *The Journal of Abnormal and Social Psychology* 63, no. 2 (1961): 326–32; William J. McGuire, "Inducing Resistance to Persuasion: Some Contemporary Approaches," *Advances in Experimental Social Psychology* 1 (1964): 191–229.

²⁶ Josh Compton, Shelley Wigley, and Sergei A Samoilenko, "Inoculation Theory and Public Relations," *Public Relations Review* 47, no. 5 (2021): 102116, <https://doi.org/https://doi.org/10.1016/j.pubrev.2021.102116>.

Through structured diagnostic reflection, mentors came to recognize that filter bubbles, echo chambers, and emotionally charged disinformation cascades could easily circumvent their unexamined adherence to *tawasuth*. In primary inoculation literature, this subjective acknowledgment of vulnerability termed perceived threat serves as the indispensable motivational catalyst: it triggers psychological defense mechanisms. It motivates individuals to fortify their beliefs against future persuasive incursions actively.²⁷

- b. Refutational Pre-emption (Cognitive Scaffolding through Flipped Discourse): The flipped classroom model provided the dialectical framework required to turn affective vigilance into functional resistance. During the asynchronous phase, mentors were exposed to carefully curated, attenuated doses of radical discourse (e.g., weaponized Qur'anic texts on jihad, sectarian takfiri propaganda, and unregulated online charity appeals). In subsequent synchronous classroom sessions, student co-researchers facilitated Focus Group Discussions (FGDs) in which mentors systematically dissected the logical fallacies, manipulative framing, and emotional exploitation underlying these messages. Grounded in the interpretive ethics of *tawasuth* (moderation), *tawazun* (balance), and *i'tidal* (justice), the mentors formulated counter-arguments and theological rebuttals before encountering uncensored online extremist narratives.²⁸

This pedagogical intervention transformed the mentors from passive knowers of dogma into active practitioners of verification (*tabayyun*). In contrast to conventional pesantren learning where religious moderation is taught through passive memorization,²⁹ inoculation transformed moderation into an operative cognitive filter. Crucially, the marked post-test improvements across assessment items that were not explicitly addressed during the training sessions confirm the presence of generalized cognitive resistance (refutational difference). Inoculation did not merely condition rote resistance to specific arguments; it broadened participants' cognitive schema, equipping them with the epistemological flexibility to identify, evaluate, and counter novel variants of cyber-extremism across diverse online contexts.

3.4. Behavioral Transformation: Creative Content Production and Dissemination

The cultivation of cognitive resistance did not remain confined to test performance; it translated directly into active, prosocial agency within the digital sphere, supported by robust empirical portfolio data. By shifting from passive information consumers to active "peer inoculators," the 12 student mentors designed original educational infographics in Canva, synthesizing critical digital literacy with normative religious ethics.³⁰ Independent expert evaluation of these digital artifacts yielded a mean score of 3.95 out of 4.00, demonstrating theological fidelity, contextual salience, and nuanced visual communication tailored to youth culture. This creative production served as an external manifestation of their

²⁷ Fransen et al., "Sixty Years Later: A Replication Study of McGuire's First Inoculation Experiment."

²⁸ Siti Juhaeriyah, "Internalisasi Nilai Moderasi Beragama Sebagai Upaya Pencegahan Radikalisme Pada Santri Pondok Pesantren Al-Qur'an Ath-Thabraniyah," *Al-Hikmah: Jurnal Studi Agama-Agama* 8, no. 1 (2022): 131–36.

²⁹ Nurhayati Ali and Imar Putra, "From Textualism to Contextualism : A New Direction for the Reform of the Islamic Education Curriculum in Indonesia," *Khulasah Islamic Studies Journal* 08, no. 02 (2026): 83–99, <https://doi.org/10.55656/ksij.v8i2.654>.

³⁰ Dirawan, Andayani, and Pertiwi, "Peningkatan Literasi Digital Pada Santri Pesantren Syeh Hasan Yamani Memanfaatkan Aplikasi Canva."

consolidated digital literacy identities.³¹ Rather than merely retreating from harmful online material, the mentors actively constructed persuasive counter-narratives to contest polarizing ideologies in open cyberspace.

The systemic ripple effect of this intervention was empirically evidenced during the subsequent peer-led dissemination phase involving the 71 seventh-grade students. To empirically substantiate the depth of value internalization among this impact cohort, their reflective peace essays were evaluated across four core analytical dimensions, as summarized in Table 3.

Table 3. Literacy Portfolio Evaluation for Seventh-Grade Impact Group (N=71)

Assessment Indicators	Mean Score (Scale 1–4)	Qualitative Classification
Threat Identification (<i>Perceived Threat</i>)	3.90	Excellent
Internalization of <i>Tawasuth</i> Logic	3.92	Excellent
Counter-Narrative Production	3.94	Excellent
Actionable Creative Solutions	3.84	Excellent
Collective Aggregate Mean	3.90	Excellent

Source: Processed portfolio evaluation data (2026)

As delineated in Table 3, the impact cohort attained an aggregate mean score of 3.90 (categorized as Excellent), led by their performance in Counter-Narrative Production (3.94) and Internalization of *Tawasuth* Logic (3.92). These empirical figures indicate that the transmission of moderation principles was not diluted across the cascading hierarchy. Instead, horizontal peer mentoring dismantled the affective resistance and authoritative detachment often associated with formal didactic instruction.³² Mentors translated complex theological hermeneutics into relatable vernacular dialogue, allowing seventh-grade students to perceive digital discernment not as an external institutional mandate, but as an essential personal ethic.

This internalized perspective was vividly captured in a seventh-grade student's reflective essay, *Friends in Diversity*: "Truth does not require speed; it requires scrutiny." This reflection encapsulates an authentic adoption of the Quranic principle of *tabayyun* over online impulsivity and clickbait-driven validation.³³ Rather than becoming ensnared by viral polarization, the younger students demonstrated the empirical capacity to deconstruct affective contagion,

³¹ Burhanudin Jauhari, M Yunan Hidayat, and Sukari, "Pengaruh Literasi Budaya, Literasi Digital Dan Kesadaran Beragama Terhadap Moderasi Beragama Siswa," *Al-Mausu'ah: Jurnal Studi Islam* 5, no. 7 (2024): 2024.

³² Laura Gehreke, Hannes Schilling, and Simone Kauffeld, "Effectiveness of Peer Mentoring in the Study Entry Phase: A Systematic Review," *Review Education* 12, no. 1 (2024): 1–32, <https://doi.org/10.1002/rev3.3462>.

³³ Siti Nurrahmatin Ni'mah, "Peran Guru Pesantren Dalam Menanamkan Nilai-Nilai Moderasi Islam Pada Santri," *Psikosopen* 2, no. 1 (2026): 863–74.

formulating coherent counter-narratives that reposition the digital realm as a space for civic peace and inclusive dialogue.³⁴

4. Conclusion

This study confirms that the "Ideological Vaccination" model, operationalized through an integrated flipped classroom and cascading peer-mentoring framework, provides an effective, sovereign pedagogical defense against digital radicalism within Islamic boarding schools. To ensure theoretical and empirical alignment, the conclusions are synthesized across the study's three core research objectives: First, regarding empirical efficacy (Objective 1), inferential statistical analysis confirmed that the intervention generated a significant and substantial increase in mentors' cognitive immunity, elevating the mean score from 87.50% at pre-test to 98.33% at post-test ($t(11) = -6.842$, $p = 0.0001$, 95% CI [7.35, 14.31]). The high Normalized Gain score (N-Gain = 0.83), alongside an exceptional effect size (Cohen's $d = 1.98$), demonstrates strong pedagogical effectiveness. Crucially, the intervention successfully remediated the cohort's most critical baseline vulnerability—exploitative jihad donation scams (item S4)—surging from 66.7% to 91.6%, while simultaneously fostering generalized cognitive resistance across insufficiently trained extremist scenarios.

Second, regarding the operationalization of moderation (Objective 2), the flipped peer-mentoring architecture successfully activated *tawasuth* (moderation), *tawazun* (balance), and *tabayyun* (verification) as an operative cognitive heuristic (*manhaj al-fikr*). Rather than functioning as abstract theological dogma, pre-emptive exposure to weakened extremist rhetoric (threat induction) coupled with peer-guided counter-framing (refutational pre-emption) empowered santri to actively interrogate manipulative cyber-narratives and neutralize online affective contagion before uncured exposure occurs.

Third, regarding behavioral transformation (Objective 3), the intervention successfully catalyzed santri from passive consumers of digital content into sovereign, proactive producers of moderate religious discourse. This behavioral antibody was empirically substantiated by the production of 12 high-quality digital civility infographics created through Canva by the mentors (mean score 3.95/4.00) and 71 reflective peace literacy essays composed by the Grade 7 impact cohort (mean score 3.90/4.00), demonstrating that moderation values were deeply internalized across peer hierarchies. Collectively, these findings emphasize the critical necessity for Islamic educational institutions to shift digital literacy frameworks from mere technical-instrumental proficiency toward values-based cognitive defense mechanisms. Despite these robust outcomes, this study is constrained by its single-institution, one-group quasi-experimental design and the absence of longitudinal tracking. Future investigations should employ randomized controlled trial (RCT) designs with external control groups across diverse pesantren typologies (*salaf* and modern), while systematically tracking the longitudinal retention of cognitive resistance against rapidly evolving generative AI-driven extremist manipulation tactics.

References

Abas, Zainul, Anas Tajudin, Arif Nugroho, and Hedy Ramadhan Putra. "Guardian of Faith: Comparative Insights of Teaching Islamic Theology in Indonesian

³⁴ Qurroh A'yuniyah, Akhmad Mahfud, and Imamul Muttakin, "Islam Nusantara Dalam Kerukunan Umat Beragama," *Karakter: Jurnal Riset Ilmu Pendidikan Islam* 2, no. 2 (January 2, 2025): 259–68, <https://doi.org/10.61132/karakter.v2i2.592>.

- and Malaysian Pesantren.” *Nazhruna: Jurnal Pendidikan Islam* 8, no. 2 (2025): 382–99. <https://doi.org/10.31538/nzh.v8i2.5645>.
- Ahmad Afnan, Anshori, and Pohl, Florian. “Environmental Education and Indonesia’s Traditional Islamic Boarding Schools: Curricular and Pedagogical Innovation in the Green Pesantren Initiative.” *Supporting Modern Teaching in Islamic Schools* 30, no. 3 (2025): 3276–87.
- Ali, Nurhayati, and Imar Putra. “From Textualism to Contextualism: A New Direction for the Reform of the Islamic Education Curriculum in Indonesia.” *Khulasah Islamic Studies Journal* 8, no. 2 (2026): 83–99. <https://doi.org/10.55656/ksij.v8i2.654>.
- Amrullah, Amrullah. “Peran Pesantren Dalam Menangkal Radikalisme Digital: Studi Terhadap Santri Era Media Sosial.” *MODELING: Jurnal Program Studi PGMI* 12, no. 2 (2025): 183–96. <https://doi.org/10.69896/modeling.v12i2.2934>.
- Anantasia, Gisela, and Sulastri Rini Rindrayani. “Metodologi Penelitian Quasi Eksperimen.” *ADIBA: Journal of Education* 5, no. 2 (2025): 183–92.
- Banas, John A., and Stephen A. Rains. “A Meta-Analysis of Research on Inoculation Theory.” *Communication Monographs* 77, no. 3 (2010): 281–311. <https://doi.org/10.1080/03637751003758193>.
- Braun, Virginia, and Victoria Clarke. “Using Thematic Analysis in Psychology.” *Qualitative Research in Psychology* 3, no. 2 (2006): 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
- Compton, Josh, Shelley Wigley, and Sergei A. Samoilenko. “Inoculation Theory and Public Relations.” *Public Relations Review* 47, no. 5 (2021): 102116. <https://doi.org/10.1016/j.pubrev.2021.102116>.
- Dirawan, Gufran Darma, Dyah Darma Andayani, and Nurlita Pertiwi. “Peningkatan Literasi Digital Pada Santri Pesantren Syekh Hasan Yamani Melalui Pemanfaatan Aplikasi Canva.” *Jurnal Masyarakat Madani Indonesia* 2, no. 4 (2023): 366–71.
- Fransen, Marieke L., Saar Mollen, Stephan A. Rains, Enny Das, and Ivar E. Vermeulen. “Sixty Years Later: A Replication Study of McGuire’s First Inoculation Experiment.” *Journal of Media Psychology: Theories, Methods, and Applications* 36 (2023): 69–78. <https://api.semanticscholar.org/CorpusID:261843636>.
- Gehreke, Laura, Hannes Schilling, and Simone Kauffeld. “Effectiveness of Peer Mentoring in the Study Entry Phase: A Systematic Review.” *Review of Education* 12, no. 1 (2024): 1–32. <https://doi.org/10.1002/rev3.3462>.
- Harianto, Eko, Reni Dian Lestari, Try Riduwan Santoso, and Jamaludin. “Algorithmic Spiritual Authority.” *ESENSIA: Jurnal Ilmu-Ilmu Ushuluddin* 26, no. 2 (2025): 96–108. <https://doi.org/10.14421/esensia.v26i2.7772>.
- Ja’far, Ali. “Literasi Digital Pesantren: Perubahan dan Kontestasi.” *Islamic Review: Jurnal Riset dan Kajian Keislaman* 8, no. 1 (2019): 17–35. <https://doi.org/10.35878/islamicreview.v8i1.156>.
- Jamil, Zawaqi Afdal, Nurhuda, Alfian, Siti Mariah Ulfah, Rasidin, and Ahmad Youssef. “Program Evaluation Study on Islamic Religious Education in Pesantren: Addressing Educational Degradation in the Digital Era.” *Jurnal Pendidikan Agama Islam* 22, no. 1 (2025): 122–39. <https://doi.org/10.14421/jpai.v22i1.11358>.

- Jauhari, Burhanudin, M. Yunan Hidayat, and Sukari. "Pengaruh Literasi Budaya, Literasi Digital, Dan Kesadaran Beragama Terhadap Moderasi Beragama Siswa." *Al-Mausu'ah: Jurnal Studi Islam* 5, no. 7 (2024): 2024.
- Juhaeriyah, Siti. "Internalisasi Nilai Moderasi Beragama sebagai Upaya Pencegahan Radikalisme pada Santri Pondok Pesantren Al-Qur'an Ath-Thabraniyah." *Al-Hikmah: Jurnal Studi Agama-Agama* 8, no. 1 (2022): 131–36.
- Lohlker, Rüdiger, and Soleh Hasan Wahid. "The Transformation of Islamic Religious Authority." *Religions* 17, no. 4 (2026): 493. <https://doi.org/10.3390/rel17040493>.
- Ma'arif, Muhammad Anas, Muhamad Arif, Mauhibur Rokhman, Ali Usman Hali, Ari Kartiko, and Akhmad Sirojuddin. "Model of Kiai Leadership Based on Local Wisdom: Preventing Radicalism and Building Education in the Global South." *Kharisma: Jurnal Administrasi dan Manajemen Pendidikan* 5, no. 1 (2026). <https://doi.org/10.59373/kharisma.v5i1.149>.
- McGuire, William J. "Inducing Resistance to Persuasion: Some Contemporary Approaches." *Advances in Experimental Social Psychology* 1 (1964): 191–229. [https://doi.org/10.1016/S0065-2601\(08\)60052-0](https://doi.org/10.1016/S0065-2601(08)60052-0).
- McGuire, William J. "Resistance to Persuasion Conferred by Active and Passive Prior Refutation of the Same and Alternative Counterarguments." *The Journal of Abnormal and Social Psychology* 63, no. 2 (1961): 326–32. <https://doi.org/10.1037/h0048344>.
- Moustakas, Clark. *Phenomenological Research Methods*. Thousand Oaks: SAGE Publications, Inc., 2011. <https://doi.org/10.4135/9781412995658>.
- Mustopa, Nawawi Fuad, and Bisri. "Edukasi Kontra Narasi Intoleran Dan Radikalisme Melalui Literasi Media Online Kepada Santri Pondok Pesantren Annida Cirebon." *Abdimas Galuh* 5, no. 2 (2023): 1026–36.
- Ni'mah, Siti Nurrahmatin. "Peran Guru Pesantren Dalam Menanamkan Nilai-Nilai Moderasi Islam Pada Santri." *Psikosopen* 2, no. 1 (2026): 863–74.
- Nurhayati, Lilis, and Abdul Karim. "Integrating Digital Citizenship and Religious Moderation in Open and Distance Education: A Holistic Approach." *Asian Association of Open Universities Journal* 20, no. 3 (2026): 261–76. <https://doi.org/10.1108/AAOUJ-02-2025-0025>.
- Nurhayati, N., and L. Lumiaty. "Transforming Islamic Elementary Education: The Role of Curriculum Management in Quality Improvement." *Journal of Social Studies and Education* 3, no. 2 (2025): 115–31. <https://doi.org/10.61987/jsse.v3i2.1293>.
- Osigwe, Obinna, Ntalo Nasulu Hussein, and Asiimwe Specioza. "Digital Influence and Cultural Identity: A Narrative Review on How Online Content Shapes Cultural Orientation among Ugandan Youth." *International Journal of Research and Innovation in Social Science* 10, no. 4 (2026): 2565–77. <https://doi.org/10.47772/ijriss.2026.100400196>.
- Permatasari, Olga Sancaya Dyah, M. Hanita, and H. Purwanto. "Strategi Internalisasi Nilai-Nilai Pancasila Pada Generasi Milenial." *Responsive* 6, no. 2 (2023): 101–18. <https://doi.org/10.24198/responsive.v6i2.46659>.
- Qurroh A'yuniyah, Akhmad Mahfud, and Imamul Muttakin. "Islam Nusantara Dalam Kerukunan Umat Beragama." *Karakter: Jurnal Riset Ilmu Pendidikan Islam* 2, no. 2 (2025): 259–68. <https://doi.org/10.61132/karakter.v2i2.592>.

- Rahmadini, Rahayu, Jauharotul Maknunah, and Serly Arsita. "Penguatan Moderasi Beragama Melalui Literasi Digital Dan Pemahaman Kritis Atas Algoritma Media Sosial Bagi Para Santri Di Pondok Pesantren Belitung." *Journal of Community Development* 6, no. 3 (2026): 1–12.
- Ridwan, Ahmad, and Ansor Feri Mahmudi. "Internalisasi Nilai-Nilai Wasathiyah Melalui Pendidikan Berbasis Pesantren." *Journal of Educational Research and Practice* 1, no. 1 (2023): 30–45. <https://doi.org/10.70376/jerp.v1i1.30>.
- Saleh, M. Nurul Ikhsan. "Perceptions of Pesantren Leaders towards Islamic Moderation Approaches in Combating Radicalism and Terrorism." *Cogent Arts & Humanities* 12, no. 1 (2025): 2474826. <https://doi.org/10.1080/23311983.2025.2474826>.
- Setyowati, Eko Mety, Ahmad Happy Nashrullah, Syafrozi Haqi, Wishnu Wardana, and Bagus Shandy Narmaditya. "Penggunaan Teknologi Artificial Intelligence (AI) dalam Pendidikan Bisnis (Analisis Tinjauan Literatur Sistematis)." *JISPENDIORA: Jurnal Ilmu Sosial, Pendidikan dan Humaniora* 4, no. 2 (2025): 308–29.
- Zakariyah, Zakariyah, Abdul Muhid, and Miftah Arifin. "Transformation of Al-Qur'an Memorization Learning in the Digital Age through the Blended Learning Flipped Classroom Approach." *Nazhruna: Jurnal Pendidikan Islam* 8, no. 3 (2025): 594–606. <https://doi.org/10.31538/nzh.v8i3.193>.