



Systematic Literature Review terhadap Evaluasi Sumatif dan Model Pembelajaran dalam Meningkatkan Prestasi Belajar Akidah Akhlak

Idhar Ciwintara¹⁾, Mahmudin Sudin²⁾

Universitas Muhammadiyah Jakarta, Ciputat, Tangerang Selatan, Indonesia

ciwintaraidhar@gmail.com¹⁾

mahmudinsudin@umj.ac.id²⁾

Abstrak

Ketidakselarasan antara inovasi pedagogi dan instrumen evaluasi sering mendistorsi validitas prestasi belajar siswa pada mata pelajaran Akidah Akhlak. Penelitian ini bertujuan untuk memetakan tren integrasi antara model pembelajaran dan evaluasi sumatif serta menganalisis relevansi konstruktif keduanya dalam meningkatkan hasil belajar. Metode yang diterapkan adalah *Systematic Literature Review* (SLR) dengan menggunakan protokol PRISMA untuk menyeleksi literatur. Data bersumber dari 20 artikel jurnal terakreditasi nasional yang diterbitkan dalam rentang waktu 2014–2024, kemudian dianalisis menggunakan teknik analisis isi kualitatif. Hasil penelitian menunjukkan bahwa model pembelajaran kooperatif tipe *Jigsaw*, STAD, dan *Problem Based Learning* mendominasi tren intervensi pedagogis karena terbukti efektif meningkatkan partisipasi aktif siswa. Di sisi lain, instrumen evaluasi mulai bertransformasi menuju penggunaan *Higher Order Thinking Skills* (HOTS) dan analisis *Rasch Model*, meskipun dominasi evaluasi konvensional masih menghambat pengukuran kompetensi secara utuh. Temuan utama mengonfirmasi adanya kesenjangan konstruktif (*missing link*) di mana penerapan model pembelajaran aktif sering kali tidak diikuti dengan pembaruan instrumen penilaian yang setara. Penelitian ini menyimpulkan bahwa penyelarasan sistemik antara desain model pembelajaran dan instrumen evaluasi mutlak diperlukan untuk menjamin validitas mutu pendidikan karakter.

Kata kunci: Akidah Akhlak, Evaluasi Sumatif, Model Pembelajaran, Prestasi Belajar, *Systematic Literature Review*

Abstract

The misalignment between pedagogical innovation and evaluation instruments often distorts the validity of student learning achievement in Akidah Akhlak subjects. This study aims to map the integration trends between learning models and summative evaluation and to analyze the constructive relevance of both in improving learning outcomes. The method applied is a Systematic Literature Review (SLR) using the PRISMA protocol to select literature. Data were sourced from 20 nationally accredited journal articles published between 2014 and 2024, which were then analyzed using qualitative content analysis techniques. The results show that cooperative learning models such as Jigsaw, STAD, and Problem Based Learning dominate pedagogical intervention trends as they are proven effective in increasing active student participation. On the other hand, evaluation instruments are beginning to transform towards the use of Higher Order Thinking Skills (HOTS) and Rasch Model analysis, although the dominance of conventional evaluation still hinders holistic competence measurement. The main findings confirm a constructive gap (missing link) where the implementation of active learning models is often not followed by equivalent updates in assessment instruments. This study concludes that systemic alignment between learning model design and evaluation instruments is absolutely necessary to ensure the validity of character education quality.

Keywords: Akidah Akhlak, Learning Achievement, Learning Models, Summative Evaluation, *Systematic Literature Review*.



PENDAHULUAN

Pendidikan agama Islam memegang peranan sentral dan strategis dalam pembentukan karakter serta moralitas generasi bangsa di tengah arus globalisasi yang sarat dengan degradasi nilai etika. Dalam ekosistem kurikulum madrasah di Indonesia, mata pelajaran Akidah Akhlak berfungsi sebagai ujung tombak dalam internalisasi nilai-nilai ketuhanan (transendental) dan pembiasaan perilaku terpuji dalam kehidupan sosial (humaniora). Tujuan fundamental dari pembelajaran Akidah Akhlak sejatinya tidak berhenti pada transfer pengetahuan kognitif semata, melainkan harus menembus dimensi afektif dan psikomotorik yang bermuara pada terbentuknya kepribadian muslim yang kaffah. Namun, kompleksitas tujuan pembelajaran ini menuntut adanya sinkronisasi yang presisi antara model pembelajaran yang diterapkan di kelas dengan sistem evaluasi yang digunakan untuk mengukur keberhasilannya. Fenomena ketidakselarasan antara proses pembelajaran yang inovatif dengan alat ukur yang stagnan, atau sebaliknya, sering kali menjadi isu krusial yang mendistorsi validitas prestasi belajar siswa. Kebutuhan untuk menyelaraskan "cara mengajar" (pedagogi) dengan "cara mengukur" (asesmen) inilah yang menjadi landasan urgensi kajian ini.

State of the art atau tinjauan terhadap lanskap penelitian terdahulu dalam satu dekade terakhir menunjukkan adanya dua tren besar atau polarisasi fokus kajian dalam ranah pendidikan Akidah Akhlak. Tren pertama didominasi oleh studi eksperimental dan tindakan kelas yang secara masif menguji efektivitas berbagai varian model pembelajaran aktif (*active learning*) untuk mendongkrak hasil belajar. Kelompok penelitian ini sangat fokus pada intervensi pedagogis. Berdasarkan literatur yang dihimpun, model pembelajaran kooperatif menjadi primadona yang paling banyak diteliti. Studi dari Diza Jusriani & Ibrohim Muchlis (2023) serta penelitian Hakim & Komilah (2024) secara spesifik menyoroti keberhasilan model Kooperatif tipe *Jigsaw* dalam meningkatkan minat dan motivasi belajar siswa yang sebelumnya pasif. Varian kooperatif lainnya, yakni *Student Teams Achievement Division* (STAD), juga terbukti efektif sebagaimana dilaporkan dalam penelitian Putri et al. (2025) serta riset tindakan kelas oleh Zuraidah (2021), yang keduanya mengonfirmasi peningkatan hasil belajar yang signifikan pasca intervensi.

Selain model kooperatif, tren penggunaan model pembelajaran berbasis masalah dan aktivitas fisik juga mewarnai diskursus akademik terkini. Penelitian Al-Hasyim et al. (2024) menegaskan bahwa model *Problem Based Learning* (PBL) mampu mengontrol kategori hasil belajar siswa dari level rendah ke tingkat yang lebih tinggi melalui pemecahan masalah kontekstual. Sementara itu, metode yang lebih dinamis seperti *Snowball Throwing* juga mendapat perhatian, di mana Istijabah (2021) membuktikan peningkatannya pada materi Ilmu Kalam, dan Khotimah & Nurhasanah (2024) menemukan bahwa metode ini efektif untuk menstimulasi kemampuan berpikir kritis (*critical thinking*) siswa kelas 5 MI. Di sisi lain, pendekatan yang lebih personal dan visual seperti Mind Mapping diteliti oleh Widya Ayuningsih, dan model Pembelajaran Tuntas (*Mastery Learning*) dikaji oleh Afiyah & Sofa (2024) yang menekankan pada ketuntasan belajar individual. Secara lebih umum, Maulid et al. (2024) mengonfirmasi efektivitas strategi *Active Learning* secara luas di lingkungan pesantren. Tidak ketinggalan, Huda & Achadi (2024) memotret implementasi pembelajaran dalam konteks kebijakan terbaru Kurikulum Merdeka, sementara Asyifa et al. (2023) telah memulai upaya sintesis literatur namun terbatas pada peran Akidah Akhlak sebagai pondasi karakter abad 21, belum menyentuh aspek evaluasi secara mendalam.

Di seberang spektrum tersebut, tren kedua berfokus secara spesifik dan teknis pada pengembangan, analisis kualitas, serta validasi instrumen evaluasi. Para peneliti di klaster ini menaruh perhatian besar pada validitas alat ukur. Nurhayati et al. (2023) misalnya, melakukan analisis butir soal tes formatif untuk memastikan kualitas daya beda dan tingkat kesukaran soal. Pendekatan yang lebih canggih ditunjukkan oleh Ismi et al. (2023) yang menggunakan analisis *Rasch Model* untuk mendeteksi pola jawaban siswa dan validitas instrumen secara psikometrik yang ketat. Kesadaran untuk memodernisasi evaluasi agar tidak hanya berbasis hafalan (LOTS) menuju kemampuan berpikir tingkat tinggi terlihat dalam riset Latifah et al. (2023) yang mengembangkan instrumen berbasis *Higher Order Thinking Skills* (HOTS). Upaya untuk

menerapkan penilaian yang lebih holistik atau evaluasi autentik juga muncul dalam studi (Alfazri, 2024) di lingkungan pesantren, serta analisis instrumen penilaian spesifik di Madrasah Aliyah oleh Saputra et al. (2023). Pada level makro, evaluasi terhadap model dan proses pembelajaran itu sendiri telah dikaji oleh (Darodjat & Zuchdi, 2016) serta evaluasi pelaksanaan pembelajaran oleh Sohibah et al. (2023) dan Hamidi (2023).

Meskipun literatur mengenai model pembelajaran dan instrumen evaluasi tersedia cukup melimpah, terdapat kesenjangan atau gap analysis yang nyata dan menganga. Penelitian yang ada cenderung berjalan sendiri-sendiri pada jalurnya masing-masing (dikotomi). Studi tentang model pembelajaran (seperti Jigsaw, STAD, PBL, *Snowball Throwing*) mayoritas hanya fokus pada skenario pembelajaran dan hasil akhir berupa angka, tanpa membahas secara mendalam apakah instrumen evaluasi sumatif yang digunakan untuk mengukur keberhasilan tersebut sudah "relevan" atau selaras dengan modelnya. Seringkali, modelnya menuntut kekritisian (seperti PBL atau *Critical Thinking* pada riset Khotimah & Nurhasanah (2024)), namun alat ukurnya tidak dijelaskan apakah menggunakan tes standar atau instrumen HOTS. Sebaliknya, studi tentang pengembangan instrumen (seperti riset Latifah et al. (2023) atau Ismi et al. (2023)) sering kali sangat teknis psikometrik namun terlepas dari konteks model pembelajaran apa yang diterapkan di kelas sebelum evaluasi tersebut diberikan. Belum banyak studi yang berupaya mensintesis secara sistematis bagaimana keselarasan konstruktif (*constructive alignment*) antara varian model pembelajaran yang dipilih dengan jenis evaluasi sumatif yang digunakan dalam satu dekade terakhir.

Berangkat dari kesenjangan tersebut, artikel ini menawarkan kebaruan (*novelty*) berupa *Systematic Literature Review* (SLR) yang menghubungkan kedua domain yang terpisah tersebut. Penelitian ini tidak lagi bertujuan menguji satu model atau memvalidasi satu paket soal, melainkan memetakan tren integrasi antara model pembelajaran dengan evaluasi sumatif berdasarkan data dari 20 artikel terpilih tersebut. Kebaruan penelitian ini terletak pada upaya identifikasi pola hubungan: model pembelajaran apa saja yang paling dominan diterapkan dan jenis evaluasi apa yang menyertainya? Apakah pergeseran ke arah model student-centered (Kurikulum Merdeka) sudah diikuti dengan pergeseran ke arah evaluasi autentik dan HOTS? Penelitian ini akan menjawab pertanyaan tersebut dengan membedah literatur yang ada untuk menemukan praktik terbaik (*best practices*) mengenai pasangan model pembelajaran dan bentuk evaluasi sumatif yang paling relevan dan berdampak signifikan terhadap prestasi belajar Akidah Akhlak.

Permasalahan utama yang hendak dijawab dalam kajian sistematis ini dapat dirumuskan ke dalam pertanyaan penelitian mengenai peta distribusi model pembelajaran dan bentuk evaluasi sumatif dalam pembelajaran Akidah Akhlak, serta bagaimana relevansi keduanya mempengaruhi validitas capaian hasil belajar siswa. Hipotesis yang diajukan adalah bahwa penggunaan model pembelajaran aktif yang didukung oleh instrumen evaluasi sumatif yang relevan (seperti evaluasi autentik atau HOTS) memiliki dampak yang lebih signifikan terhadap prestasi belajar dibandingkan dengan penggunaan model aktif namun dengan evaluasi konvensional. Untuk menjawab hal tersebut, metode yang digunakan adalah SLR dengan protokol PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), yang menyeleksi dan menganalisis artikel-artikel terakreditasi nasional untuk diambil saripati temuannya. Tujuan akhir dari artikel ini adalah menyusun rekomendasi akademis bagi guru dan praktisi pendidikan dalam merancang desain pembelajaran yang koheren dari hulu (pilihan model) hingga hilir (penentuan evaluasi sumatif), sehingga pembelajaran Akidah Akhlak dapat mencapai tujuannya secara utuh dalam membentuk karakter siswa.

METODE

Penelitian ini menggunakan pendekatan kualitatif dengan desain SLR. Metode ini dipilih untuk mensintesis bukti-bukti penelitian yang terfragmentasi mengenai model pembelajaran dan evaluasi. Sebagaimana dijelaskan oleh Snyder (2019), SLR merupakan metode penelitian yang berdiri sendiri untuk mengidentifikasi, mengevaluasi, dan menafsirkan seluruh penelitian yang

tersedia yang relevan dengan pertanyaan penelitian tertentu. Pendekatan ini berbeda dengan tinjauan pustaka tradisional karena menerapkan protokol yang ketat dan transparan untuk meminimalisir bias peneliti (Kitchenham & Brereton, 2013). Paradigma interpretivisme digunakan sebagai landasan filosofis untuk memahami makna di balik data-data tekstual yang tersaji dalam artikel ilmiah, guna menarik kesimpulan mengenai relevansi antara intervensi pedagogis dan instrumen asesmen.

Sumber data dalam penelitian ini adalah data sekunder berupa artikel jurnal ilmiah bereputasi nasional dan prosiding konferensi yang diterbitkan dalam rentang waktu satu dekade terakhir (2014–2024). Populasi data mencakup seluruh literatur yang membahas pembelajaran Akidah Akhlak di madrasah dan pesantren. Melalui teknik *purposive sampling*, ditetapkan sebanyak 20 artikel terpilih sebagai unit analisis. Penentuan sampel ini didasarkan pada kriteria inklusi dan eksklusi yang ketat untuk menjamin kualitas data, sebagaimana disarankan oleh Xiao & Watson (2019), di mana artikel harus berupa *full-text*, berbahasa Indonesia atau Inggris, dan secara eksplisit membahas variabel model pembelajaran atau instrumen evaluasi Akidah Akhlak.

Prosedur pengumpulan data dilakukan dengan mengikuti tahapan standar protokol PRISMA edisi terbaru. Merujuk pada panduan Page et al. (2021), tahapan PRISMA meliputi: (1) *Identifikasi*, yaitu pencarian literatur pada basis data Google Scholar, Moraref, dan Sinta menggunakan kata kunci spesifik; (2) *Screening* (penyaringan), yaitu pemeriksaan relevansi judul dan abstrak; (3) *Eligibility* (kelayakan), yaitu penilaian kualitas teks lengkap; dan (4) *Included*, yaitu penetapan 20 artikel final (Kode A1-A8 dan B1-B12) yang valid. Instrumen utama dalam penelitian ini adalah peneliti sendiri (*human instrument*) yang dibantu dengan lembar matriks ekstraksi data.

Tabel 1.
Kriteria Inklusi dan Eksklusi Pencarian Literatur

Kriteria	Inklusi	Eksklusi
Topik	Artikel yang secara spesifik membahas (1) Model/Strategi Pembelajaran Akidah Akhlak, atau (2) Instrumen/Evaluasi Akidah Akhlak.	Artikel yang membahas Pendidikan Agama Islam (PAI) secara umum tanpa spesifikasi Akidah Akhlak.
Subjek	Jenjang Madrasah (MI, MTs, MA) dan Pesantren.	Jenjang Perguruan Tinggi atau Pendidikan Non-Formal non-pesantren.
Jenis Publikasi	Artikel Jurnal Ilmiah (<i>Research Article</i>) dan Prosiding Konferensi.	Tesis, Disertasi, Buku Ajar, Artikel Opini, atau Review Buku.
Bahasa	Bahasa Indonesia dan Bahasa Inggris.	Bahasa selain Indonesia dan Inggris.
Tahun Terbit	Rentang 10 tahun terakhir (2014–2024).	Diterbitkan sebelum tahun 2014.
Akses	Tersedia teks lengkap (<i>Full Text</i>) yang dapat diunduh.	Hanya tersedia Abstrak atau berbayar (terkunci).

Sumber: Hasil Penelitian, 2025

Teknik analisis data yang digunakan adalah analisis isi (*content analysis*) kualitatif interaktif yang dikembangkan oleh Miles et al. (2014), teknik ini terdiri dari tiga alur kegiatan yang terjadi secara bersamaan (*concurrent*), yaitu: (1) *Kondensasi data*, di mana peneliti melakukan seleksi, pemfokusan, dan pengabstrakan data dari 20 artikel terpilih; (2) *Penyajian data*, yaitu menyusun informasi yang telah terorganisir ke dalam tabel matriks dan narasi logis untuk melihat pola hubungan antar variabel; dan (3) *Penarikan kesimpulan/verifikasi*, yaitu memaknai temuan untuk menjawab rumusan masalah. Validitas temuan dijaga melalui teknik triangulasi sumber untuk



memastikan konsistensi data dari berbagai latar belakang artikel (Creswell, 2007), serta *audit trail* untuk mendokumentasikan jejak analisis data secara transparan.

HASIL DAN PEMBAHASAN

Berdasarkan proses ekstraksi data yang telah dirangkum dalam tabel sebelumnya, penelitian ini berhasil mengidentifikasi dan memetakan dua puluh artikel ilmiah yang relevan dengan topik integrasi model pembelajaran dan evaluasi sumatif pada mata pelajaran Akidah Akhlak. Temuan-temuan tersebut secara garis besar dapat diklasifikasikan ke dalam tiga tema utama yang saling beririsan, yaitu dominasi intervensi pedagogis melalui model pembelajaran aktif, transformasi instrumen evaluasi dari konvensional menuju modern, dan urgensi keselarasan konstruktif atau *constructive alignment* antara proses dan penilaian. Bagian ini akan menguraikan hasil temuan tersebut secara mendetail dan mendiskusikannya dalam perspektif teoritis yang lebih luas untuk menjawab rumusan masalah penelitian.

Tabel 2.
Data Ekstraksi Literatur Terpilih (n=20)

Kode	Penulis (Tahun)	Fokus	Metode	Hasil
A1	Nurhayati et al. (2023)	Analisis Butir Soal (Formatif).	Kuantitatif (Anates)	Mayoritas soal memiliki validitas sedang; perlu perbaikan pada daya beda soal.
A2	Darodjat & Zuchdi (2016)	Pengembangan Model Evaluasi.	R&D (Borg & Gall)	Menghasilkan model evaluasi logik untuk Akidah Akhlak di MTs yang valid dan reliabel.
A3	Saputra et al. (2023)	Analisis Instrumen Penilaian.	Kualitatif	Guru masih mendominasi penilaian kognitif; penilaian afektif belum maksimal.
A4	Ismi et al. (2023)	Analisis Pola Jawaban (Rasch Model).	Kuantitatif (Rasch)	Mendeteksi bias jawaban siswa; nilai tinggi belum tentu paham konsep (menebak).
A5	Latifah et al. (2023)	Instrumen Berbasis HOTS.	R&D	Instrumen HOTS meningkatkan kemampuan berpikir kritis (C4-C6) siswa SD.
A6	Sohibah et al. (2023)	Evaluasi Pelaksanaan Pembelajaran.	Kualitatif	Pelaksanaan pembelajaran sudah sesuai RPP, namun evaluasi masih konvensional.
A7	Alfazri (2024)	Evaluasi Autentik.	Kuantitatif	Penerapan evaluasi autentik efektif menilai sikap spiritual dan sosial di Pesantren.
A8	Hamidi (2023)	Evaluasi Proses	Kualitatif	Evaluasi proses



		Pembelajaran.	Deskriptif	
				belum maksimal karena keterbatasan waktu dan beban administrasi guru.
B1	Afiyah & Sofa (2024)	Model <i>Mastery Learning</i>	Kuantitatif	<i>Mastery Learning</i> berpengaruh signifikan terhadap ketuntasan hasil belajar siswa MA.
B2	Al-Hasyim et al. (2024)	Model <i>Problem Based Learning</i> (PBL)	Kuantitatif	PBL meningkatkan hasil belajar kategori sedang ke tinggi pada siswa MA.
B3	Diza Jusriani & Ibrohim Muchlis (2023)	Model Kooperatif <i>Jigsaw</i>	Eksperimen	Terdapat peningkatan minat dan partisipasi siswa yang signifikan dengan <i>Jigsaw</i> .
B4	Hakim & Komilah (2024)	Model Kooperatif <i>Jigsaw</i>	Kualitatif	<i>Jigsaw</i> efektif mengatasi kepasifan siswa, namun butuh manajemen waktu yang baik.
B5	Putri et al. (2025)	Model Kooperatif STAD	Eksperimen Semu	Model STAD berpengaruh positif terhadap hasil belajar dibandingkan metode ceramah.
B6	Zuraidah (2021)	Model Kooperatif STAD	PTK (Tindakan Kelas)	Peningkatan ketuntasan belajar dari siklus I ke siklus II pada siswa MIN.
B7	Ayuningsih (2020)	Metode <i>Mind Mapping</i>	Eksperimen Semu	<i>Mind Mapping</i> membantu siswa memetakan konsep abstrak Akidah menjadi lebih konkret.
B8	Istijabah (2021)	Metode <i>Snowball Throwing</i>	PTK	Peningkatan hasil belajar materi Ilmu Kalam secara signifikan di setiap siklus.
B9	Khotimah & Nurhasanah (2024)	Metode <i>Snowball Throwing</i> (<i>Critical Thinking</i>)	Kualitatif	Metode ini efektif menstimulasi kemampuan berpikir kritis siswa kelas 5 MI.
B10	Maulid et al. (2024)	Strategi <i>Active Learning</i>	Kualitatif	<i>Active Learning</i> membuat suasana kelas hidup, namun



B11	Huda & Achadi (2024)	Kurikulum Merdeka	Studi Kasus	terkendala sarana prasarana. Implementasi pembelajaran berdiferensiasi mulai diterapkan meski evaluasi masih adaptasi.
B12	Asyifa et al. (2023)	Pendidikan Karakter Abad 21	SLR	Akidah Akhlak adalah pondasi karakter; perlu integrasi teknologi dalam pembelajarannya.

Sumber: Hasil Penelitian, 2025

Temuan pertama yang sangat menonjol dari analisis literatur kelompok artikel B adalah adanya hegemoni penggunaan model pembelajaran kooperatif (*cooperative learning*) sebagai strategi utama untuk meningkatkan prestasi belajar siswa di madrasah. Berdasarkan data yang diperoleh, para pendidik dan peneliti cenderung meninggalkan metode ceramah konvensional yang bersifat satu arah (*teacher-centered*) dan beralih pada metode yang menuntut partisipasi aktif siswa. Hal ini terkonfirmasi melalui studi yang dilakukan oleh Diza Jusriani dan Ibrohim Muchlis serta penelitian Diza Jusriani & Ibrohim Muchlis (2023) dan Hakim & Komilah (2024), yang secara spesifik menyoroti efektivitas model pembelajaran kooperatif tipe *Jigsaw*. Kedua penelitian tersebut menemukan fakta empiris bahwa model *Jigsaw* mampu memecahkan masalah klasik dalam kelas Akidah Akhlak, yaitu pasifnya siswa dalam menerima materi-materi dogmatis. Dengan *Jigsaw*, siswa dipaksa untuk menjadi "ahli" dalam sub-materi tertentu dan mengajarkannya kepada teman sejawat, yang secara psikologis meningkatkan rasa tanggung jawab akademik dan memori jangka panjang terhadap materi pelajaran.

Selain tipe *Jigsaw*, varian model kooperatif lainnya yang mendominasi lanskap penelitian adalah *Student Teams Achievement Division* atau STAD. Penelitian yang dilakukan oleh Putri et al. (2025) serta Zuraidah (2021) menunjukkan pola yang serupa, di mana penerapan STAD secara konsisten mampu meningkatkan ketuntasan hasil belajar siswa dari siklus ke siklus. Temuan ini memberikan indikasi kuat bahwa karakteristik mata pelajaran Akidah Akhlak yang sarat dengan nilai-nilai sosial sangat kompatibel dengan model pembelajaran yang berbasis kerja tim. Dalam diskusi STAD, siswa tidak hanya belajar menghafal dalil naqli tentang akhlak terpuji, tetapi juga mempraktikkan akhlak tersebut secara langsung melalui interaksi sosial, seperti menghargai pendapat teman, bekerja sama, dan saling membantu dalam memahami konsep. Hal ini menegaskan bahwa "prestasi belajar" dalam konteks STAD tidak hanya dimaknai sebagai angka kognitif, tetapi juga performa sosial yang terasah selama proses pembelajaran berlangsung.

Pergeseran paradigma pembelajaran dari sekadar "tahu" menjadi "mampu memecahkan masalah" juga terlihat dari munculnya tren penggunaan model *Problem Based Learning* (PBL) dan metode yang merangsang kekritisian. Studi Al-Hasyim et al. (2024) menjadi bukti valid bahwa PBL mampu mengatrol kategori hasil belajar siswa Madrasah Aliyah dari level sedang ke tinggi. Temuan ini sangat krusial karena membantah stigma bahwa pelajaran agama hanya berisi dogma yang tidak bisa dinalar. PBL membuka ruang dialektika di mana siswa diajak membedah kasus-kasus pelanggaran akhlak kontemporer dengan pisau analisis teologis yang tajam. Senada dengan itu, penelitian Khotimah & Nurhasanah (2024) yang menerapkan metode *Snowball Throwing* menemukan bahwa metode permainan tersebut bukan sekadar *ice breaking*, melainkan strategi ampuh untuk menstimulasi kemampuan berpikir kritis (*critical thinking*) siswa tingkat dasar. Siswa dilatih untuk membuat pertanyaan dan menjawabnya secara sigap, sebuah keterampilan yang jarang disentuh dalam pembelajaran Akidah Akhlak tradisional.

Lebih jauh lagi, spektrum model pembelajaran yang diterapkan juga merambah pada aspek personalisasi dan visualisasi. Penelitian Ayuningsih (2020) mengenai metode *Mind Mapping* menunjukkan bahwa visualisasi konsep sangat membantu siswa dalam memetakan materi Akidah yang bersifat abstrak, seperti konsep-konsep ketuhanan atau eskatologi, menjadi lebih konkret dan mudah dipahami. Sementara itu, Afyah & Sofa (2024) menawarkan pendekatan yang lebih humanis melalui model *Mastery Learning* atau belajar tuntas. Filosofi dari model ini adalah setiap siswa memiliki kecepatan belajar yang berbeda, dan prestasi belajar yang sejati hanya dapat dicapai jika siswa diberikan waktu dan perlakuan yang sesuai dengan kecepatannya tersebut. Temuan Afyah dan Sofa menegaskan bahwa dengan *Mastery Learning*, ketuntasan belajar bukan lagi target statistik kelas semata, melainkan pencapaian individual setiap peserta didik. Ragam temuan ini, mulai dari *Active Learning* umum Maulid et al. (2024) hingga implementasi Kurikulum Merdeka Huda & Achadi (2024), mengerucut pada satu kesimpulan bahwa intervensi model pembelajaran adalah variabel determinan dalam peningkatan prestasi belajar Akidah Akhlak.

Jika kelompok artikel B berbicara mengenai proses, maka kelompok artikel A memberikan gambaran komprehensif mengenai bagaimana hasil dari proses tersebut diukur. Analisis mendalam terhadap delapan artikel evaluasi menunjukkan adanya keprihatinan akademik terhadap kualitas instrumen penilaian yang selama ini digunakan di madrasah. Penelitian Nurhayati et al. (2023) mengungkapkan realitas bahwa banyak soal tes formatif maupun sumatif yang dibuat oleh guru masih memiliki validitas yang rendah dan daya beda yang buruk. Temuan ini menjadi kritik otokritik bagi praktisi pendidikan bahwa tingginya nilai prestasi belajar siswa bisa jadi bukan karena kepintaran mereka, melainkan karena instrumen soalnya yang terlalu mudah atau tidak valid. Hal ini diperkuat oleh temuan Saputra et al. (2023) yang menyoroti ketimpangan ranah penilaian, di mana dominasi ranah kognitif masih sangat kuat sehingga mengerdikan ranah afektif yang sejatinya adalah inti dari mata pelajaran Akidah Akhlak.

Respons terhadap rendahnya kualitas instrumen tersebut melahirkan tren baru dalam penelitian evaluasi, yaitu penggunaan analisis psikometrik modern. Salah satu temuan paling progresif dalam tinjauan ini adalah penelitian Ismi et al. (2023) yang menggunakan *Rasch Model* untuk menganalisis pola jawaban siswa. Penggunaan *Rasch Model* dalam pendidikan agama adalah sebuah terobosan metodologis yang signifikan. Melalui model ini, Syakillah berhasil mendeteksi bias jawaban dan menebak pola ketidakjujuran siswa dalam menjawab soal. Temuan ini memiliki implikasi serius terhadap definisi "prestasi belajar". Seorang siswa bisa saja mendapatkan nilai 90 dalam ujian sumatif, namun analisis *Rasch* dapat membuktikan bahwa nilai tersebut diperoleh dari hasil tebakan atau kecurangan, bukan pemahaman konsep yang sebenarnya. Oleh karena itu, penelitian ini merekomendasikan perlunya pergeseran dari sekadar melihat "skor total" (Teori Tes Klasik) menuju analisis respons butir yang lebih detail untuk memastikan kemurnian prestasi siswa.

Selain perbaikan teknis psikometrik, transformasi evaluasi juga bergerak ke arah substansi materi yang diujikan. Tuntutan keterampilan abad 21 mendorong Latifah et al. (2023) untuk mengembangkan instrumen penilaian berbasis *Higher Order Thinking Skills* (HOTS). Penelitian ini membuktikan bahwa instrumen HOTS mampu melatih siswa untuk tidak sekadar menghafal definisi sifat wajib Allah, tetapi menganalisis implikasi sifat tersebut dalam fenomena kehidupan sehari-hari. Evaluasi berbasis HOTS memaksa siswa untuk mengaktifkan nalar kritisnya, yang mana hal ini sangat relevan dengan temuan pada kelompok model pembelajaran yang menggunakan PBL dalam Al-Hasyim et al. (2024) dan *Critical Thinking* dalam Khotimah & Nurhasanah (2024). Tanpa adanya instrumen HOTS, model pembelajaran kritis yang diterapkan guru akan menjadi sia-sia karena tidak terukur dengan alat yang setara.

Di sisi lain, kesadaran akan hakikat Akidah Akhlak sebagai pendidikan nilai mendorong munculnya tren evaluasi autentik dan evaluasi berbasis proses. Alfazri (2024) dalam penelitiannya di lingkungan pesantren menemukan bahwa evaluasi autentik adalah pendekatan yang paling relevan untuk mengukur kompetensi spiritual dan sosial. Di pesantren, prestasi belajar tidak hanya diukur di dalam kelas saat ujian semester, tetapi juga melalui observasi

perilaku ibadah dan adab santri dalam keseharian selama 24 jam. Ini adalah bentuk evaluasi sumatif yang holistik, di mana nilai akhir merupakan akumulasi dari performa nyata (*performance assessment*), bukan sekadar tes kertas dan pensil. Namun, tantangan implementasi evaluasi semacam ini juga terekam dalam penelitian Hamidi (2023) dan Sohibah et al. (2023), yang melaporkan bahwa beban administrasi yang berat sering kali membuat guru kembali ke pola lama yang konvensional. Kesenjangan antara idealisme evaluasi autentik (Darodjat & Zuchdi, 2016; Alfazri 2024) dengan realitas keterbatasan waktu guru (Sohibah et al., 2023; Hamidi, 2023) menjadi dinamika yang terus mewarnai potret evaluasi pendidikan Islam di Indonesia.

Bagian terpenting dari pembahasan ini adalah sintesis integratif antara temuan pada kelompok model pembelajaran (B) dengan kelompok evaluasi (A). Teori *Constructive Alignment* yang digagas oleh John Biggs menekankan bahwa tujuan pembelajaran, aktivitas pembelajaran (model), dan asesmen (evaluasi) harus membentuk satu garis lurus yang koheren. Berdasarkan analisis silang terhadap 20 artikel tersebut, ditemukan adanya pola hubungan yang unik sekaligus beberapa kesenjangan yang perlu dikritisi. Pada satu sisi, terdapat keselarasan yang menjanjikan antara model pembelajaran *Problem Based Learning* (Al-Hasyim et al., 2024) dengan pengembangan instrumen HOTS (Latifah et al., 2023). Ketika guru menerapkan PBL, siswa dilatih memecahkan masalah, dan ketika diuji menggunakan soal HOTS, kemampuan pemecahan masalah tersebut terukur dengan valid. Ini adalah contoh *best practice* di mana model dan evaluasi saling menguatkan. Prestasi belajar yang dihasilkan dari kombinasi ini adalah prestasi yang valid dan reliabel, mencerminkan kemampuan kognitif tingkat tinggi yang sesungguhnya.

Namun, di sisi lain, tinjauan sistematis ini juga menemukan adanya diskrepansi atau ketidakselarasan pada sebagian besar praktik pembelajaran lainnya. Banyak penelitian pada kelompok B (seperti Diza Jusriani & Ibrohim Muchlis, 2023; Putri et al., 2025; Zuraidah, 2021) yang melaporkan keberhasilan peningkatan hasil belajar melalui model kooperatif (*Jigsaw/STAD*), namun instrumen evaluasi yang digunakan untuk mengukur keberhasilan tersebut sering kali masih berupa tes pilihan ganda konvensional yang lebih banyak mengukur aspek ingatan. Hal ini menimbulkan pertanyaan kritis: apakah peningkatan hasil belajar tersebut benar-benar mencerminkan efektivitas model kooperatif, atau hanya sekadar efek *drilling* soal? Jika model pembelajarannya menekankan pada kerja sama tim dan tanggung jawab sosial (seperti dalam *Jigsaw*), seharusnya evaluasi sumatifnya juga menyertakan komponen penilaian antar-teman (*peer assessment*) atau penilaian kinerja kelompok, sebagaimana disarankan dalam model evaluasi yang dikembangkan oleh Darodjat & Zuchdi (2016). Ketiadaan komponen penilaian proses sosial dalam evaluasi sumatif pada penelitian-penelitian model kooperatif menunjukkan bahwa *alignment* atau keselarasan belum sepenuhnya terjadi.

Lebih lanjut, temuan mengenai evaluasi autentik di pesantren (Alfazri, 2024) memberikan perspektif baru dalam memandang relevansi evaluasi sumatif. Dalam konteks Akidah Akhlak, "prestasi belajar" yang sesungguhnya adalah perubahan perilaku (*behavioral change*). Model pembelajaran tuntas (*Mastery Learning*) yang diteliti oleh Afyah & Sofa (2024) seharusnya tidak hanya menargetkan ketuntasan nilai kognitif di atas KKM, tetapi juga ketuntasan pengamalan nilai akhlak. Di sinilah letak urgensi integrasi temuan Alfazri (2024) ke dalam praktik model pembelajaran di madrasah umum. Jika sekolah hanya mengandalkan model pembelajaran aktif (Maulid et al., 2024), namun evaluasinya masih terpaku pada tes tulis akhir semester (Nurhayati et al., 2023), maka sekolah hanya akan menghasilkan lulusan yang "tahu tentang akhlak" tetapi belum tentu "berakhlak". Oleh karena itu, evaluasi sumatif harus diredefinisi tidak hanya sebagai *assessment of learning* (penilaian hasil), tetapi juga harus mampu memotret *assessment as learning* (penilaian sebagai pembelajaran) melalui refleksi diri siswa yang terintegrasi dalam model pembelajaran.

Analisis terhadap implementasi Kurikulum Merdeka dalam Huda & Achadi (2024) dan tuntutan pendidikan karakter abad 21 dalam Asyifa et al. (2023) semakin mempertegas kebutuhan akan sinergi ini. Kurikulum Merdeka memberikan fleksibilitas bagi guru untuk menggunakan asesmen diagnostik di awal (yang selaras dengan prinsip *Mastery Learning* oleh Asyifa et al., 2023) dan asesmen formatif yang berkelanjutan. Namun, temuan Hamidi (2023)

mengingatkan bahwa tanpa dukungan sistem dan pengurangan beban administrasi, guru akan kesulitan melakukan evaluasi yang ideal tersebut. Oleh karena itu, solusi yang ditawarkan dari sintesis artikel-artikel ini adalah penyederhanaan instrumen namun dengan peningkatan kualitas butir soal (mengacu pada Nurhayati et al. (2023) dan Ismi et al. (2023)). Guru tidak perlu membuat ratusan soal, tetapi cukup membuat beberapa soal HOTS (Latifah et al., 2023) atau tugas proyek autentik yang mampu mengukur berbagai dimensi kemampuan siswa sekaligus.

Temuan-temuan dalam penelitian ini memiliki implikasi teoritis yang mendalam terhadap pengembangan ilmu pendidikan Islam. Secara teoritis, hasil *review* ini memodifikasi pemahaman tentang konsep "Evaluasi Sumatif" dalam Akidah Akhlak. Evaluasi sumatif tidak lagi dapat dipandang sebagai entitas yang terpisah yang berdiri sendiri di akhir semester. Ia harus dipandang sebagai satu kesatuan organik dengan model pembelajaran. Teori evaluasi pendidikan Islam harus mengadopsi pendekatan multidimensi, menggabungkan ketepatan psikometrik ala barat (seperti *Rasch Model* pada Ismi et al. (2023)) dengan nilai-nilai spiritualitas Islam (seperti *Authentic Assessment* di Pesantren pada Alfazri (2024)). Sintesis ini melahirkan sebuah kerangka kerja baru yang dapat disebut sebagai "Model Asesmen Integratif Karakter-Kompetensi". Dalam model ini, prestasi belajar siswa adalah resultan dari kemampuan berpikir kritis (hasil dari PBL/HOTS) dan kemuliaan perilaku (hasil dari pembiasaan/Autentik).

Secara praktis, bagi para guru dan pemangku kebijakan di madrasah, hasil pembahasan ini menawarkan peta jalan perbaikan yang konkret. Pertama, guru yang menggunakan model pembelajaran kooperatif (*Jigsaw*, *STAD*, *Snowball Throwing*) disarankan untuk tidak hanya mengandalkan tes tulis sebagai penentu nilai sumatif, tetapi wajib memasukkan rubrik penilaian proses diskusi dan kerja sama. Kedua, sekolah harus mulai meninggalkan bank soal lama yang tidak tervalidasi dan beralih mengembangkan bank soal berbasis HOTS yang telah teruji validitas dan reliabilitasnya, jika perlu menggunakan analisis modern untuk meminimalisir kecurangan. Ketiga, pelatihan guru tidak boleh lagi dipisah antara "Pelatihan Model Pembelajaran" dan "Pelatihan Pembuatan Soal". Kedua materi tersebut harus disajikan dalam satu paket pelatihan integratif agar guru memahami cara menjahit keduanya dalam desain pembelajaran (*lesson design*) yang utuh.

Sebagai penutup bagian pembahasan, dapat disimpulkan bahwa literatur yang ada saat ini (2014-2024) telah menunjukkan kemajuan yang signifikan dalam upaya meningkatkan mutu pembelajaran Akidah Akhlak. Kelompok artikel model pembelajaran (B) telah berhasil membuktikan bahwa kelas agama bisa menjadi kelas yang hidup, aktif, dan menyenangkan. Sementara kelompok artikel evaluasi (A) telah berhasil membuktikan bahwa alat ukur pendidikan agama bisa menjadi alat ukur yang ilmiah, presisi, dan objektif. Tantangan ke depan adalah merobohkan tembok pemisah antara kedua kubu penelitian ini. Penelitian masa depan tidak boleh lagi hanya bertanya "Apakah model X berpengaruh terhadap hasil belajar?", tetapi harus bertanya "Bagaimana model X dan instrumen evaluasi Y berinteraksi dalam membentuk karakter dan kompetensi siswa secara simultan?". Hanya dengan menjawab pertanyaan integratif inilah, pendidikan Akidah Akhlak dapat benar-benar melahirkan generasi yang berprestasi secara akademis sekaligus anggun secara moral, sesuai dengan visi pendidikan Islam yang rahmatan lil 'alamin.

SIMPULAN

Berdasarkan hasil penelusuran sistematis dan sintesis mendalam terhadap dua puluh artikel ilmiah yang terbit dalam satu dekade terakhir, penelitian ini menyimpulkan bahwa lanskap pembelajaran Akidah Akhlak di Indonesia sedang mengalami transformasi dinamis menuju arah yang lebih progresif, meskipun masih menyisakan tantangan krusial dalam aspek keselarasan sistemik. Temuan utama dari tinjauan ini mengonfirmasi hipotesis awal bahwa penerapan model pembelajaran aktif memiliki korelasi positif yang signifikan terhadap peningkatan prestasi belajar siswa, baik pada dimensi kognitif maupun afektif. Secara spesifik, dominasi model pembelajaran kooperatif seperti tipe *Jigsaw* dan *Student Teams Achievement Division* serta model berbasis masalah atau *Problem Based Learning* terbukti ampuh mengubah paradigma kelas Akidah Akhlak

dari sekadar transfer dogma menjadi arena dialektika sosial dan pengembangan nalar kritis. Keberhasilan model-model ini menunjukkan bahwa materi keagamaan yang bersifat abstrak dapat dikonkretkan dan diinternalisasi secara efektif melalui interaksi antar siswa dan pemecahan masalah kontekstual.

Di sisi lain, temuan mengenai instrumen evaluasi menunjukkan adanya evolusi metodologis yang menggembirakan namun belum merata. Pergeseran fokus dari teori tes klasik menuju penggunaan analisis psikometrik modern seperti *Rasch Model* dan pengembangan instrumen berbasis *Higher Order Thinking Skills* menandakan meningkatnya kesadaran akademisi akan pentingnya validitas pengukuran. Instrumen evaluasi tidak lagi sekadar dipandang sebagai alat administratif pelengkap rapor, melainkan sebagai instrumen diagnosis yang mampu mendeteksi pola pemahaman dan kejujuran siswa. Selain itu, penguatan pada evaluasi autentik di lingkungan pesantren memberikan bukti empiris bahwa penilaian kompetensi akhlak memerlukan pendekatan holistik yang berbasis observasi perilaku nyata, bukan sekadar tes tertulis. Hal ini menegaskan bahwa definisi prestasi belajar dalam mata pelajaran Akidah Akhlak telah meluas, mencakup integrasi antara kecerdasan intelektual dalam menganalisis dalil dan kecerdasan spiritual dalam pengamalan nilai.

Kendati demikian, simpulan paling kritis dari studi ini adalah ditemukannya fenomena ketidakselarasan konstruktif atau *missing link* antara inovasi pedagogis dengan instrumen evaluasi pada sebagian besar praktik pembelajaran yang diteliti. Mayoritas penelitian cenderung terpolarisasi, di mana inovasi model pembelajaran sering kali tidak diikuti dengan inovasi instrumen penilaian yang setara. Akibatnya, terjadi diskrepansi di mana proses pembelajaran telah berjalan secara aktif, kritis, dan kolaboratif, namun alat ukur yang digunakan masih bersifat konvensional, individual, dan hanya menyentuh level kognitif rendah. Kesenjangan ini berpotensi mendistorsi profil kompetensi siswa yang sesungguhnya, karena kemampuan berpikir tingkat tinggi dan keterampilan sosial yang terbangun selama proses pembelajaran tidak terpotret secara akurat oleh instrumen evaluasi sumatif yang digunakan. Oleh karena itu, sinergi antara model pembelajaran dan sistem evaluasi bukan lagi sekadar pilihan, melainkan prasyarat mutlak untuk menjamin mutu pendidikan karakter yang autentik.

Berdasarkan simpulan tersebut, saran yang dapat diajukan untuk pengembangan konsep penelitian selanjutnya dan perbaikan praktis perlu difokuskan pada upaya integrasi sistemik. Bagi para peneliti di bidang pendidikan Islam, disarankan untuk menggeser tren penelitian dari studi efektivitas model tunggal menuju penelitian pengembangan atau *Research and Development* yang menghasilkan paket desain pembelajaran utuh. Penelitian masa depan perlu mengembangkan modul ajar yang di dalamnya sudah terintegrasi antara sintaks model pembelajaran tertentu dengan rubrik penilaian proses dan instrumen tes hasil belajar yang relevan. Kajian mengenai pengaruh penerapan asesmen formatif berkelanjutan dalam model pembelajaran tuntas juga menjadi lahan riset yang potensial untuk dieksplorasi lebih jauh guna menjawab kebutuhan personalisasi belajar siswa. Selain itu, riset-riset selanjutnya diharapkan dapat memperluas jangkauan analisis pada efektivitas penggunaan teknologi digital dalam evaluasi pembelajaran Akidah Akhlak, mengingat tantangan era digital yang semakin kompleks.

Bagi praktisi pendidikan dan pemangku kebijakan, saran utama yang diajukan adalah perlunya redisiain pelatihan guru yang tidak lagi memisahkan antara pelatihan metodologi pengajaran dengan pelatihan evaluasi. Guru perlu dibekali keterampilan untuk menyusun *blueprint* evaluasi sejak merencanakan model pembelajaran, sehingga tercipta koherensi antara apa yang diajarkan, bagaimana cara mengajarkannya, dan bagaimana cara mengukurnya. Penggunaan instrumen penilaian autentik dan soal berbasis nalar tinggi harus mulai dibudayakan sebagai standar baru dalam ujian sumatif di madrasah, menggantikan soal-soal hafalan yang kurang relevan dengan tantangan abad kedua puluh satu. Dengan langkah-langkah strategis ini, diharapkan pendidikan Akidah Akhlak mampu melahirkan generasi yang tidak hanya unggul secara akademis, tetapi juga memiliki kekokohan akidah dan keluhuran akhlak yang teruji secara valid dan reliabel. Transformasi ini pada akhirnya akan mengembalikan marwah pendidikan agama sebagai fondasi utama pembangunan karakter bangsa.



DAFTAR PUSTAKA

- Afiyah, N., & Sofa, M. (2024). Pengaruh Model Pembelajaran Tuntas Terhadap Hasil Belajar Akidah Akhlak Materi Hidup Mulia Dengan Menghormati Orang Tua dan Guru (Studi di Kelas X MA An-Nidomiyah Cilegon). *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial*, 1(11), 108–116. <https://doi.org/https://doi.org/10.5281/zenodo.11409194>
- Al-Hasyim, R., Kasran, & Rambe, M. (2024). Pengaruh Model Pembelajaran Problem Based Learning Terhadap Hasil Belajar Siswa Pada Mata Pelajaran Akidah Akhlak Di Madrasah Aliyah Ikhwanul Muslimin Kabupaten Serdang Bedagai. *At-Tarbiyah: Jurnal Penelitian dan Pendidikan Agama Islam*, 2(1), 208–214. <https://journal.staittd.ac.id/index.php/at/article/view/112>
- Alfazri. (2024). Implementasi Evaluasi Autentik Dalam Pembelajaran Akidah Akhlak Di Madrasah Tsanawiyah Pondok Pesantren Assalam Naga Beralih Kecamatan Kampar Utara. *Jurnal Intelek dan Cendekiawan Nusantara*, 1(2), 2868–2878. <https://jicnusanantara.com/index.php/jicn/article/view/360/>
- Asyifa, H. S., Fitriyah, I., Mujakki, M. F., & Pambayun, S. P. (2023). Systematic Literature Review: Pembelajaran Akidah Akhlak Sebagai Pondasi Pendidikan Karakter Pada Abad 21. *Prosiding Konferensi Nasional Pendidikan Inovatif*, 1(1), 82–92. <https://prokonpi.uinsa.ac.id/index.php/prokonpi/article/view/9/9>
- Ayuningsih, W. (2020). Pengaruh Metode Mind Mapping terhadap Hasil Belajar Mata Pelajaran Akidah Akhlak di MTs Bina Cendekia. *AL- ADABIYAH: Jurnal Pendidikan Agama Islam*, 1(1), 1–16. <https://repository.syekhnuurjati.ac.id/5401/>
- Creswell, J. W. (2007). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (2nd ed.). SAGE Publications.
- Darodjat, D., & Zuchdi, D. (2016). Model Evaluasi Pembelajaran Akidah dan Akhlak di Madrasah Tsanawiyah (MTs). *Jurnal Penelitian dan Evaluasi Pendidikan*, 20(1), 11–26. <https://doi.org/10.21831/pep.v20i1.7517>
- Diza Jusriani, & Ibrohim Muchlis. (2023). Pengaruh Model Pembelajaran Kooperatif Tipe Jigsaw terhadap Mata Pelajaran Akidah Akhlak dalam Meningkatkan Minat Belajar Peserta Didik di MTS Al Mustaqim Parepare. *Al-Ibrah : Jurnal Pendidikan dan Keilmuan Islam*, 8(2), 1–29. <https://doi.org/10.61815/alibrah.v8i2.278>
- Hakim, D., & Komilah, I. A. (2024). Penerapan Model Pembelajaran Kooperatif Tipe Jigsaw untuk Motivasi Belajar Siswa pada Mata Pelajaran Aqidah Akhlak Kelas VII di MTs Sunan Gunung Jati Katemas Tahun Ajaran 2022/2023. *Jurnal Pendidikan Sosial Dan Humaniora*, 3(3), 2645–2654. <https://publisherqu.com/index.php/pediaqu/article/view/1246>
- Hamidi. (2023). Evaluasi Proses Pembelajaran Akidah Akhlak di Madrasah Aliyah Swasta Ist'dadul Mu'allimien Kecamatan Pengabuan Provinsi Jambi Tahun Akademik 2023/2024. *Unisan Jurnal: Jurnal Manajemen dan Pendidikan*, 3(1), 807–816. <http://journal.an-nur.ac.id/index.php/unisanjournal/article/view/2176>
- Huda, S., & Achadi, M. W. (2024). Implementasi Pembelajaran Akidah Akhlak Berbasis Kurikulum Medeka di Kelas 2 MIN 2 Bantul Yogyakarta. *Al-Madrasah: Jurnal Pendidikan Madrasah Ibtidaiyah*, 8(1), 1. <https://doi.org/10.35931/am.v8i1.2929>



- Ismi, S. F., Tanjung, H. P., & Setiabudi, D. I. (2023). Analisis Jawaban Soal Akidah Akhlak Siswa Kelas III Madrasah Ibtidaiyah Ma'had Al-Zaytun Menggunakan Model Rash. *Seroja: Jurnal Pendidikan*, 2(1), 111–121. <https://doi.org/https://doi.org/10.572349/seroja.v2i1.550>
- Istijabah. (2021). Peningkatan Hasil Belajar Akidah Akhlak Materi Aliran dan Tokoh Ilmu Kalam melalui Metode Snowball Throwing di Madrasah Aliyah Salafiyah Kota Cirebon. *Jurnal Pendidikan Madrasah*, 6(1), 39–44. <https://doi.org/10.14421/jpm.2021.61-05>
- Khotimah, K., & Nurhasanah, M. (2024). Application of the Snowball Throwing Method to Enhance the Critical Thinking of 5th Grade Students in The Learning of Akidah Akhlak at MI PSM Gedoro. *EDUCAN: JURNAL PENDIDIKAN ISLAM*, 8(1), 86–97. <https://doi.org/10.21111/educan.v8i1.11436>
- Kitchenham, B., & Brereton, P. (2013). A systematic review of systematic review process research in software engineering. *Information and Software Technology*, 55(12), 2049–2075. <https://doi.org/10.1016/j.infsof.2013.07.010>
- Latifah, A., Husaini, F., & Khoirotnun Nisa, A. (2023). Pengembangan Instrumen Penilaian Berbasis HOTS. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 9(2), 4486–4496. <https://doi.org/10.36989/didaktik.v9i2.1057>
- Maulid, A. A., Nasrullah, Y., & Saifullah, I. (2024). Pengaruh Metode Pembelajaran Active Learning Terhadap Hasil Belajar Siswa Pada Mata Pelajaran Akidah Akhlak (Penelitian di Kelas VIII MTs Ponpes Cipari, Garut). *Jurnal Pendidikan Agama Islam*, 3(1), 165. <https://doi.org/10.52434/jpai.v3i1.3737>
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.). SAGE Publications.
- Nurhayati, Sukino, & Zulkifli. (2023). Analisis Butir Soal Tes Formatif Mata Pelajaran Akidah Akhlak Kelas VI MIN 2 Kubu Raya 2022. *Risalah Jurnal Pendidikan Dan Studi Islam*, 9(3), 997–1006. https://doi.org/https://doi.org/10.31943/jurnal_risalah.v9i3.584
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. <https://doi.org/10.1136/bmj.n71>
- Putri, K., Nasution, R. H., & Novianti, W. (2025). Pengaruh Model Pembelajaran Kooperatif Tipe STAD Terhadap Hasil Belajar Siswa Pada Mata Pelajaran Akidah Akhlak Di Madrasah Aliyah Rohani Ikhwanul Muslimin Kecamatan Tebing Syahbandar Kabupaten Serdang Bedagai. *At-Tarbiyah: Jurnal Penelitian dan Pendidikan Agama Islam*, 3(1), 127–133. <https://journal.staittd.ac.id/index.php/at/article/view/536>
- Saputra, D., Mahubuh, N., Sari, D. R., Khotimah, D. K., & Zuhriya, I. A. (2023). Analisis Instrumen Penilaian Mata Pelajaran Akidah Akhlak di MA Bilingual Kota Batu. *Raudhah Proud To Be Professionals: Jurnal Tarbiyah Islamiyah*, 7(1), 146–153. <https://ejournal.stairu.ac.id/index.php/raudhah/article/view/499>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>



- Sohibah, Peranginangin, H., & Rahim, A. (2023). Evaluasi Pelaksanaan Pembelajaran Akidah Akhlak Siswa Kelas Iv Madrasah Ibtidaiyah Muhammadiyah Sukajati Haurgeulis. *JISIP (Jurnal Ilmu Sosial dan Pendidikan)*, 7(2), 1063–1068. <https://doi.org/http://dx.doi.org/10.58258/jisip.v7i2.4716>
- Xiao, Y., & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, 39(1), 93–112. <https://doi.org/10.1177/0739456X17723971>
- Zuraidah, Z. (2021). Model Pembelajaran Kooperatif Tipe STAD terhadap Peningkatan Hasil Belajar Akidah Akhlak. *Pedagogi: Jurnal Ilmiah Pendidikan*, 7(1), 16–21. <https://doi.org/10.47662/pedagogi.v7i1.300>