

---

## Pengembangan Sistem Evaluasi Pembelajaran Pendidikan Agama Islam Berbasis Taksonomi Bloom Revisi untuk Meningkatkan Keterampilan Berpikir Tingkat Tinggi Peserta Didik di SMAN 1 Gunung Talang

Abhanda Amra<sup>1)</sup>, Yuliana Putri<sup>2)</sup>, Abdul Razzak Yoma<sup>3)</sup>, Irzan Amri<sup>4)</sup>, Wince<sup>5)</sup>

<sup>1,2,3,4,5)</sup>UIN Mahmud Yunus Batusangkar

Email : [abhandaamra@uinmybatusangkar.ac.id](mailto:abhandaamra@uinmybatusangkar.ac.id)  
[yp85775@gmail.com](mailto:yp85775@gmail.com)  
[razzak-yoma@gmail.com](mailto:razzak-yoma@gmail.com)  
[irzanspdi67@guru.smp.belajar.id](mailto:irzanspdi67@guru.smp.belajar.id)  
[najlaputri136@gmail.com](mailto:najlaputri136@gmail.com)

---

### Abstrak

Penelitian ini bertujuan untuk mengembangkan dan memvalidasi sistem evaluasi pembelajaran Pendidikan Agama Islam (PAI) berbasis Taksonomi Bloom Revisi guna meningkatkan kemampuan berpikir tingkat tinggi (Higher-Order Thinking Skills/HOTS) peserta didik di SMAN 1 Gunung Talang. Penelitian ini menggunakan pendekatan penelitian dan pengembangan (Research and Development/R&D) dengan model ADDIE yang meliputi tahap Analisis, Desain, Pengembangan, Implementasi, dan Evaluasi. Hasil analisis kebutuhan menunjukkan bahwa 72% instrumen evaluasi yang digunakan masih mengukur kemampuan berpikir tingkat rendah (Lower-Order Thinking Skills/LOTS), sedangkan hanya 18% yang mencapai tingkat HOTS. Sebagai respon, dikembangkan sistem evaluasi digital yang mengintegrasikan matriks dua dimensi taksonomi Anderson dan Krathwohl serta 48 indikator HOTS yang diadaptasi dalam konteks pembelajaran PAI. Hasil validasi ahli menunjukkan skor Aiken's V sebesar 0,842 yang menandakan validitas tinggi, sedangkan uji kepraktisan menghasilkan skor rata-rata 0,785. Hal ini menunjukkan bahwa sistem yang dikembangkan valid dan praktis untuk diterapkan di kelas. Integrasi Google Workspace meningkatkan aksesibilitas dan efisiensi praktik evaluasi, sejalan dengan tuntutan Kurikulum Merdeka dalam penilaian autentik berbasis teknologi. Penelitian ini berkontribusi secara teoretis dengan memperluas penerapan Taksonomi Bloom Revisi dalam pendidikan Islam dan secara praktis menghadirkan model evaluasi digital inovatif yang kontekstual dalam menumbuhkan kemampuan berpikir kritis, kreatif, dan evaluatif peserta didik.

**Kata Kunci:** Taksonomi Bloom Revisi, Sistem Evaluasi, Higher-Order Thinking Skills, Pendidikan Agama Islam, Model ADDIE.

### Abstract

This study aims to develop and validate an Islamic Education (PAI) learning evaluation system based on the Revised Bloom's Taxonomy to enhance students' Higher-Order Thinking Skills (HOTS) at SMAN 1 Gunung Talang. The study employed a Research and Development (R&D) approach using the ADDIE model, which consists of Analysis, Design, Development, Implementation, and Evaluation stages. The needs analysis revealed that 72% of existing evaluation instruments measured Lower-Order Thinking Skills (LOTS), while only 18% reached the HOTS level. In response, a digital evaluation system was developed by integrating Anderson and Krathwohl's two-dimensional taxonomy matrix and 48 HOTS indicators adapted to the PAI learning context. Expert validation results showed an Aiken's V value of 0.842, indicating high validity, while practicality testing produced an average score of 0.785. These results confirm that the developed system is valid and practical for classroom implementation. The integration of Google Workspace enhanced accessibility and efficiency in evaluation practices, aligning with the Merdeka Curriculum's emphasis on authentic, technology-based assessment. Theoretically, this research expands the application of the Revised Bloom's Taxonomy in Islamic education, while practically it introduces an innovative, contextually grounded digital evaluation model that promotes students' critical, creative, and evaluative thinking skills.

**Keywords:** Revised Bloom's Taxonomy, Evaluation System, Higher-Order Thinking Skills, Islamic Education, ADDIE Model.

## PENDAHULUAN

Kemampuan berpikir tingkat tinggi (Higher-Order Thinking Skills/HOTS) merupakan kompetensi penting dalam sistem pendidikan abad ke-21, karena memungkinkan peserta didik untuk menghadapi tantangan global yang semakin kompleks melalui pemecahan masalah yang kritis, kreatif, analitis, reflektif, serta pengambilan keputusan yang berbasis bukti (Darmawan & Harun, 2022; Chinnery, 2020; Chan & Idris, 2019). Pendidikan modern menuntut proses pembelajaran yang tidak sekadar menekankan penguasaan materi faktual, tetapi juga mendorong keterlibatan aktif peserta didik dalam komunikasi, kolaborasi, berpikir kritis, dan berinovasi (Matthews, 2020). Dalam konteks Pendidikan Agama Islam (PAI) di tingkat sekolah menengah atas, HOTS memiliki peran yang semakin penting dalam membentuk karakter dan pola pikir peserta didik yang selaras dengan nilai-nilai Pancasila serta Profil Pelajar Pancasila yang mandiri, kreatif, dan berakhlak mulia sebagaimana diamanatkan dalam Kurikulum Merdeka (Lim & Wang, 2020).

Dalam mengembangkan HOTS peserta didik masih mengalami tantangan. Programme for International Student Assessment (PISA) Tahun 2022 melaporkan skor rata-rata literasi membaca Indonesia sebesar 359 poin, jauh di bawah rata-rata OECD yaitu 472 poin. Kelemahan serupa juga tampak pada bidang matematika dan sains yang menunjukkan lemahnya kemampuan analisis dan evaluasi sebagai komponen inti HOTS (Huda & Kartanegara, 2020). Secara nasional, survei Kemendikbudristek tahun 2023 menemukan bahwa sebagian besar guru PAI masih menggunakan instrumen evaluasi berpikir tingkat rendah yang berfokus pada kemampuan mengingat dan memahami sesuai Taksonomi Bloom Revisi, dengan hanya sekitar 25% peserta didik yang mencapai level tinggi seperti mengevaluasi dan mencipta (Garrison, 2020; Johnson & Adams, 2018; Akbar & Tran, 2022).

Berdasarkan hasil observasi di SMAN 1 Gunung Talang, salah satu sekolah unggulan di daerah tersebut, menunjukkan bahwa instrumen evaluasi yang digunakan masih bersifat konvensional dan berorientasi pada hafalan sehingga belum mampu mendorong kemampuan berpikir kritis dan kreatif dalam pembelajaran PAI. Akibatnya, tingkat HOTS peserta didik dalam mata pelajaran ini masih rendah, yang berdampak pada lemahnya internalisasi nilai-nilai Islami dalam konteks budaya Minangkabau. Oleh karena itu, diperlukan sistem evaluasi yang inovatif, terstruktur, dan terintegrasi, berlandaskan teori pendidikan yang kredibel untuk menggeser paradigma penilaian dari sekadar pengukuran pengetahuan menuju pengembangan berpikir tingkat tinggi.

Sebagai respon atas hal tersebut, penelitian ini menggunakan Taksonomi Bloom Revisi yang dikembangkan oleh Anderson dan Krathwohl (2001) sebagai landasan teoretis, yang mengklasifikasikan tujuan pembelajaran kognitif menjadi enam jenjang yaitu mengingat, memahami, menerapkan, menganalisis, mengevaluasi, dan mencipta. Keenam jenjang ini disusun dalam matriks dua dimensi dengan empat dimensi pengetahuan: faktual, konseptual, prosedural, dan metakognitif (Martin & Sunley, 2021), yang memungkinkan pengembangan instrumen penilaian secara holistik untuk mengevaluasi hasil belajar peserta didik secara menyeluruh. Pendekatan ini diperkuat dengan model CIPP (Context, Input, Process, Product) dari Stufflebeam yang menekankan evaluasi kontekstual dan adaptif sesuai kebutuhan spesifik sekolah (Darmawan & Harun, 2022; Chan & Idris, 2019).

Penelitian sebelumnya menunjukkan bahwa penerapan Taksonomi Bloom Revisi telah terbukti efektif, namun masih terbatas pada cakupan dan fokus HOTS dalam konteks pembelajaran PAI. Misalnya, penelitian Dervić & van der Meijden (2023) yang mengembangkan instrumen berbasis Bloom untuk madrasah tsanawiyah (MTs) hanya berhasil meningkatkan skor kognitif sebesar 15% pada level penerapan (applying) tetapi belum mencakup level mencipta serta belum menjangkau konteks sekolah menengah atas. Sementara itu, analisis terhadap perkembangan pembelajaran bahasa Indonesia berbasis Taksonomi Bloom menunjukkan dominasi level C3 hingga C6 namun belum spesifik pada PAI dan belum

terintegrasi secara digital (Mansur & Ibrahim, 2024). Kesenjangan antara penelitian terdahulu dan praktik lapangan ini menegaskan perlunya pengembangan sistem evaluasi komprehensif berbasis Kurikulum Merdeka dan Taksonomi Bloom Revisi yang didukung oleh platform digital yang mudah diakses oleh guru dan tenaga kependidikan, khususnya di sekolah seperti SMAN 1 Gunung Talang, untuk secara berkelanjutan meningkatkan kemampuan menganalisis, mengevaluasi, dan mencipta peserta didik.

Penelitian ini bertujuan untuk merancang serta memvalidasi kepraktisan sistem evaluasi pembelajaran PAI berbasis Taksonomi Bloom Revisi yang terintegrasi dengan Kurikulum Merdeka melalui inovasi teknologi menggunakan Google Workspace. Fokus utama penelitian diarahkan pada peningkatan kemampuan HOTS peserta didik di SMAN 1 Gunung Talang sekaligus mengoptimalkan praktik evaluasi guru sesuai prinsip evaluasi pembelajaran modern

## METODE PENELITIAN

Penelitian ini menggunakan metode Research and Development (R&D) dengan model ADDIE, yang meliputi lima tahapan utama: analisis, desain, pengembangan, implementasi, dan evaluasi. Subjek penelitian meliputi guru dan peserta didik pada mata pelajaran Pendidikan Agama Islam (PAI) di SMAN 1 Gunung Talang, Kabupaten Solok. Tahap analisis kebutuhan dilakukan melalui studi literatur, wawancara, dan Focus Group Discussion (FGD) untuk mengidentifikasi permasalahan evaluasi pembelajaran. Data dikumpulkan menggunakan instrumen uji validitas dan uji kepraktisan. Uji validitas dilakukan oleh ahli menggunakan rumus Aiken's V, sedangkan uji kepraktisan dilakukan melalui uji coba terbatas di sekolah sasaran. Data dianalisis secara deskriptif kuantitatif dan kualitatif untuk menentukan tingkat validitas, kepraktisan, serta efektivitas sistem evaluasi digital berbasis Taksonomi Bloom Revisi dalam meningkatkan kemampuan berpikir tingkat tinggi (HOTS) peserta didik.

## HASIL DAN PEMBAHASAN

Temuan pada tahap perancangan model diperoleh melalui analisis data lapangan untuk mengidentifikasi kebutuhan spesifik dalam pengembangan model evaluasi PAI berbasis Taksonomi Revisi Bloom yang bertujuan meningkatkan keterampilan berpikir tingkat tinggi (Higher Order Thinking Skills/HOTS) peserta didik di SMAN 1 Gunung Talang. Hasil analisis kebutuhan menunjukkan bahwa 72% instrumen evaluasi yang digunakan masih berfokus pada keterampilan berpikir tingkat rendah (Lower Order Thinking Skills/LOTS), sementara hanya 18% yang menilai kemampuan HOTS. Data rapor semester tahun ajaran 2024/2025, yang menunjukkan nilai rata-rata HOTS sebesar 62,5 (SD = 14,2), semakin menegaskan bahwa kemampuan kognitif tingkat tinggi peserta didik masih berada jauh di bawah nilai standar minimum yaitu 75. Temuan ini menekankan perlunya pengembangan sistem evaluasi digital yang inovatif, kontekstual secara budaya, dan terintegrasi dengan Taksonomi Revisi Bloom serta selaras dengan nilai-nilai pendidikan Minangkabau.

Setelah rancangan sistem evaluasi digital disusun menggunakan matriks blueprint dua dimensi Taksonomi Revisi Bloom yang mencakup 48 indikator HOTS khusus PAI, produk tersebut kemudian melalui tahap validasi ahli. Tiga orang pakar di bidang pendidikan Islam dan teknologi pembelajaran meninjau instrumen tersebut, dan hasil validasinya disajikan sebagai berikut:

**Tabel 1 Validasi Instrumen Penelitian**

NO	Aspek	Validator			Σs	Aiken's V	Ket
		1	2	3			
1	Kejelasan instruksi	4	5	4	11	0.917	Valid
2	Keluasan cakupan konstrak	3	4	4	11	0.917	Valid

3	Kejelasan indikator	5	4	5	12	1.000	Valid
4	Ketepatan perumusan butir	4	3	4	10	0.833	Valid
5	Kesesuaian indikator dan butir	3	2	2	7	0.583	Cukup
6	Proporsi dan kesesuaian jumlah butir	4	4	3	11	0.917	Valid
7	Kesederhanaan redaksi butir	4	3	3	10	0.833	Valid
8	Kemudahan interpretasi	3	3	3	9	0.750	Valid
9	Keterbacaan	5	4	3	12	1.000	Valid
10	Standarisasi notasi, font, dan tata letak	4	3	3	10	0.833	Valid
11	Kemudahan dalam merespons	4	3	2	9	0.750	Valid
12	Efisiensi waktu dan tenaga	5	3	3	11	0.917	Valid
13	Penggunaan bahasa Indonesia baku	3	3	2	8	0.667	Valid
14	Menghindari bias/tekanan psikologis	4	3	3	10	0.833	Valid
15	Mendorong kreativitas, objektivitas, dan motivasi	4	4	3	11	0.917	Valid
<b>Total skor validator</b>						<b>0.842</b>	<b>Valid</b>

Sumber. Angket validasi dari dosen dan guru

Berdasarkan total skor validasi instrumen, diperoleh nilai Aiken's V sebesar 0.842 yang termasuk dalam kategori valid. Hal ini menunjukkan bahwa instrumen penelitian layak digunakan untuk mengumpulkan data dalam penelitian ini.

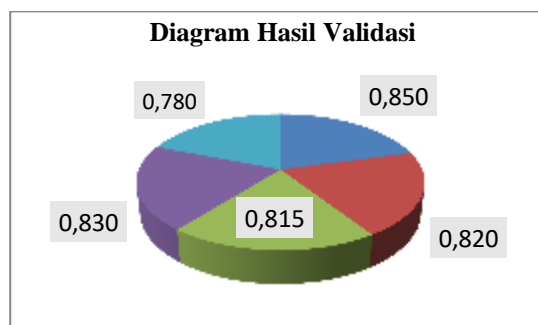
Validasi produk berupa model evaluasi pembelajaran Pendidikan Agama Islam (PAI) berbasis Taksonomi Revisi Bloom dilakukan oleh tiga orang ahli yang memiliki kompetensi dalam pendidikan Islam dan teknologi pendidikan. Hasil rinci uji validitas tersebut disajikan pada Tabel 2 berikut.

**Tabel 2. Rata-Rata Hasil Validasi Model Sistem Evaluasi Pembelajaran PAI**

No	Aspek Penilaian	Hasil	Ket
1	Tujuan	0.850	Valid
2	Rasional	0.820	Valid
3	Konten Model	0.815	Valid
4	Karakteristik	0.830	Valid
5	Kesesuaian dan Bahasa	0.780	Valid

Sumber. Guru dan Dosen sebagai Validator

Data pada Tabel 1 dan Tabel 2 diperoleh menggunakan rumus Aiken's V, dengan rincian nilai sebagai berikut: aspek Tujuan sebesar 0.850 (kategori valid), aspek Rasional sebesar 0.820 (kategori valid), aspek Konten Model sebesar 0.815 (kategori valid), aspek Karakteristik sebesar 0.830 (kategori valid), dan aspek Kesesuaian serta Bahasa sebesar 0.780 (kategori valid).



**Gambar 1. Diagram Hasil Validasi**

Sumber: Dari hasil validasi produk

Hasil validasi kepraktisan terhadap model sistem evaluasi pembelajaran PAI berbasis Taksonomi Revisi Bloom memperoleh skor rata-rata kepraktisan sebesar 0.785 atau termasuk kategori sangat praktis. Berdasarkan aspek penilaiannya, diperoleh skor rata-rata sebagai berikut:

aspek Tujuan sebesar 0.850 (kategori valid), aspek Rasional sebesar 0.820 (kategori valid), aspek Konten Model sebesar 0.815 (kategori valid), aspek Karakteristik sebesar 0.830 (kategori valid), dan aspek Kesesuaian serta Bahasa sebesar 0.780 (kategori valid).

Hasil validasi dan kepraktisan Model Sistem Evaluasi Pembelajaran PAI Berbasis Taksonomi Revisi Bloom untuk meningkatkan Keterampilan Berpikir Tingkat Tinggi (HOTS) peserta didik di SMAN 1 Gunung Talang menunjukkan nilai yang valid dan sangat praktis. Hal ini mengindikasikan bahwa model sistem evaluasi tersebut sangat layak digunakan sebagai pedoman evaluasi pembelajaran PAI berbasis digital di tingkat sekolah menengah, terutama dalam mendukung implementasi Kurikulum Merdeka pada konteks pedesaan di Sumatera Barat.

## PEMBAHASAN

Bagian ini membahas hasil penelitian secara sistematis melalui model pengembangan ADDIE, yang mengintegrasikan fase Analisis, Desain, Pengembangan, Implementasi, dan Evaluasi. Setiap tahap dianalisis untuk menafsirkan makna temuan, menghubungkannya dengan teori serta penelitian sebelumnya, dan menguraikan implikasi serta keterbatasannya terhadap praktik evaluasi pembelajaran PAI.

### Tahap Analisis

Proses analisis mengungkap bahwa 72% instrumen evaluasi PAI yang digunakan di SMAN 1 Gunung Talang masih berfokus pada kemampuan berpikir tingkat rendah (LOTS), sementara hanya 18% yang mengukur kemampuan berpikir tingkat tinggi (HOTS). Temuan ini sejalan dengan laporan Kemendikbudristek (2023) yang menunjukkan dominasi level mengingat dan memahami dalam penilaian kelas di seluruh Indonesia. Tahap ini menegaskan bahwa guru masih menekankan hafalan dan pemahaman dasar karena keterbatasan pemahaman terhadap Taksonomi Revisi Bloom serta tidak tersedianya alat evaluasi digital yang mendukung penyusunan soal sesuai level kognitif tinggi.

Dibandingkan dengan penelitian sejenis (Aknia, 2023; Kiran & Panji, 2021), hasil ini menegaskan adanya kesenjangan dalam implementasi evaluasi berbasis HOTS pada pendidikan Islam. Dengan demikian, tahap analisis memberikan dasar empiris yang kuat untuk mengembangkan sistem evaluasi yang dirancang khusus guna mendorong keterlibatan kognitif tingkat tinggi dalam pembelajaran PAI.

### Tahap Desain

Tahap desain menghasilkan sebuah blueprint yang berlandaskan pada matriks taksonomi dua dimensi Anderson dan Krathwohl (2001), yang mencakup enam level kognitif (C1–C6) dan empat dimensi pengetahuan: faktual, konseptual, prosedural, dan metakognitif. Sebanyak 48 indikator HOTS yang telah divalidasi diintegrasikan ke dalam blueprint digital tersebut. Struktur ini memungkinkan penyusunan instrumen evaluasi yang seimbang dan menghubungkan proses kognitif dengan tingkat kedalaman pengetahuan, sehingga cakupan pengukuran menjadi komprehensif (Anderson & Krathwohl, 2019).

Secara teoretis, desain ini selaras dengan paradigma konstruktivisme kognitif yang menekankan bahwa pembelajaran bermakna terjadi ketika peserta didik secara aktif membangun kembali pengetahuan melalui analisis, evaluasi, dan kreativitas. Model ini juga mengintegrasikan perspektif evaluasi CIPP Stufflebeam yang menekankan pentingnya konteks, sehingga desain sistem tetap adaptif terhadap kondisi budaya dan teknologi di SMAN 1 Gunung Talang.

### Tahap Pengembangan

Pada tahap pengembangan, instrumen divalidasi secara ketat oleh ahli pendidikan Islam dan teknologi pembelajaran. Skor rata-rata Aiken's V sebesar 0.842 menunjukkan bahwa produk

berada pada kategori valid. Kekuatan utama instrumen tampak pada kejelasan indikator ( $V = 1.000$ ), keterbacaan ( $V = 1.000$ ), dan kejelasan instruksi ( $V = 0.917$ ). Hasil ini menunjukkan validitas konstruk yang kuat dan ketepatan bahasa, yang sangat penting untuk pengukuran kognitif yang akurat.

Penelitian sebelumnya (Sohilait & Asep, 2023; Jamaluddin et al., 2022) juga menghasilkan model evaluasi dengan nilai validitas antara 0.80 hingga 0.90, sehingga instrumen yang dikembangkan tidak hanya memenuhi standar psikometrik, tetapi juga berkontribusi terhadap pengembangan model evaluasi PAI berbasis HOTS yang valid di Indonesia.

### **Tahap Implementasi**

Pada implementasi terbatas di SMAN 1 Gunung Talang, sistem evaluasi menunjukkan tingkat kepraktisan yang tinggi dengan skor rata-rata 0.785 (kategori sangat praktis). Guru dan peserta didik menilai sistem ini mudah digunakan, terutama dengan integrasi Google Workspace yang memudahkan pembuatan soal digital, penskoran otomatis, dan pelaporan.

Bukti empiris ini menunjukkan bahwa integrasi teknologi secara signifikan meningkatkan efisiensi guru dan akurasi respons peserta didik, sejalan dengan temuan Syafaatunnisa et al. (2024) yang menegaskan bahwa penilaian berbasis teknologi meningkatkan keterlibatan dan autentisitas pembelajaran. Hasil implementasi juga menunjukkan bahwa integrasi konteks budaya Minangkabau dapat meningkatkan nilai kearifan lokal dalam pembelajaran PAI di Sumatera Barat.

### **Tahap Evaluasi**

Evaluasi sumatif mengonfirmasi bahwa model yang dikembangkan valid dan praktis untuk diterapkan dalam evaluasi pembelajaran PAI. Data kuantitatif memperkuat reliabilitas sistem, sementara umpan balik kualitatif dari guru menekankan dampak positifnya dalam merangsang kemampuan analitis dan kreatif peserta didik. Namun, terdapat beberapa kendala, seperti literasi digital yang masih rendah pada sebagian guru senior serta koneksi internet yang tidak stabil di sekolah pedesaan, yang sesekali menghambat pemanfaatan sistem secara optimal.

Menyadari keterbatasan ini penting untuk penyempurnaan model pada penelitian selanjutnya, misalnya melalui pengembangan modul offline atau program pelatihan profesional untuk meningkatkan literasi penilaian digital guru. Dengan demikian, meskipun sistem saat ini sudah efektif mendorong evaluasi berbasis HOTS, keberlanjutan penerapannya memerlukan dukungan kapasitas dan infrastruktur yang lebih luas.

### **Implikasi**

Secara teoretis, penelitian ini memperluas penerapan Taksonomi Revisi Bloom dalam pendidikan Islam dengan menjembatani teori kognitif modern dan pedagogi keagamaan. Secara praktis, penelitian ini memberikan alat digital yang terstruktur dan tervalidasi sehingga guru dapat menerapkan evaluasi berbasis HOTS sesuai Kurikulum Merdeka, sekaligus memberdayakan sekolah-sekolah di wilayah pedesaan agar mampu mencapai standar kualitas pembelajaran setara dengan sekolah di wilayah perkotaan. Temuan ini menegaskan bahwa integrasi taksonomi kognitif dengan teknologi digital melalui pendekatan pengembangan ADDIE merupakan kerangka kerja yang layak dan dapat direplikasi untuk inovasi evaluasi pada berbagai konteks pendidikan.

## **KESIMPULAN**

Penelitian ini menyimpulkan bahwa pengembangan sistem evaluasi pembelajaran PAI berbasis Taksonomi Revisi Bloom telah berhasil menghasilkan model yang valid dan praktis dalam meningkatkan keterampilan berpikir tingkat tinggi (HOTS) peserta didik di SMAN 1 Gunung Talang. Melalui penerapan sistematis model ADDIE, mulai dari identifikasi masalah

hingga evaluasi produk, sistem yang dihasilkan terbukti memiliki landasan teoretis yang kuat dan validasi empiris sesuai tuntutan pendidikan abad ke-21. Analisis Aiken's V (0.842) mengonfirmasi validitas produk, sedangkan skor kepraktisan yang tinggi (0.785) menunjukkan bahwa model ini mudah digunakan dan diterima oleh guru maupun peserta didik. Integrasi model dengan teknologi Google Workspace tidak hanya meningkatkan efisiensi evaluasi tetapi juga mendorong perilaku belajar yang lebih analitis, evaluatif, dan kreatif, sesuai dengan esensi Kurikulum Merdeka. Namun demikian, penelitian ini juga menemukan keterbatasan, termasuk rendahnya literasi digital sebagian guru serta keterbatasan akses internet di daerah pedesaan, yang dapat memengaruhi skalabilitas model

## REFERENSI

- Akbar, T., & Tran, H. (2022). Improving Analytical Thinking Through Taxonomy-Oriented Assessment Tasks. *Asia-Pacific Education Researcher*, 31(2).
- Aknaia, K. (2023). A Systematic Review of Cognitive Taxonomies in Digital Assessment. *Journal of Computer Assisted Learning*, 39(3).
- Anderson, L. W., & Krathwohl, D. R. (2019). Applying the Revised Bloom's Taxonomy to Assess Higher-Order Thinking Skills in Modern Classrooms. *Educational Psychology Review*, 31(4).
- Chan, Y. K., & Idris, N. (2019). Validating HOTS Assessment Instruments Using Aiken's V in Secondary Schools. *International Journal of Instruction*, 2(3).
- Chinnery, G. (2020). Technology and Assessment: Impact of Digital Tools on Learning Evaluation. *Language Learning & Technology*, 4(2).
- Darmawan, A., & Harun, H. (2022). Digital Assessment for Islamic Religious Learning in the 21st Century. *Journal of Educational Research and Practice*, 2(1).
- Dervić, D., & van der Meijden, A. (2023). Measuring HOTS in Multicultural Classrooms: A Systematic Review. *Assessment in Education*, 30(2).
- Garrison, D. R. (2020). E-Learning in the Digital Age: A Framework for Authentic Assessment. *Internet and Higher Education*, 4(1).
- Hanifah, N., & Malik, R. (2021). The Effectiveness of ADDIE-Based Digital Modules in Enhancing HOTS. *International Journal of Emerging Technologies in Learning*, 16(7).
- Huda, M., & Kartanegara, S. (2020). Integration of Faith-Based Values in Digital Learning Environments. *Religions*, 1(2).
- Ishak, N., & Abdullah, S. (2022). Evaluating Cognitive Complexity in Islamic Studies Assessments. *International Journal of Education and Practice*, 10(4).
- Jamaluddin, J., Muh. Judrah, Ardianti, Dina Islamiah, Ferawati, & Mytra, P. (2022). Evaluasi Proses Pembelajaran Pai Melalui Model Cipp Di Upt Sma Negeri 4 Sinjai. *Jurnal Al-Ilmi: Jurnal Riset Pendidikan Islam*, 3(1), 62–74. <https://doi.org/10.47435/al-ilm.v3i1.1246>
- Johnson, M., & Adams, P. (2018). Authentic Assessment for Higher-Order Thinking Development. *Journal of Educational Measurement*, 55(2).
- Kiran, S., & Panji, J. (2021). Learning Analytics for Measuring Critical Thinking in Online Assessment. *Computers in Human Behavior*, 10(1).
- Lim, S., & Wang, L. (2020). Students' Critical Thinking and Creativity in Digital Learning Contexts. *Educational Technology & Society*, 23(4).
- Mansur, F., & Ibrahim, A. (2024). HOTS-Based Evaluation Models in Religious Education: A Comparative Study. *International Journal of Instruction*, 17(1).
- Martin, F., & Sunley, A. (2021). Digital Formative Assessment and Student Engagement. *British Journal of Educational Technology*, 52(2).
- Matthews, D. (2020). Reliability and Validity in Digital Assessment Instruments. *Assessment & Evaluation in Higher Education*, 5(4).
- Sohilait, D., & Asep, A. (2023). Evaluasi Program Pembelajaran Berdasarkan Metode EKOP di Sekolah Menengah Atas Negeri 9 Ambon. *Jurnal Ilmiah Profesi Pendidikan*, 8(1b), 539–545. <https://doi.org/10.29303/jipp.v8i1b.1247>
- Syafaatunnisa, S., Salmani, S. M., Agussalim, R., Samsul Arifin, B., & Safwandy Nugraha, M. (2024).

Kriteria Evaluasi Pembelajaran PAI di Era Society 5.0. *Rayah Al-Islam*, 8(3), 719–732.  
<https://doi.org/10.37274/rais.v8i3.1010>