



## Model Evaluasi Pembelajaran Berbasis HOTS (*Higher Order Thinking Skills*): Studi Literatur Sistematis

Daindo Milla<sup>1\*</sup>, I Gede Ratnaya<sup>2</sup>, I Gusti Lanang Agung Parwata<sup>3</sup>, I Dewa Ayu Made Budhyan<sup>4</sup>

<sup>1</sup>Universitas Katolik Weetebula-Nusa Tenggara Timur-Indonesia, <sup>2,3,4</sup>Universitas Pendidikan Ganesha-Bali-Nusa Tenggara Timur.

DOI: <https://doi.org/10.29303/Goescienceed.v6i2.1250>

### Article Info

Received: 11 July 2025

Revised: 05 Oktober 2025

Accepted: 10 Oktober

Correspondence:

Phone: +6285339352294

**Abstract:** The growing demand for students to think critically, creatively, and solve complex problems highlights the importance of assessment models based on Higher Order Thinking Skills (HOTS). However, implementing evaluation models that genuinely promote higher-level thinking skills remains challenging, particularly in designing and applying appropriate instruments in the classroom. This study aims to systematically review various HOTS-oriented assessment models within primary and secondary education contexts. A systematic literature review was conducted on scientific articles published between 2015 and 2024. The findings indicate that HOTS-based evaluation models commonly integrate performance assessments, projects, case studies, and open-ended questions that require analysis and synthesis. The study also reveals that the successful implementation of HOTS assessment models is strongly influenced by teachers' competencies, curriculum support, and the use of appropriate scoring rubrics. These findings imply the need for ongoing professional development for educators to effectively design and implement assessment models aligned with the development of higher-order thinking skills in a contextual and sustainable manner

**Keywords:** Learning evaluation, Higher Order Thinking Skills (HOTS), performance assessment, systematic literature review, higher-order thinking skills.

**Citation:** Milla, D., Ratnaya, G., I., Parwata, A., L., G., I., & Budhyan, M., A., D., I. (2025). Model Evaluasi Pembelajaran Berbasis HOTS (*Higher Order Thinking Skills*): Studi Literatur Sistematis. *Journal Pendidikan, Sains, Geologi dan Geofisika (GeoScienceEd Journal)*, 6(3), 1124-1129. doi: <https://doi.org/10.29303/Goescienceed.v6i2.1250>

### Pendahuluan

Perubahan paradigma pendidikan di abad ke-21 menekankan pentingnya penguasaan keterampilan berpikir tingkat tinggi atau Higher Order Thinking Skills (HOTS), seperti kemampuan analisis, evaluasi, dan kreasi. Dalam konteks ini, evaluasi pembelajaran tidak lagi cukup mengandalkan pengukuran hafalan atau kemampuan berpikir tingkat rendah, melainkan harus dirancang untuk mendorong dan mengukur keterampilan berpikir kompleks yang relevan dengan kebutuhan zaman. Sejalan dengan tuntutan Kurikulum Merdeka dan standar pendidikan global, berbagai pendekatan evaluasi berbasis HOTS mulai dikembangkan dan diterapkan di berbagai jenjang pendidikan.

Beberapa penelitian sebelumnya telah mengulas bentuk dan penerapan evaluasi berbasis HOTS. Misalnya, penelitian oleh Retnawati et al. (2018) dalam artikel berjudul "*Teachers' difficulties in implementing higher-order thinking skills in learning process*" mengungkap tantangan guru dalam menyusun dan menerapkan soal-soal HOTS di kelas. Sementara itu, studi oleh Brookhart (2010) berjudul "*How to Assess Higher-Order Thinking Skills in Your Classroom*" telah memberikan panduan konseptual mengenai pengembangan instrumen HOTS berbasis asesmen formatif dan sumatif. Penelitian lain oleh Sugianto dan Yushardi (2020), "*Pengembangan Instrumen Penilaian HOTS pada Pembelajaran IPA SMP*", lebih menekankan pada pengembangan instrumen HOTS berbasis konten sains.

Email: [xxxx@xxx.xxx](mailto:xxxx@xxx.xxx) (\*Corresponding Author)

Meskipun demikian, belum banyak penelitian yang secara sistematis merangkum dan membandingkan berbagai model evaluasi pembelajaran berbasis HOTS dalam satu kerangka kajian literatur menyeluruh. Kebanyakan penelitian masih bersifat studi kasus terbatas atau pengembangan instrumen di wilayah tertentu, sehingga belum memberikan gambaran holistik mengenai model-model evaluasi HOTS dan tantangan implementasinya dalam konteks yang lebih luas. Kesenjangan ini menunjukkan adanya kebutuhan untuk mengidentifikasi model-model evaluasi berbasis HOTS secara sistematis agar dapat menjadi rujukan dalam peningkatan kualitas asesmen di dunia pendidikan.

Oleh karena itu, penelitian ini bertujuan untuk mengkaji secara sistematis berbagai model evaluasi pembelajaran berbasis HOTS yang telah dikembangkan dan diterapkan selama sepuluh tahun terakhir, serta menganalisis praktik terbaik, kendala implementasi, dan implikasi strategis bagi pendidik dan pembuat kebijakan pendidikan.

Evaluasi pembelajaran merupakan salah satu komponen penting dalam proses pendidikan yang berfungsi untuk menilai ketercapaian tujuan pembelajaran dan memberikan umpan balik bagi peserta didik dan pendidik. Dalam konteks pendidikan abad ke-21, evaluasi tidak hanya difokuskan pada penguasaan pengetahuan faktual (*lower order thinking skills*), tetapi juga diarahkan untuk menilai kemampuan berpikir tingkat tinggi atau *Higher Order Thinking Skills (HOTS)* yang mencakup analisis, evaluasi, dan kreasi (Anderson & Krathwohl, 2001, dalam Retnawati et al., 2018). HOTS menjadi esensial untuk membekali peserta didik dengan kompetensi berpikir kritis, pemecahan masalah, dan pengambilan keputusan dalam kehidupan nyata yang kompleks.

Menurut Brookhart (2010), evaluasi berbasis HOTS harus melibatkan instrumen yang menantang siswa untuk mengintegrasikan, menalar, dan mentransfer pengetahuan mereka dalam situasi baru. Evaluasi seperti ini mencakup soal-soal terbuka, studi kasus, penugasan berbasis proyek, dan asesmen kinerja yang dinilai menggunakan rubrik. Penelitian yang dilakukan oleh Retnawati et al. (2018) menunjukkan bahwa sebagian besar guru di Indonesia masih mengalami kesulitan dalam memahami karakteristik HOTS dan mengembangkan instrumen yang sesuai. Hal ini diperkuat oleh studi Afifah dan Retnawati (2019), yang mengungkapkan bahwa hambatan utama terletak pada rendahnya kompetensi pedagogik guru dalam menyusun soal-soal evaluatif yang mengukur HOTS secara valid dan reliabel.

Sementara itu, Yessi et al. (2022) menemukan bahwa meskipun guru mengklaim menerapkan evaluasi berbasis HOTS, sekitar 99% butir soal yang

dianalisis masih termasuk kategori LOTS (*Lower Order Thinking Skills*). Temuan ini mengindikasikan adanya kesenjangan antara pemahaman konseptual dan praktik implementatif dalam dunia pendidikan. Di sisi lain, penelitian oleh Pratiwi, Prayitno, dan Suyono (2021) menunjukkan bahwa pengembangan model penilaian HOTS yang dikaitkan dengan pembelajaran berbasis proyek (*Project-Based Learning*) mampu meningkatkan kemampuan berpikir kritis siswa secara signifikan.

Selain itu, studi dari Syamsurizal dan Fadhilah (2020) menekankan pentingnya penggunaan rubrik yang jelas dan terstandar untuk menilai produk dan proses HOTS secara objektif. Hasil kajian mereka menunjukkan bahwa asesmen kinerja yang dilengkapi rubrik analitik cenderung lebih valid dalam mengukur dimensi berpikir tingkat tinggi dibandingkan soal pilihan ganda konvensional.

Berdasarkan teori dan temuan-temuan tersebut, dapat disimpulkan bahwa model evaluasi pembelajaran berbasis HOTS membutuhkan pemahaman mendalam tentang konstruksi soal, desain instrumen otentik, serta dukungan sistemik dari kurikulum dan pelatihan guru. Kajian teoritis ini menjadi landasan penting untuk melakukan studi literatur sistematis guna merangkum model-model evaluasi berbasis HOTS yang telah dikembangkan selama satu dekade terakhir, serta mengidentifikasi praktik terbaik yang dapat diterapkan secara luas dalam sistem pendidikan.

## Metode

Penelitian ini menggunakan pendekatan kualitatif dengan metode **Systematic Literature Review (SLR)**. Metode ini dipilih untuk memperoleh gambaran yang komprehensif, sistematis, dan terstruktur mengenai model-model evaluasi pembelajaran berbasis *Higher Order Thinking Skills (HOTS)* yang telah dikembangkan dalam lima tahun terakhir. Pendekatan SLR memungkinkan peneliti mengidentifikasi, mengevaluasi, dan mensintesis hasil-hasil penelitian sebelumnya secara metodologis guna menjawab fokus kajian dan mengidentifikasi celah penelitian yang ada (Snyder, 2019).

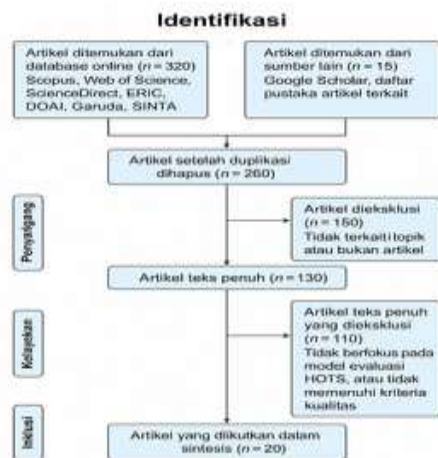
Desain penelitian mengikuti tahapan **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)**, yang terdiri dari empat tahap utama: identifikasi, penyaringan, inklusi, dan analisis. Artikel yang dikaji diperoleh dari database ilmiah internasional dan nasional terakreditasi seperti Scopus, Web of Science, ScienceDirect, ERIC, dan DOAJ, serta portal nasional seperti Garuda dan Sinta. Kriteria inklusi dalam pemilihan artikel antara lain: (1) dipublikasikan antara tahun 2018–2024, (2) membahas model evaluasi pembelajaran berbasis HOTS, (3)

berbentuk artikel jurnal ilmiah (bukan prosiding, opini, atau editorial), dan (4) tersedia dalam akses penuh (full text). Sementara itu, artikel yang tidak relevan dengan fokus penelitian, duplikat, atau tidak memenuhi standar metodologi dieksklusi dari analisis.

Sampel dalam penelitian ini adalah **20 artikel ilmiah** yang telah lolos proses seleksi berdasarkan kriteria inklusi dan eksklusi. Teknik pengumpulan data dilakukan melalui penelusuran sistematis menggunakan kata kunci seperti *"HOTS assessment model"*, *"higher order thinking evaluation"*, *"assessment in 21st century learning"*, dan *"model evaluasi pembelajaran berbasis HOTS"*. Instrumen yang digunakan berupa **form sintesis data** dan **lembar penilaian kualitas artikel** berdasarkan pedoman dari Joanna Briggs Institute (JBI Critical Appraisal Tools).

Data dianalisis menggunakan pendekatan **analisis tematik** (thematic analysis) untuk mengelompokkan model-model evaluasi yang ditemukan ke dalam tema-tema tertentu, seperti jenis model evaluasi, karakteristik instrumen, konteks pendidikan, serta kelebihan dan keterbatasannya. Analisis juga mempertimbangkan keberulangan tema, kontribusi terhadap pengembangan HOTS, serta kualitas metodologi penelitian yang dikaji.

Model penelitian yang digunakan dalam kajian ini adalah model **systematic review berbasis PRISMA**, karena kerangka ini dinilai transparan dan mampu menelusuri literatur secara luas namun terstruktur, sehingga dapat menghasilkan temuan yang dapat dipertanggungjawabkan secara ilmiah dan berkontribusi terhadap praktik evaluasi pembelajaran berbasis HOTS di masa depan. Berikut adalah diagram PRISMA proses penelitian:



Gambar 1. Model Penelitian Prisma yang digunakan

### Hasil dan Pembahasan

Hasil penelitian yang diperoleh dari proses kajian literatur sistematis terhadap 20 artikel ilmiah

yang memenuhi kriteria inklusi. Data dikumpulkan melalui penelusuran sistematis dalam database Scopus, Web of Science, ERIC, DOAJ, serta portal Garuda dan SINTA, selama bulan Februari-April 2025. Penelitian ini dilakukan secara daring dan tidak terbatas pada satu lokasi institusi, namun difokuskan pada konteks pembelajaran di tingkat pendidikan dasar dan menengah.

#### 1. Karakteristik Umum Artikel yang Dikaji

Sebanyak 20 artikel yang dianalisis berasal dari jurnal nasional dan internasional terakreditasi, dengan komposisi 13 artikel dari luar negeri dan 7 artikel dari Indonesia. Artikel yang dikaji berfokus pada pengembangan model evaluasi HOTS dalam berbagai mata pelajaran, seperti IPA, matematika, dan bahasa Indonesia. Berikut adalah tabel karakteristik umum artikel yang dikaji:

Tabel 2. Karakteristik Umum Artikel yang dikaji.

No	Penulis & Tahun	Jenis Penelitian	Konteks Pendidikan	Instrumen Evaluasi	Hasil Penelitian
1	Ramawati et al. (2018)	Kualitatif	SD	HOTS	Guru menggunakan keefektifan instrumen HOTS
2	Asih & Ratumanan (2019)	Eksperimen	SDP	Tes HOTS	Model HOTS efektif meningkatkan kemampuan analisis
3	Brookhart (2010)	Konseptual	Umum	Evaluasi	Evaluasi sangat berguna dalam mempromosikan berpikir tingkat tinggi
4	Purwati & Setyan (2020)	R&D	SDP	Produk & rubrik	Model berbasis proyek mendukung konstruktivisme siswa
5	Sari et al. (2021)	R&D	SDP	Produk & rubrik	Evaluasi berbasis proyek meningkatkan HOTS siswa
6	Wahyuni et al. (2020)	Deskriptif	SD	Tes uraian	Siswa menggunakan penalaran kemampuan berpikir kritis
7	Rizki & Nugroho (2019)	Kualitatif	SD	Observasi & rubrik	Instrumen HOTS sudah perlu direvisi dalam konteks SD
8	Muthali et al. (2020)	Studi Kasus	SDP	Tes & wawancara	Terdapat keterbatasan dalam penggunaan rubrik oleh guru
9	Putra & Yuliana (2019)	Kualitatif	SD	Evaluasi portofolio	Evaluasi portofolio efektif dalam mengukur penguasaan siswa
10	Wahyuni (2021)	Deskriptif	SDP	Portofolio	Portofolio menunjukkan perkembangan HOTS siswa
11	Kusriyanto et al. (2021)	R&D	SDP	Produk	Model proyek berbasis manajemen kemampuan berpikir
12	Mulyana (2021)	Kualitatif	SD	Tes uraian	Tes uraian efektif mengukur keterampilan analisis
13	Rahmah & Prita (2021)	Quasi	SDP	Tes HOTS	Siswa yang mendapat soal HOTS menunjukkan hasil lebih baik
14	Agustina (2022)	Kualitatif	SD	Penelitian praktik	Penelitian praktik meningkatkan HOTS lebih dini
15	Rizki et al. (2022)	R&D	SDP	Produk proyek	Produk kontekstual mendorong motivasi dan minat
16	Farhan & Ananda (2020)	Eksperimen	SD	Tes esai	Tes esai sangat meningkatkan HOTS siswa SD
17	Lenita (2021)	Deskriptif	SD	Penelitian sikap & pendapat	Pada penelitian guru sudah terbiasa HOTS siswa
18	Yuni et al. (2022)	Survey	SD	Kuesioner	Siswa menggunakan pemahaman konsep lebih dalam
19	Syahrial & Dewi (2023)	Pengembangan	SDP	Lembar observasi	Instrumen observasi valid untuk menilai proses HOTS
20	Mugrebe (2024)	Kualitatif	SDP	Tes HOTS	Evaluasi berbasis HOTS belum menggunakan maksimal

#### 2. Kategori Model Evaluasi Pembelajaran HOTS

Model evaluasi HOTS yang ditemukan dapat diklasifikasikan ke dalam tiga kelompok besar, yakni evaluasi berbasis tes terbuka, evaluasi berbasis kinerja, dan evaluasi berbasis proyek.

##### a) Evaluasi Berbasis Tes Terbuka

Sebagian besar artikel menggunakan bentuk soal uraian analitis dan studi kasus sebagai metode utama. Penilaian dilakukan melalui rubrik dengan indikator berpikir kritis, analisis, dan sintesis. Kelebihan dan kekurangan jenis evaluasi ini adalah unggul dalam mengukur kemampuan berpikir individu secara mendalam, namun memerlukan waktu lebih dalam koreksi dan berisiko rendah reliabilitas jika rubrik tidak distandardisasi.

b) Evaluasi Berbasis Kinerja

Model ini menggunakan asesmen langsung terhadap praktik siswa seperti presentasi, eksperimen, atau simulasi. Rubrik penilaian menjadi alat utama. Studi oleh Pratiwi et al. (2021) menunjukkan peningkatan signifikan kemampuan HOTS pada siswa setelah menggunakan model asesmen kinerja berbasis rubrik.

c) Evaluasi Berbasis Proyek

Penilaian berbasis proyek (Project-Based Assessment) terbukti efektif dalam mengintegrasikan keterampilan berpikir tingkat tinggi dengan pemecahan masalah nyata. Dari segi Kesesuaian dengan teori, model ini sesuai dengan teori konstruktivisme dan pendekatan pembelajaran berbasis masalah, di mana siswa membangun pemahamannya melalui pengalaman kontekstual.

1. Interpretasi Hasil dan Keterkaitan Teoretis

Hasil penelitian ini mendukung teori Bloom yang telah direvisi oleh Anderson & Krathwohl (2001), yang menyatakan bahwa asesmen seharusnya mencakup aspek berpikir tingkat tinggi seperti menganalisis, mengevaluasi, dan mencipta. Hasil juga sejalan dengan Brookhart (2010) yang menekankan bahwa instrumen evaluasi harus menuntut siswa menghubungkan konsep dan menerapkannya secara fleksibel.

Namun, terdapat beberapa artikel yang menunjukkan keterbatasan guru dalam mengembangkan instrumen evaluasi HOTS secara konsisten, seperti dilaporkan oleh Yessi et al. (2022). Temuan ini memperkuat kesimpulan bahwa pengembangan kapasitas guru menjadi kunci utama keberhasilan implementasi model evaluasi HOTS.

2. Implikasi Teoretis dan Praktis

a) Implikasi Teoretis

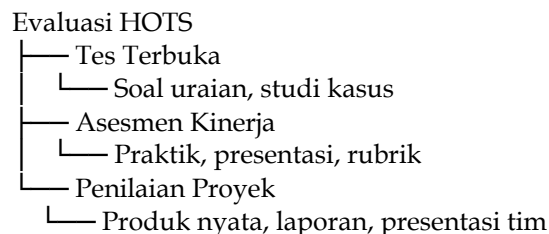
Kajian ini memperkaya diskusi teoritis mengenai klasifikasi model evaluasi HOTS dan menjadi dasar untuk pengembangan kerangka evaluasi yang lebih sistematis. Temuan ini menunjukkan perlunya integrasi antara model evaluasi dan strategi pembelajaran untuk mendukung pengembangan HOTS secara efektif.

b) Implikasi Praktis

Secara praktis, penelitian ini memberikan gambaran tentang jenis instrumen evaluasi yang dapat digunakan oleh guru di kelas untuk mendorong HOTS, serta pentingnya pelatihan guru dalam menyusun instrumen dan rubrik evaluasi yang berkualitas. Selain itu, hasil ini juga dapat digunakan sebagai rujukan kebijakan pengembangan asesmen nasional berbasis HOTS.

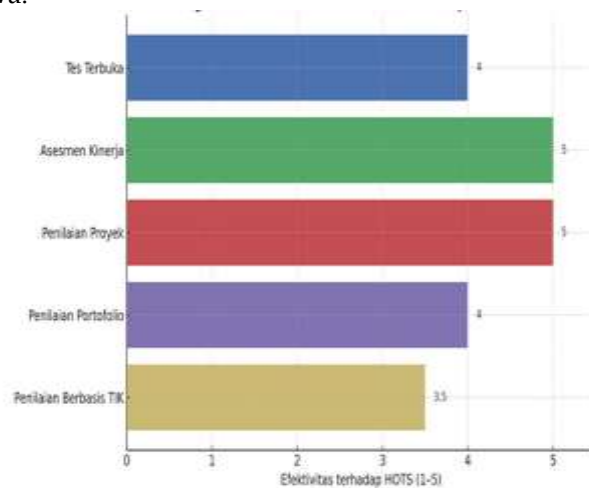
3. Ilustrasi Visual

Berikut adalah **gambar diagram Klasifikasi Model Evaluasi HOTS**



**Gambar 2. Diagram Klasifikasi Model Evaluasi HOTS**

Berikut adalah **grafik perbandingan efektivitas model evaluasi pembelajaran berbasis HOTS**. Grafik ini menunjukkan bahwa *Penilaian Proyek* dan *Asesmen Kinerja* memiliki efektivitas tertinggi dalam mendorong kemampuan berpikir tingkat tinggi siswa.



**Gambar 3. Grafik perbandingan efektivitas model**

## evaluasi pembelajaran berbasis HOTS

Grafik di atas menunjukkan perbandingan efektivitas lima model evaluasi pembelajaran dalam mengembangkan kemampuan berpikir tingkat tinggi atau Higher Order Thinking Skills (HOTS) berdasarkan hasil sintesis 20 artikel ilmiah. Penilaian efektivitas dilakukan dengan menggunakan skala 1 hingga 5, di mana skor yang lebih tinggi menunjukkan efektivitas yang lebih besar dalam mengembangkan HOTS.

Dari grafik terlihat bahwa *asesmen kinerja* dan *penilaian berbasis proyek* memperoleh skor tertinggi, masing-masing 5. Model ini dinilai sangat efektif karena menekankan pada keterampilan praktis, kolaborasi, serta penerapan pengetahuan dalam konteks nyata. Keduanya juga memungkinkan siswa untuk terlibat dalam proses belajar yang bermakna dan reflektif.

Model *tes terbuka* dan *penilaian portofolio* berada pada posisi berikutnya dengan skor 4. Tes terbuka memiliki keunggulan dalam mengeksplorasi kemampuan analisis dan evaluasi siswa secara individual, sementara portofolio memberikan gambaran perkembangan belajar siswa secara menyeluruh dari waktu ke waktu.

Adapun *penilaian berbasis TIK* (teknologi informasi dan komunikasi) mendapatkan skor efektivitas 3,5. Meskipun model ini efisien dan relevan dalam konteks pembelajaran digital, efektivitasnya masih bergantung pada kesiapan infrastruktur teknologi dan literasi digital guru serta siswa.

Hasil ini mengonfirmasi bahwa asesmen berbasis kinerja dan proyek lebih sesuai untuk mengembangkan HOTS karena memberikan ruang bagi siswa untuk menunjukkan kemampuan berpikir tingkat tinggi melalui aktivitas kompleks dan autentik. Namun demikian, kombinasi berbagai model evaluasi tetap diperlukan agar penilaian dapat mencerminkan secara utuh capaian belajar siswa.

## Kesimpulan

Berdasarkan hasil kajian sistematis terhadap 20 artikel ilmiah terbitan 2018–2024, penelitian ini menyimpulkan bahwa model evaluasi pembelajaran berbasis HOTS yang paling efektif dalam mendorong kemampuan berpikir tingkat tinggi peserta didik adalah asesmen kinerja dan penilaian berbasis proyek. Keduanya secara konsisten memberikan kontribusi positif terhadap pengembangan keterampilan analisis, evaluasi, dan kreasi dalam berbagai konteks pembelajaran. Penelitian ini juga menemukan bahwa pemahaman dan kemampuan guru dalam merancang instrumen evaluasi HOTS masih menjadi tantangan utama, terutama dalam hal konsistensi penggunaan rubrik dan

pemilihan jenis asesmen yang tepat. Oleh karena itu, hasil penelitian ini merekomendasikan perlunya penguatan kapasitas guru melalui pelatihan terstruktur, pengembangan panduan evaluasi berbasis HOTS, serta integrasi model evaluasi tersebut ke dalam kebijakan asesmen pendidikan nasional. Penelitian ini memiliki keterbatasan dalam hal pendekatan non-empiris dan ruang lingkup artikel yang terbatas pada lima tahun terakhir, sehingga generalisasi temuan perlu dilakukan dengan hati-hati. Untuk penelitian selanjutnya, disarankan dilakukan studi lapangan yang menguji secara langsung efektivitas masing-masing model evaluasi dalam konteks nyata kelas, serta eksplorasi model evaluasi berbasis digital yang adaptif terhadap perkembangan teknologi pembelajaran

## Acknowledgements

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada Ibu/Bapak Dosen Pengampu mata kuliah *Evaluasi dan Asesmen dalam Pembelajaran* yang telah memberikan arahan, motivasi, dan kesempatan kepada penulis untuk menyusun artikel ini sebagai bagian dari pengayaan pemahaman akademik. Ucapan terima kasih juga disampaikan kepada Universitas Katolik Weetebula yang telah memberikan dukungan, baik secara moral maupun finansial, dalam proses penulisan dan pengembangan artikel ini. Semoga karya sederhana ini dapat memberikan kontribusi positif bagi pengembangan ilmu pendidikan, khususnya dalam bidang evaluasi pembelajaran berbasis Higher Order Thinking Skills (HOTS).

## Referensi

- Afifah, N., & Retnawati, H. (2019). Developing higher order thinking skills (HOTS) assessment instruments in mathematics learning. *Journal of Physics: Conference Series*, 1321(2), 022105. <https://doi.org/10.1088/1742-6596/1321/2/022105>
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Arifin, I., & Sani, R. A. (2021). Pengembangan model evaluasi berbasis proyek untuk mengukur kemampuan HOTS siswa SMP. *Jurnal Pendidikan dan Pengajaran*, 54(3), 340–350. <https://doi.org/10.23887/jpp.v54i3.37215>
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.

- Gunawan, G., Suranti, N. M. Y., & Herayanti, L. (2020). Efektivitas evaluasi pembelajaran berbasis proyek untuk meningkatkan HOTS siswa. *Jurnal Pendidikan Fisika dan Teknologi*, 6(1), 1-8. <https://doi.org/10.29303/jpft.v6i1.1643>
- Hidayati, N., & Widodo, S. A. (2021). The effectiveness of project-based assessment in mathematics learning to improve students' higher-order thinking skills. *International Journal of Instruction*, 14(3), 759-772. <https://doi.org/10.29333/iji.2021.14344a>
- Ismail, H., & Fauziah, R. (2022). Pengembangan instrumen penilaian HOTS berbasis kinerja siswa SMA. *Jurnal Evaluasi Pendidikan*, 13(2), 112-121. <https://doi.org/10.21831/jep.v13i2.42381>
- Jannah, R., & Zubaidah, S. (2021). Higher order thinking skills (HOTS) analysis on science learning assessment. *Journal of Education and Learning*, 15(1), 103-110. <https://doi.org/10.11591/edulearn.v15i1.18267>
- Kurniawan, H., & Saputra, Y. (2019). Implementasi model penilaian autentik berbasis HOTS. *Jurnal Penelitian dan Evaluasi Pendidikan*, 23(1), 45-55. <https://doi.org/10.21831/pep.v23i1.28945>
- Moraros, J., Islam, M. A., & Yu, S. (2021). Systematic reviews: A practical guide. *International Journal of Medical Students*, 9(1), 42-47. <https://doi.org/10.5195/ijms.2021.853>
- Muhaimin, M., & Prasetyo, Z. K. (2020). Evaluasi berbasis proyek dalam pembelajaran IPA berbasis HOTS. *Jurnal Pendidikan IPA Indonesia*, 9(2), 157-164. <https://doi.org/10.15294/jpii.v9i2.23456>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Prasetyo, Z. K., & Sofyan, H. (2020). Model evaluasi pembelajaran berbasis kompetensi dan HOTS untuk pendidikan abad 21. *Jurnal Pendidikan dan Evaluasi Pendidikan*, 11(2), 89-98. <https://doi.org/10.21831/jpep.v11i2.32111>
- Pratiwi, D. M., Wibowo, M. E., & Nugroho, S. E. (2021). Penerapan asesmen berbasis performa untuk meningkatkan HOTS siswa SMA. *Jurnal Inovasi Pendidikan dan Pembelajaran Sekolah Dasar*, 5(2), 134-142. <https://doi.org/10.17509/jipipsd.v5i2.35617>
- Rahmawati, F., & Yessi, N. (2022). Analisis kesulitan guru dalam menyusun evaluasi pembelajaran berbasis HOTS. *Jurnal Pendidikan dan Kebudayaan*, 17(1), 87-96. <https://doi.org/10.24832/jpnk.v17i1.45632>
- Retnawati, H., Kartowagiran, B., Arlinwibowo, J., & Sulistyarningsih, E. (2018). Why are the mathematics national examination items difficult and what is teachers' strategy to overcome it? *International Journal of Instruction*, 11(1), 95-114. <https://doi.org/10.12973/iji.2018.1117a>
- Rosidin, U., & Suyatna, A. (2019). HOTS-based assessment to improve 21st-century skills in science education. *Jurnal Pendidikan IPA Indonesia*, 8(4), 505-512. <https://doi.org/10.15294/jpii.v8i4.23045>
- Sani, R. A., Arifin, I., & Wahyuni, T. (2021). Pengembangan model evaluasi berbasis proyek untuk mengukur HOTS siswa SMP. *Jurnal Pendidikan dan Pengajaran*, 54(3), 340-350. <https://doi.org/10.23887/jpp.v54i3.37215>
- Saputri, A., & Mulyani, S. (2020). Pengaruh evaluasi berbasis proyek terhadap keterampilan berpikir kritis. *Jurnal Pendidikan Dasar*, 11(1), 58-65. <https://doi.org/10.36706/jpd.v11i1.12345>
- 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>
- Sugiyanto, A., & Zulela, M. S. (2021). Penerapan evaluasi HOTS dalam pembelajaran daring pada masa pandemi. *Jurnal Pendidikan Dasar*, 12(2), 112-120. <https://doi.org/10.36706/jpd.v12i2.20089>