

# Menganalisis Kemampuan Mahasiswa dalam Pengembangan Instrumen Tes dengan Taksonomi Bloom Revisi

\*Munali Munali, Siti Alifah, Syifa Afidah Nurul Arifin

Universitas Indraprasta PGRI, Jakarta, Indonesia

\*Email: [munali2206@gmail.com](mailto:munali2206@gmail.com) (Corresponding Author)



DOI: <https://doi.org/10.53621/jider.v6i4.811>

## Informasi Artikel

### Riwayat Artikel:

Diterima: 22 Mei 2026

Revisi Akhir: 16 Agustus 2026

Disetujui: 20 Agustus 2026

Terbit: 31 Agustus 2026

### Kata Kunci:

Higher Order Thinking Skills;  
Instrumen Tes;

Lower Order Thinking Skills;  
Taksonomi Bloom Revisi.



## ABSTRAK

Tujuan penelitian ini adalah menganalisis kemampuan mahasiswa dalam mengembangkan instrumen tes berdasarkan Taksonomi Bloom revisi. Penelitian ini menggunakan metodologi kuantitatif dengan analisis data menggunakan statistik deskriptif. Sampel penelitian adalah mahasiswa pascasarjana Universitas Indraprasta PGRI Program Studi Magister Pendidikan Bahasa Indonesia semester gasal tahun akademik 2024/2025 sebanyak 29 mahasiswa pada mata kuliah Tes dan Pengukuran. Hasil penelitian: 1) kemampuan mahasiswa dalam pengembangan instrumen tes berdasarkan Taksonomi Bloom revisi dengan kategori sangat baik sebesar 83%, 2) rerata kemampuan mahasiswa dalam pengembangan instrumen tes sebesar 85,28 dengan kategori sangat baik, dan 3) terdapat 13 mahasiswa atau 45% yang belum memahami penulisan butir soal berbasis HOTS, yaitu: C4 (menganalisis), C5 (mengevaluasi), dan C6 (menciptakan). Temuan ini menunjukkan bahwa kemampuan mahasiswa dalam menulis butir soal berbasis HOTS belum optimal. Solusi yang dapat dilakukan bagi mahasiswa atau guru dalam pengembangan butir soal berbasis HOTS adalah membiasakan menulis butir soal berbasis (HOTS) dalam kegiatan ulangan harian, ujian tengah semester, dan ujian akhir semester, sehingga pengalaman dan kemampuan guru dalam menulis butir soal dapat meningkat.

## PENDAHULUAN

Permendiknas Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan merupakan proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar peserta didik (Kurniawati, 2020). Penilaian merupakan salah satu komponen penting dalam pembelajaran. Penilaian memberikan informasi terkait ketercapaian tujuan pembelajaran, perkembangan kemampuan peserta didik, dan efektivitas pembelajaran yang telah dilaksanakan. Penilaian merupakan proses yang dilakukan untuk mengetahui kualitas pembelajaran dengan menggunakan instrumen tertentu (Nasrum, 2020). Hasil penilaian tidak hanya berfungsi untuk menentukan hasil akhir peserta didik, tetapi juga menjadi dasar bagi guru dalam memperoleh informasi terkait penguasaan kompetensi, mengidentifikasi kesulitan belajar, memberikan umpan balik, dan menentukan tindak lanjut pembelajaran (Phafiandita et al., 2022). Penilaian tidak hanya meliputi pengumpulan data siswa, tetapi juga memperoleh gambaran tentang proses dan hasil belajar siswa (Rivaldi et al., 2025).

Penilaian merupakan komponen penting dalam proses pendidikan karena berfungsi memberikan informasi mengenai ketercapaian kompetensi peserta didik, efektivitas proses pembelajaran, serta menjadi dasar pengambilan keputusan dalam menentukan tindak lanjut pembelajaran (Hamdi et al., 2022). Penilaian yang berkualitas membutuhkan instrumen yang mampu merepresentasikan kompetensi yang hendak diukur secara akurat. Oleh karena itu, instrumen tes tidak dapat dipandang hanya sebagai kumpulan pertanyaan yang diberikan kepada peserta didik, tetapi sebagai alat ukur yang harus dikembangkan berdasarkan konstruk, indikator, tujuan pembelajaran, karakteristik materi, dan tingkat proses kognitif yang telah ditetapkan. Kualitas instrumen akan menentukan kualitas informasi yang diperoleh guru terkait kemampuan peserta didik, sehingga kemampuan mengembangkan instrumen merupakan salah

satu kompetensi penting yang perlu dimiliki oleh guru. Untuk mencapai kualitas instrumen tersebut, guru membiasakan menulis butir soal berbasis *Higher Order Thinking Skills* (HOTS) dalam kegiatan ulangan harian, ujian tengah semester, dan ujian akhir semester, sehingga pengalaman dan kemampuan guru dalam menulis butir soal dapat meningkat.

Selain kemampuan menulis butir soal, pengembang instrumen juga perlu memahami aspek validitas dan reliabilitas (Hadiyanti & Susongko, 2024). Validitas berkaitan dengan sejauh mana bukti yang diperoleh mendukung interpretasi skor sesuai tujuan pengukuran, sedangkan reliabilitas berkaitan dengan konsistensi hasil pengukuran. Dalam pengembangan instrumen pendidikan, validitas isi menjadi tahap awal yang penting karena memastikan bahwa butir instrumen telah merepresentasikan konstruk dan indikator yang ditetapkan. Proses validasi dapat dilakukan melalui *expert judgment* atau telaah ahli. Selanjutnya, instrumen dapat diuji secara empiris untuk mengetahui karakteristik butir dan kualitas pengukuran. Temuan penelitian menunjukkan perlunya pengembangan instrumen berbasis HOTS bagi guru.

Kualitas penilaian berpengaruh terhadap instrumen yang digunakan (Magdalena et al., 2020). Instrumen tes yang digunakan untuk mengukur hasil belajar berkaitan dengan penguasaan bahan pelajaran yang sesuai dengan tujuan pembelajaran (Fitrianty et al., 2022). Kebutuhan akan instrumen penilaian yang berkualitas semakin meningkat seiring dengan tuntutan pendidikan abad ke-21. Peserta didik tidak cukup memiliki kemampuan mengingat informasi dan memahami konsep, tetapi perlu menggunakan pengetahuan untuk menganalisis permasalahan, memecahkan masalah, mengevaluasi informasi, mengambil keputusan, dan menghasilkan ide atau produk baru. Perubahan orientasi tersebut menuntut perubahan dalam praktik penilaian, karena penilaian pada tingkat mengingat, memahami, dan mengaplikasikan belum mampu memberikan gambaran yang komprehensif terkait HOTS (Eliza et al., 2022).

HOTS menjadi salah satu kemampuan yang penting untuk dikembangkan melalui pembelajaran dan penilaian. HOTS berkaitan dengan kemampuan peserta didik dalam menganalisis, mengevaluasi dan menciptakan. Butir soal yang menggunakan kata kerja: menganalisis, mengevaluasi, dan menciptakan belum tentu termasuk kategori HOTS, jika penyelesaiannya melalui proses mengingat informasi, memahami, dan menerapkan prosedur rutin. Dengan demikian, butir soal HOTS harus dilihat dari aktivitas kognitif yang benar-benar diperlukan peserta didik untuk merespons butir soal, bukan dari kata kerja operasional yang digunakan dalam menulis butir soal. Temuan penelitian menunjukkan bahwa pengembangan instrumen berbasis HOTS memenuhi kriteria: validitas, reliabilitas, daya pembeda dan tingkat kesukaran (Miladanta et al., 2024).

Pengembangan instrumen merupakan proses yang membutuhkan keterpaduan antara landasan teoritis dan pengalaman menulis butir soal. Instrumen tidak cukup dikembangkan berdasarkan pengalaman pengajar atau kemampuan penulis pertanyaan, tetapi perlu diawali dengan identifikasi kompetensi yang akan diukur. Kompetensi tersebut selanjutnya diterjemahkan ke dalam indikator yang dapat diamati melalui respons peserta didik. Oleh sebab itu, pengembang instrumen tes harus memahami hubungan antara konstruk, variabel, indikator, menulis stimulus, menulis butir soal, kunci jawaban, dan rubrik penskoran. Ketidaksesuaian pada salah satu komponen dapat menyebabkan instrumen tidak sepenuhnya merepresentasikan kemampuan yang menjadi tujuan pengukuran. Proses pengembangan instrumen dengan demikian merupakan kegiatan akademik yang membutuhkan kemampuan analitis dan sistematis (Miladanta et al., 2024).

Pada penelitian ini, pengembangan instrumen yang dilakukan adalah pengembangan instrumen tes. Pengembangan instrumen tes yang sering digunakan berdasarkan Taksonomi Bloom revisi. Taksonomi Bloom revisi mempunyai enam dimensi: mengingat (C1), memahami (C2), menerapkan (C3), menganalisis (C4), mengevaluasi (C5), dan menciptakan (C6). Dimensi C1 sampai C3 disebut *Lower Order Thinking Skills* (LOTS) dan dimensi C4 sampai C6 disebut *Higher Order Thinking Skills* (HOTS). Oleh karena itu, pemahaman terhadap dimensi tersebut menjadi dasar penting bagi guru dan calon guru dalam mengembangkan instrumen yang sesuai dengan tujuan pengukuran. Temuan penelitian menunjukkan bahwa pengembangan instrumen

berbasis HOTS dengan model Tessmer mempunyai kualitas baik dan layak digunakan (Nugrahnastiti & Kamaludin, 2024).

Pengembangan instrumen tes berdasarkan Taksonomi Bloom revisi membutuhkan kemampuan menulis kisi-kisi instrumen: mengonstruksi variabel, dimensi variabel, indikator, menulis butir, menulis stimulus, dan menentukan rubrik penskoran. Ketidaksesuaian dimensi dan indikator dapat menyebabkan instrumen tidak mengukur kompetensi yang dimaksud. Oleh karena itu, kemampuan mengembangkan instrumen membutuhkan teori pengukuran dan pengalaman praktis dalam menulis butir soal (Hadiyanti & Susongko, 2024).

Mengonstruksi variabel menjadi dasar penting karena kemampuan yang hendak diukur dalam pendidikan pada umumnya bersifat abstrak. Oleh karena itu, dalam pengembangan instrumen perlu dimiliki kemampuan membaca dan menganalisis teori sebelum menentukan dimensi dan indikator pengukuran. Kemampuan tersebut diperlukan agar dimensi dan indikator yang dikembangkan tidak sekadar merupakan daftar aktivitas, tetapi benar-benar memiliki keterkaitan dengan konstruk yang akan diukur. Dengan demikian, pengembangan instrumen juga menjadi sarana untuk melatih kemampuan mahasiswa dalam menghubungkan konsep teoritis dengan praktik pengukuran. Temuan penelitian menunjukkan bahwa pengembangan instrumen tes menggunakan model Borg and Gall memenuhi kriteria validitas, reliabilitas, daya pembeda, dan tingkat kesukaran butir (Rahmawati et al., 2022).

Setelah konstruksi variabel ditetapkan, proses berikutnya adalah mengembangkan dimensi dan indikator. Dimensi diperlukan ketika suatu variabel memiliki beberapa aspek yang berbeda, sedangkan indikator digunakan untuk menjelaskan karakteristik yang dapat diamati dari setiap dimensi. Ketepatan dalam menentukan dimensi dan indikator akan menentukan kisi-kisi dan butir soal yang akan dikembangkan. Apabila indikator terlalu luas, pengembang akan mengalami kesulitan dalam menentukan butir soal dan respons yang sesuai. Sebaliknya, indikator yang terlalu sempit dapat menyebabkan instrumen hanya mengukur sebagian kecil dari konstruk. Oleh karena itu, mahasiswa perlu diberikan latihan untuk mengevaluasi apakah indikator yang disusun telah spesifik, relevan, terukur, dan konsisten dengan definisi konstruk.

Setelah dimensi dan indikator ditetapkan, ditulis kisi-kisi instrumen. Melalui kisi-kisi instrumen, pengembang dapat menentukan: indikator, butir soal, nomor butir soal, level kognitif, stimulus, kunci jawaban, dan rubrik penskoran. Dalam pengembangan instrumen HOTS, kisi-kisi juga berfungsi untuk memastikan bahwa soal yang dibuat benar-benar tersebar pada kompetensi yang ditetapkan dan tidak hanya terkonsentrasi pada satu level (Rumtini et al., 2022).

Tahap penulisan butir soal menjadi tantangan tersendiri karena pengembang instrumen harus menerjemahkan indikator ke dalam butir soal yang mampu memunculkan proses kognitif sesuai tujuan pengukuran. Hasil penelitian menunjukkan bahwa pengembangan instrumen berbasis HOTS model ADDIE memenuhi kriteria validitas, reliabilitas, daya pembeda, dan tingkat kesukaran butir soal dengan rata-rata kemampuan HOTS siswa kategori cukup (Siagian & Sagala, 2021).

Aspek penting lainnya dalam pengembangan instrumen tes berbasis HOTS adalah menulis stimulus. Stimulus berfungsi untuk memberikan informasi yang dapat digunakan peserta didik dalam proses berpikir. Stimulus dapat berupa: teks, tabel, grafik, data, gambar, kasus, fenomena atau permasalahan kontekstual. Oleh karena itu, kemampuan menulis stimulus menjadi bagian penting dalam pengembangan instrumen HOTS (Nurmaliati et al., 2024).

Kualitas stimulus juga menentukan apakah butir soal benar-benar mengukur HOTS (Nurmaliati et al., 2024). Stimulus yang terlalu sederhana dan tidak membutuhkan pengolahan informasi tidak akan menghasilkan proses kognitif tingkat tinggi meskipun pertanyaannya menggunakan kata kerja menganalisis. Sebaliknya, stimulus yang kaya informasi dapat memberikan kesempatan kepada peserta didik untuk membandingkan, mengelompokkan, menghubungkan, menilai, dan menghasilkan solusi. Dengan demikian, guru maupun calon guru perlu memiliki kemampuan untuk merancang stimulus yang relevan dengan indikator dan level kognitif yang ditargetkan. Pengembangan kemampuan tersebut membutuhkan latihan yang berulang serta proses telaah terhadap kualitas soal yang telah dikembangkan.

Karakteristik butir soal HOTS juga berkaitan dengan kemampuan pengembangan butir soal non rutin. Butir soal rutin dapat diselesaikan hanya dengan mengingat contoh, sehingga dapat membatasi proses berpikir peserta didik. Sebaliknya, butir soal non rutin yang membutuhkan penerapan konsep pada konteks berbeda dapat memberikan kesempatan bagi peserta didik untuk melakukan penalaran. Oleh karena itu, mahasiswa perlu dilatih membedakan soal rutin dan nonrutin serta mengidentifikasi proses kognitif yang diperlukan untuk menyelesaikannya. Pengalaman tersebut dapat membantu mahasiswa memahami bahwa HOTS bukan sekadar tingkat kesulitan soal, melainkan kualitas proses berpikir yang diperlukan dalam menyelesaikan permasalahan (Siagian & Sagala, 2021).

Aspek lain yang perlu mendapatkan perhatian adalah validitas instrumen (Fadzillah Karira et al., 2023). Validitas tidak hanya berkaitan dengan apakah suatu soal terlihat baik, tetapi juga apakah bukti yang diperoleh mendukung penggunaan skor sesuai dengan tujuan pengukuran. Pada tahap awal, validitas isi dapat diperiksa melalui telaah ahli dengan mempertimbangkan kesesuaian materi, konstruk, indikator, bahasa, dan karakteristik butir. Setelah itu, pengujian empiris dapat dilakukan untuk memperoleh informasi mengenai validitas butir, reliabilitas, tingkat kesukaran, dan daya pembeda. Uji empiris memberikan informasi yang tidak dapat diperoleh hanya melalui telaah teoritis. Butir yang dianggap baik oleh pengembang atau ahli belum tentu menunjukkan karakteristik pengukuran yang baik ketika digunakan pada peserta didik. Sebaliknya, hasil analisis empiris dapat menunjukkan butir yang perlu dipertahankan, direvisi, atau dihilangkan.

Kemampuan melakukan validasi dan analisis instrumen menjadi penting bagi guru atau mahasiswa karena mereka merupakan calon pendidik sekaligus calon peneliti. Dalam penelitian pendidikan, instrumen sering digunakan sebagai sumber utama pengumpulan data. Apabila instrumen tidak memiliki kualitas yang memadai, kesalahan pengukuran dapat mempengaruhi interpretasi hasil penelitian. Oleh karena itu, mahasiswa perlu memahami bahwa pengembangan instrumen tidak berakhir ketika butir telah selesai ditulis. Mereka perlu memiliki kemampuan untuk menilai kembali instrumen melalui validasi, uji coba, dan analisis data. Kompetensi tersebut dapat membentuk sikap akademik bahwa setiap alat ukur perlu memperoleh bukti yang mendukung penggunaannya (Afifah Aulia Zayrin et al., 2025).

Permasalahan pengembangan instrumen juga dapat muncul ketika guru atau mahasiswa belum memiliki pengalaman yang cukup dalam mengubah teori menjadi produk pengukuran. Pengetahuan mengenai evaluasi pembelajaran secara konseptual belum tentu menghasilkan kemampuan menyusun instrumen secara otomatis. Mahasiswa perlu mengalami sendiri proses menentukan konstruk, mengembangkan indikator, membuat kisi-kisi, menulis butir, menyusun rubrik, serta menelaah kualitas instrumen. Penelitian Pujiastuti dan Kulup (2021) menunjukkan bahwa mahasiswa mengalami berbagai kendala ketika menyusun instrumen penilaian kognitif berbasis HOTS, termasuk ketidaksesuaian instrumen berbasis teks dan ketidaktepatan dalam membuat rubrik penilaian. Pembelajaran problem-based learning dan peer assessment dapat meningkatkan pengembangan instrumen berbasis HOTS (Pujiastuti & Kulup, 2021).

Temuan tersebut menunjukkan bahwa pembelajaran pengembangan instrumen perlu memberikan pengalaman praktik yang lebih intensif. Mahasiswa tidak cukup hanya diberi contoh instrumen yang sudah jadi, tetapi perlu dilibatkan dalam proses menghasilkan instrumen secara bertahap. Melalui kegiatan tersebut, mahasiswa dapat menemukan kesulitan yang sebenarnya muncul ketika teori diterapkan. Mereka dapat mengetahui bahwa menentukan indikator, misalnya, tidak selalu mudah; demikian pula, mengubah indikator menjadi soal C5 atau C6 membutuhkan pertimbangan stimulus, konteks, respons, dan rubrik. Pengalaman tersebut dapat meningkatkan pemahaman karena mahasiswa belajar melalui proses menghasilkan dan mengevaluasi produk mereka sendiri (Pujiastuti & Kulup, 2021).

Salah satu pendekatan yang dapat digunakan adalah pembelajaran berbasis masalah yang dipadukan dengan peer assessment (Pujiastuti & Kulup, 2021). Dalam pendekatan tersebut, mahasiswa menghadapi persoalan autentik berupa kebutuhan untuk menghasilkan instrumen yang memenuhi kriteria tertentu. Setelah instrumen dikembangkan, mahasiswa dapat menelaah karya teman berdasarkan rubrik yang disediakan. Proses tersebut memungkinkan mahasiswa

memperoleh perspektif berbeda mengenai kualitas instrumen. Peer assessment juga memberikan kesempatan kepada mahasiswa untuk mengembangkan kemampuan mengevaluasi. Ketika mahasiswa menilai instrumen yang dikembangkan oleh teman, mereka harus menentukan apakah konstruk, indikator, kisi-kisi, butir, dan rubrik telah sesuai. Kegiatan tersebut secara tidak langsung menempatkan mahasiswa pada proses kognitif tingkat tinggi karena mereka harus menganalisis dan mengevaluasi kualitas suatu produk. Dengan demikian, kegiatan peer assessment memiliki dua fungsi sekaligus, yaitu sebagai strategi pembelajaran dan sebagai sarana untuk meningkatkan kemampuan mahasiswa dalam melakukan telaah instrumen.

Pengembangan instrumen berbasis HOTS bagi guru atau mahasiswa tersebut menjadi penting karena instrumen yang digunakan guru atau mahasiswa akan mempengaruhi informasi yang diperoleh terkait kemampuan peserta didik (Prima Kurniawati & Rahmantika Hadi, 2021). Jika instrumen lebih banyak mengukur kemampuan mengingat, memahami, dan mengaplikasikan, maka informasi tentang kemampuan menganalisis, mengevaluasi, dan menciptakan menjadi terbatas. Sebaliknya, instrumen yang dirancang secara proporsional berdasarkan tujuan pembelajaran dapat memberikan informasi yang lebih lengkap terkait kemampuan peserta didik.

Kondisi kemampuan guru dalam pengembangan instrumen tes berbasis HOTS menjadi salah satu persoalan yang perlu mendapat perhatian. Guru memiliki posisi strategis, karena instrumen yang digunakan dalam pembelajaran sehari-hari sebagian besar dikembangkan oleh guru. Apabila guru mempunyai kemampuan dalam pengembangan instrumen yang baik, peserta didik akan memperoleh pengalaman dalam merespons butir soal berbasis HOTS. Sebaliknya, apabila instrumen yang digunakan berbasis LOTS, kesempatan peserta didik menjadi terbatas (Billa et al., 2024).

Pengembangan instrumen memerlukan rubrik penskoran terutama pada butir soal uraian atau tugas yang menuntut peserta didik menghasilkan solusi. Hal tersebut penting, karena instrumen berbasis HOTS membutuhkan respons yang kompleks sehingga proses penskoran harus dirancang dengan jelas. Rubrik penskoran dapat membantu menjaga konsistensi penskoran dan memberikan gambaran terkait kualitas respons peserta didik (Kristiana Wabang et al., 2026).

Kesenjangan penelitian tersebut dapat terlihat dari perbedaan antara kemampuan memahami teori dan kemampuan menerapkan teori. Mahasiswa telah mempelajari konsep variabel, dimensi, indikator, validitas, reliabilitas, dan Taksonomi Bloom revisi, tetapi penerapan konsep tersebut dalam menghasilkan instrumen berkualitas membutuhkan kemampuan keduanya. Mahasiswa harus mampu menghubungkan teori dan praktik serta menerapkan teori. Mereka perlu menentukan teori yang tepat, mengidentifikasi konstruk, menjabarkan konstruk menjadi dimensi dan indikator, indikator menjadi butir soal, level kognitif yang digunakan, kunci jawaban, dan rubrik (Hanik et al., 2021).

Pada penelitian ini, kemampuan pengembangan instrumen dipandang sebagai rangkaian kompetensi yang saling berkaitan. Tahap pertama adalah kemampuan mencari teori yang relevan untuk menentukan konstruk. Tahap kedua kemampuan mengkonstruksi variabel berdasarkan teori. Tahap ketiga adalah menentukan dimensi dan indikator. Tahap keempat adalah menulis kisi-kisi instrumen. Tahap kelima adalah menulis butir soal berdasarkan Taksonomi Bloom revisi. Tahap keenam adalah menulis rubrik penskoran. Tahap ketujuh adalah melakukan validasi konten melalui teman sejawat atau ahli. Setiap tahapan memberikan informasi terkait penguasaan mahasiswa terhadap konsep dan praktik pengukuran (Chasanah et al., 2022).

Kemampuan pada setiap tahapan tersebut perlu dianalisis secara terpisah karena mahasiswa dapat memiliki pola kemampuan yang berbeda. Seorang mahasiswa mungkin mampu menemukan teori yang relevan, tetapi mengalami kesulitan dalam mengajarkannya menjadi indikator. Mahasiswa lain mungkin mampu membuat indikator dan kisi-kisi, tetapi mengalami kesulitan dalam menulis butir soal C5 dan C6. Mahasiswa juga mampu menulis soal, tetapi belum mampu membuat rubrik yang sesuai. Oleh karena itu, analisis kemampuan secara

komprehensif dapat memberikan informasi yang lebih bermanfaat dibandingkan dengan hanya menggunakan satu skor keseluruhan (Putu Agus Adi Saputra et al., 2021).

Permasalahan tersebut menjadi semakin relevan bagi mahasiswa pascasarjana yang mengikuti mata kuliah Tes dan Pengukuran. Mahasiswa pada jenjang tersebut diharapkan tidak hanya memahami konsep evaluasi secara teoritis, tetapi juga mampu mengembangkan instrumen yang dapat digunakan dalam penelitian maupun praktik pendidikan. Kemampuan tersebut penting karena instrumen yang dikembangkan dalam penelitian pendidikan akan menentukan kualitas data yang diperoleh. Kesalahan pada tahap konstruksi dan pengembangan butir dapat menyebabkan data penelitian tidak merepresentasikan variabel yang hendak diukur. Dengan demikian, penguasaan pengembangan instrumen merupakan kompetensi akademik sekaligus profesional bagi mahasiswa pendidikan.

Keterlibatan mahasiswa dalam penelitian ini dapat memberikan pengalaman dalam pengembangan instrumen antara teori dan praktik. Mahasiswa tidak hanya mempelajari teori pengukuran melalui buku atau perkuliahan, tetapi juga mengaplikasikan teori tersebut dalam sebuah produk instrumen. Proses tersebut memberikan kesempatan kepada mahasiswa untuk menemukan kesalahan, mendapatkan umpan balik, melakukan revisi, dan memahami konsekuensi dari keputusan yang mereka ambil ketika mengembangkan instrumen. Pembelajaran berbasis pengalaman seperti ini dapat memperkuat kompetensi pengembangan asesmen secara lebih bermakna.

Pengembangan kompetensi mahasiswa dalam pengembangan instrumen perlu dilakukan secara sistematis, karena kualitas instrumen sangat ditentukan oleh ketepatan setiap tahapan pengembangan (Litna et al., 2021). Mahasiswa tidak cukup memahami teori Taksonomi Bloom revisi, tetapi dapat mengaitkan tujuan pengukuran dengan konstruk, dimensi, indikator, stimulus, soal, dan pedoman penskoran. Pengembangan instrumen pada dasarnya meliputi: proses perencanaan, penulisan, validasi oleh ahli, ujicoba, dan revisi agar diperoleh alat ukur yang valid dan reliabel.

Pengembangan instrumen perlu dilakukan melalui tahapan sistematis, yaitu: analisis kebutuhan sampai evaluasi produk (Imania & Bariah, 2019). Dengan demikian, pengalaman mahasiswa dalam pengembangan instrumen pada mata kuliah tes dan pengukuran dapat lebih baik lagi serta dapat diterapkan pada instansi mereka. Pemahaman terhadap level kognitif pada Taksonomi Bloom revisi perlu ditulis melalui karakteristik butir soal yang sesuai dengan proses berpikir yang akan diukur. Terutama pemahaman yang mendalam terkait dimensi menganalisis, mengevaluasi, dan menciptakan.

Pengembangan instrumen berbasis HOTS pada beberapa penelitian menunjukkan bahwa kualitas butir soal diperoleh dari proses pengembangan dan pengujian, seperti pengujian validitas, reliabilitas, dan validasi konten (Hulwani & Yuhana, 2024). Kondisi tersebut sangat perlu diperhatikan oleh mahasiswa, bukan saja dari pengembangan instrumen yang ditulis, tetapi juga dari kesesuaian antara dimensi, indikator, level kognitif, penulisan butir soal, dan penulisan stimulus.

Mahasiswa dapat saling menelaah instrumen berdasarkan kriteria yang telah ditentukan. Melalui kegiatan tersebut, mahasiswa belajar mengidentifikasi kesesuaian antara indikator dan butir soal, level kognitif, menulis stimulus, konstruksi pertanyaan, penulisan bahasa Indonesia baku, dan rubrik penskoran. Proses telaah sejawat memberikan kesempatan kepada mahasiswa untuk belajar tidak hanya dari instrumen yang dibuat sendiri, tetapi juga dari kesalahan dan kelebihan instrumen mahasiswa lain. Praktik telaah dan revisi semacam ini sejalan dengan prinsip bahwa pengembangan instrumen merupakan proses iteratif yang membutuhkan evaluasi dan penyempurnaan (Desiriah et al., 2021).

Salah satu strategi yang dapat digunakan untuk meningkatkan kompetensi dalam pengembangan instrumen tes adalah pembelajaran berbasis proyek. Melalui proyek pengembangan instrumen tes, mahasiswa dapat menyelesaikan seluruh tahapan pengembangan secara terintegrasi. Dosen dapat memberikan tugas untuk memilih variabel, mencari teori, mengembangkan konstruk, menyusun indikator, membuat kisi-kisi, mengembangkan butir C1-C6, membuat rubrik, kemudian melakukan validasi. Produk akhir tidak hanya berupa kumpulan

soal, tetapi juga sebuah instrumen yang memiliki landasan teoritis dan prosedur pengembangan yang jelas. Pendekatan tersebut sesuai dengan kebutuhan pengembangan kompetensi praktis calon pendidik dalam bidang penilaian (Agusdianita, 2023).

Pendampingan dosen juga diperlukan terutama pada butir soal: C4, C5, dan C6. Dosen dapat memberikan contoh bagaimana mengubah indikator menjadi stimulus dan pertanyaan yang benar-benar menuntut untuk menganalisis, mengevaluasi, dan menciptakan. Mahasiswa kemudian dapat diminta mengembangkan soal secara mandiri, mendapatkan umpan balik, melakukan revisi, dan menjelaskan alasan mengapa soal tersebut termasuk C4, C5, atau C6. Dengan demikian, mahasiswa tidak hanya menghasilkan produk, tetapi juga mampu mempertanggungjawabkan keputusan pedagogis dan pengukuran yang digunakan dalam pengembangan instrumen.

Kegiatan tersebut juga dapat diterapkan pada pengembangan profesional guru. Guru dapat mengikuti workshop yang berbasis produk, yaitu setiap peserta membawa instrumen yang biasa digunakan, kemudian melakukan analisis level kognitif, merevisi kisi-kisi, menulis stimulus, menulis ulang butir soal C4–C6, membuat rubrik, dan melakukan telaah sejawat. Pendekatan tersebut lebih memungkinkan terjadinya transfer kompetensi dibandingkan dengan pelatihan yang hanya memberikan penjelasan konseptual. Temuan penelitian menunjukkan bahwa pengalaman guru dalam menulis butir soal berbasis HOTS menjadi salah satu aspek yang perlu diperhatikan dalam meningkatkan kemampuan penulisan butir (Hamdi et al., 2022).

Pengembangan kompetensi guru atau mahasiswa pada akhirnya perlu diarahkan pada terciptanya budaya penilaian yang berkualitas. Instrumen tes tidak seharusnya hanya digunakan untuk memberikan nilai, tetapi juga untuk memberikan informasi mengenai bagaimana peserta didik berpikir. Instrumen yang baik dapat membantu guru mengetahui apakah peserta didik mampu menganalisis informasi, mengevaluasi alternatif, memberikan alasan, dan menghasilkan solusi. Informasi tersebut selanjutnya dapat digunakan untuk memperbaiki proses pembelajaran. Dengan demikian, kualitas instrumen memiliki hubungan yang erat dengan kualitas proses pembelajaran dan keputusan pedagogis yang ditulis oleh guru atau mahasiswa (Hadiyanti & Susongko, 2024). Beberapa kajian teoretik terkait pengembangan instrumen tes sebagai berikut:

Kemampuan kognitif siswa Indonesia dalam menyelesaikan butir soal Higher Order Thinking Skills (HOTS) berada pada kategori rendah dibandingkan dengan negara lain (OECD, 2023). Hal ini dapat terlihat dari hasil PISA (Programme for International Student Assessment) pada tahun 2022. Skor kemampuan membaca sebesar 482 dengan kategori Lower Order Thinking Skills (LOTS), skor kemampuan matematika sebesar 480 dengan kategori Lower Order Thinking Skills, dan skor kemampuan sains sebesar 491 dengan kategori Lower Order Thinking Skills. Jika dibandingkan dengan hasil PISA tahun 2018, kemampuan siswa Indonesia mengalami penurunan. Isu saat ini terkait rendahnya skor Tes Kemampuan Akademik (TKA) SMA 2025 (Choir, 2025). Walaupun TKA bukan penentu kelulusan, TKA didesain sebagai asesmen opsional berbasis standar nasional yang berfokus pada pengukuran capaian akademik murid secara objektif untuk seleksi prestasi, penyetaraan hasil belajar nonformal/informal, dan pengendalian mutu pendidikan (Sukanto et al., 2026). Skor TKA ini dipandang sebagai potret nasional kemampuan akademik siswa yang tidak dapat diabaikan. Butir soal TKA dirancang tidak hanya untuk menguji hafalan, tetapi juga untuk menguji kemampuan siswa menerapkan konsep dan bernalar. Rerata skor TKA SMA 2025 untuk matematika sebesar 36,1; skor bahasa Inggris sebesar 25,93; dan skor bahasa Indonesia sebesar 55,38. Masalah kemampuan kognitif siswa tersebut memiliki implikasi yang luas jika tidak segera diselesaikan. Beberapa kajian teoretik terkait pengembangan instrumen tes meliputi:

Penelitian terhadap 21 guru matematika SMP dalam mengembangkan instrumen berbasis HOTS. Hasil penelitian menunjukkan bahwa kualitas soal yang dihasilkan oleh guru masih memiliki kendala, terutama berkaitan dengan reliabilitas dan daya pembeda. Temuan tersebut menunjukkan bahwa kemampuan mengembangkan instrumen HOTS tidak hanya memahami konsep, tetapi juga membutuhkan pengalaman dan pembiasaan (Hamdi et al., 2022).

Penelitian terhadap 15 calon guru sekolah dasar dalam mengembangkan instrumen berbasis HOTS dengan tahapan pengembangan Borg and Gall. Hasil penelitian menunjukkan

bahwa instrumen yang dikembangkan memenuhi kriteria validitas dan reliabilitas, daya pembeda, dan tingkat kesukaran yang baik. Hasil penelitian tersebut menunjukkan bahwa calon dapat dilibatkan secara langsung dalam pengembangan instrumen berbasis HOTS. Temuan tersebut menunjukkan bahwa pengembangan instrumen perlu dimulai sejak mahasiswa (Rahmawati et al., 2022).

Pengembangan instrumen berbasis HOTS menggunakan Quizizz pada materi termokimia dengan menggunakan model Plomp, kemudian divalidasi berdasarkan aspek materi, konstruksi, dan bahasa. Hasil penelitian menunjukkan bahwa instrumen yang dikembangkan memenuhi kriteria validitas dan reliabilitas yang baik, sehingga Quizizz dapat digunakan sebagai media instrumen berbasis HOTS. Temuan tersebut memperkuat bahwa pengembangan instrumen melibatkan proses pengujian validasi isi instrumen.

Pengembangan instrumen berbasis HOTS pada materi elektrokimia siswa sekolah menengah. Hasil penelitian menunjukkan bahwa penerapan Taksonomi Bloom Revisi tidak dapat dilakukan secara mekanis dengan mengganti kata kerja operasional. Pengembangan instrumen harus memahami materi, karakteristik peserta didik, konteks permasalahan, dan proses kognitif yang dilakukan.

Hasil penelitian instrumen Mathematics Assessment Contruction Scale (MACS) yang diadaptasi oleh (Akinboboye et al., 2021) sebanyak 63 guru mata pelajaran matematika SMP di Nigeria: 1) sebagian besar menggunakan tes tertulis dan tugas dalam penilaian, 2) 16,4% menulis butir soal sendiri tanpa rujukan, 3) 80% guru menulis butir soal pada level kognitif mengingat dan memahami, dan 4) 0% guru tidak menulis butir soal pada level kognitif menerapkan, menganalisis, mengevaluasi, dan menciptakan.

Hasil penelitian pengembangan instrumen tes kemampuan kognitif berdasarkan Taksonomi Bloom revisi bentuk pilihan ganda (Nurmaya et al., 2022) meliputi: 1) instrumen mempunyai validitas yang baik berdasarkan penilaian ahli dengan rerata sebesar 93,65; 2) koefisien reliabilitas sebesar 0,87 dengan kategori tinggi; 3) tingkat kesukaran butir dalam kategori sedang, daya pembeda dengan kategori baik, dan distraktor berfungsi dengan baik.

Hasil penelitian pengembangan instrumen berbasis HOTS mempunyai kualitas sangat baik berdasarkan validasi ahli dengan skor 93% (Maxnun et al., 2024), meliputi: sebanyak 10 butir esai valid dan skor reliabilitas butir sebesar 0,863 dengan kategori tinggi. Tingkat kesukaran butir: 50% sedang, 50% sulit, daya pembeda dengan kategori baik, 11 siswa dapat menyelesaikan butir soal HOTS dengan kategori sangat baik, 17 siswa kategori baik, dan 2 siswa kategori cukup.

Hasil penelitian skor rerata kemampuan 28 mahasiswa calon guru prodi Pendidikan Matematika Universitas Maritim Raja Ali Haji dalam pengembangan instrumen tes berbasis HOTS sebesar 76,2% dengan kategori baik (Izzati & Febrian, 2021). Pengukuran pengembangan instrumen tes meliputi: 1) analisis KD, 2) pengembangan indikator, 3) penulisan kisi-kisi, 4) pengembangan butir soal, dan 5) analisis butir soal.

Hasil penelitian pengembangan instrumen tes uraian berbasis HOTS menggunakan model pengembangan Wilson, Oriondo, dan Antonio melalui tiga tahap: perencanaan, uji coba terbatas, dan pengukuran (Lara Juni Ulpa et al., 2025a). Pengembangan instrumen tes pemahaman kelas VII SMP terdiri atas 20 butir soal pilihan ganda dengan Taksonomi Bloom revisi dengan sampel 31 siswa (Safitri et al., 2025) dengan dimensi pemahaman meliputi: interpreting, exemplifying, classifying, summarizing, inferring, comparing, dan explaining.

Model pengembangan instrumen menurut Retnawati (2017): 1) menentukan tujuan, 2) mencari teori yang relevan, 3) menulis indikator, 4) menulis butir soal, 5) validasi isi, 6) revisi berdasarkan validator, 7) uji coba instrumen, 8) analisis data meliputi: reliabilitas, tingkat kesukaran, dan daya pembeda, dan 9) merakit instrumen.

Berdasarkan penelitian terdahulu, terdapat tiga persoalan utama yang perlu mendapatkan perhatian. Pertama, guru masih menghadapi kesulitan dalam menghasilkan soal HOTS yang benar-benar memenuhi karakteristik analisis, evaluasi, dan kreasi. Kedua, calon guru perlu memperoleh pengalaman yang lebih sistematis dalam mengembangkan instrumen, bukan hanya memahami konsep evaluasi secara teoritis. Ketiga, pengembangan instrumen harus dilakukan melalui proses yang mencakup konstruk, indikator, kisi-kisi, butir, rubrik, validasi, dan evaluasi

kualitas instrumen. Ketiga persoalan tersebut menunjukkan pentingnya penelitian yang secara khusus menganalisis kemampuan mahasiswa dalam mengembangkan instrumen tes berdasarkan Taksonomi Bloom revisi. Beberapa penelitian juga menunjukkan bahwa pengembangan instrumen dapat menghasilkan produk yang valid dan layak apabila dilakukan melalui tahapan sistematis. Namun, fokus penelitian-penelitian tersebut lebih banyak diarahkan pada produk instrumen dan karakteristik kualitasnya daripada pemetaan kemampuan pengembangan instrumen pada setiap tahapan pengembangan.

Berdasarkan kesenjangan tersebut, penelitian ini memfokuskan kajian pada kemampuan mahasiswa dalam mengembangkan instrumen tes berdasarkan Taksonomi Bloom revisi. Fokus penelitian tidak hanya pada kemampuan menulis butir soal, tetapi mencakup keseluruhan tahapan pengembangan instrumen, yaitu mencari teori yang relevan, mengonstruksi variabel, menentukan dimensi, menyusun indikator, membuat kisi-kisi, mengembangkan butir soal C1 sampai C6, menyusun rubrik, dan melakukan validasi konten. Fokus tersebut diharapkan dapat memberikan gambaran yang lebih komprehensif mengenai kompetensi mahasiswa dalam pengembangan instrumen tes.

Penelitian ini memiliki kontribusi penting karena hasilnya dapat digunakan untuk memperbaiki pembelajaran mata kuliah Tes dan Pengukuran. Jika diketahui bahwa mahasiswa mengalami kesulitan pada tahap tertentu, dosen dapat memberikan intervensi pembelajaran yang lebih spesifik. Misalnya, kesulitan dalam menentukan dimensi dapat ditindaklanjuti dengan latihan analisis teori dan konstruk, sedangkan kesulitan dalam menulis butir soal C4-C6 dapat ditindaklanjuti dengan menulis stimulus, menulis indikator, dan menelaah butir soal dengan teman sejawat. Dengan demikian, hasil penelitian tidak hanya menggambarkan kemampuan mahasiswa.

Kontribusi penelitian juga dapat diperluas ke pengembangan profesional guru. Mahasiswa yang memperoleh pengalaman dalam mengembangkan instrumen diharapkan memiliki bekal yang lebih baik ketika memasuki dunia pendidikan. Mereka dapat menerapkan tahapan yang sama dalam menyusun instrumen di sekolah, mulai dari menentukan tujuan pengukuran sampai melakukan evaluasi kualitas butir. Dengan demikian, pendidikan guru atau mahasiswa dapat menjadi salah satu titik awal untuk memperbaiki kualitas penilaian di sekolah. Penguatan kompetensi asesmen sejak pendidikan calon guru juga dapat mengurangi kesenjangan antara pemahaman teoritis mengenai HOTS dengan praktik penulisan butir soal.

Pada akhirnya, pengembangan instrumen tes bagi guru atau mahasiswa tidak cukup dilakukan melalui pemberian teori tentang Taksonomi Bloom, tetapi perlu dilaksanakan melalui proses pembelajaran dan pengembangan profesional yang berbasis praktik, proyek, pendampingan, telaah sejawat, validasi, revisi, dan analisis kualitas instrumen. Guru perlu memperoleh pengalaman berkelanjutan dalam mengembangkan butir soal C4, C5, dan C6, sedangkan mahasiswa perlu memperoleh pengalaman serupa sejak mengikuti mata kuliah Tes dan Pengukuran. Dengan demikian, pengembangan instrumen dapat bergeser dari aktivitas membuat pertanyaan menjadi kompetensi profesional untuk merancang alat ukur yang memiliki landasan konstruk yang jelas, indikator yang tepat, stimulus yang bermakna, butir yang sesuai dengan tingkat kognitif, rubrik yang terukur, serta bukti validitas yang valid.

Berdasarkan beberapa permasalahan tersebut, solusi yang dapat dilakukan untuk mahasiswa dan guru di Indonesia adalah memberikan pemahaman terkait pengembangan instrumen tes berbasis HOTS. Penelitian ini bertujuan untuk menganalisis kemampuan mahasiswa dalam pengembangan instrumen tes berdasarkan Taksonomi Bloom revisi, meliputi kemampuan mencari teori yang relevan, mengonstruksi dimensi, menulis indikator, menulis butir soal, menulis rubrik penskoran, dan melakukan validasi konten dengan teman sejawat atau dengan guru senior. Secara khusus, penelitian ini diharapkan dapat mengidentifikasi aspek pengembangan instrumen yang telah dikuasai guru atau mahasiswa, terutama berbasis HOTS.

## METODE PENELITIAN

Terdapat beberapa model penelitian dan pengembangan: model ADDIE: 1) *analysis*, 2) *design*, 3) *development*, 4) *implementation*, dan 5) *evaluation*. Model penelitian dan pengembangan

ADDIE terbukti efektif dalam pengembangan instrumen untuk proses pembelajaran berbasis proyek di perguruan tinggi (Handoko et al., 2025). Model ADDIE merupakan model paling populer dan fleksibel.

Model penelitian dan pengembangan Borg & Gall digunakan untuk mengembangkan suatu produk. Tahapan model Borg and Gall meliputi: 1) rumusan masalah, 2) pengumpulan data, 3) desain produk, dan 4) validasi desain (Anggermawan et al., 2017).

Model pembelajaran Dick and Carey digunakan untuk mendesain pembelajaran (Rukmi Octaviana et al., 2022). Terdapat sepuluh langkah model Dick and Carey yang meliputi: 1) mengidentifikasi tujuan pembelajaran, 2) melakukan analisis pembelajaran, 3) menganalisis karakteristik siswa dan konteks pembelajaran, 4) merumuskan tujuan pembelajaran, 5) mengembangkan instrumen penilaian, 6) mengembangkan strategi pembelajaran, 7) mengembangkan dan memilih bahan ajar, 8) merancang dan mengembangkan evaluasi formatif, 9) melakukan revisi terhadap program pembelajaran, dan 10) merancang dan mengembangkan evaluasi sumatif.

Model penelitian dan pengembangan Hannafin dan Peck (Nadzir, 2023). Model Hannafin and Peck merupakan model penelitian dan pengembangan dalam mendesain pembelajaran yang menghasilkan produk. Tahapan Hannafin dan Peck meliputi: 1) penelitian kebutuhan, 2) tahap desain, dan 3) tahap pengembangan dan implementasi. Ketiga tahap tersebut selanjutnya dievaluasi dan direvisi.

Model penelitian dan pengembangan: *Decide, Design, Develop*, dan *Evaluate* (DDD-E) adalah model yang digunakan untuk mengembangkan multimedia pembelajaran (Ayu et al., 2021). Langkah-langkah model DDD-E meliputi: 1) *decide* atau menetapkan tujuan dan materi program, 2) *design*, yaitu membuat struktur program, 3) *develop*, yaitu memproduksi elemen media dan membuat tampilan multimedia, dan 4) *evaluate*, yaitu mengevaluasi seluruh proses desain dan pengembangan.

Model penelitian dan pengembangan Bergman & Moore digunakan sebagai panduan dan manajemen produksi produk video dan multimedia interaktif (Ningrum et al., 2022). Langkah-langkah model Bergman & Moore meliputi: 1) analisis, 2) desain, 3) pengembangan, 4) produksi, 5) penggabungan, dan 6) validasi.

Model pengembangan instrumen menurut Retnawati (2017): 1) menentukan tujuan, 2) mencari teori yang relevan, 3) menulis indikator, 4) menulis butir soal, 5) validasi isi, 6) revisi berdasarkan validator, 7) uji coba instrumen, 8) analisis data meliputi: reliabilitas, tingkat kesukaran, dan daya pembeda, dan 9) merakit instrumen

Pada penelitian ini, peneliti menggunakan model pengembangan instrumen tes dengan mengadaptasi model (Retnawati, 2027) yang meliputi: 1) menentukan variabel penelitian, 2) mencari teori yang relevan atau pendapat ahli, 3) mengonstruksi/menulis definisi konseptual, 4) menulis dimensi, 5) menulis indikator, 6) menulis kisi-kisi instrumen, 7) menulis butir soal, 8) menulis rubrik, dan 9) validasi isi.

Penelitian ini menggunakan pendekatan metodologi penelitian kuantitatif. Sampel penelitian adalah mahasiswa pascasarjana Universitas Indraprasta PGRI Program Studi Magister Pendidikan Bahasa Indonesia semester gasal tahun akademik 2024/2025 sebanyak 29 mahasiswa pada mata kuliah Tes dan Pengukuran. Teknik pengumpulan data dilakukan melalui tugas mandiri yang dikumpulkan setelah pertemuan terakhir perkuliahan. Tugas mandiri melalui pengembangan instrumen tes, yaitu: Saudara diminta untuk mengembangkan instrumen tes bahasa dan sastra Indonesia. Variabel pengembangan instrumen tes dapat dipilih, yaitu: variabel keterampilan membaca, keterampilan mendengarkan, keterampilan berbicara, atau keterampilan menulis. Pengembangan instrumen tes disesuaikan dengan tempat saudara bertugas. Tugas meliputi: 1) menentukan variabel penelitian, 2) mencari teori relevan (pendapat ahli) dari variabel penelitian. Minimal 5 artikel dan referensi menggunakan aplikasi Mendeley! (referensi minimal 5 tahun terakhir antara tahun 2021 s.d. tahun 2025), 3) menulis konstruk dari masing-masing variabel yang dipilih, 4) menulis dimensi dari variabel yang dipilih, 5) menulis indikator/tujuan pembelajaran, 6) menulis kisi-kisi

instrumen, 7) menulis 10 butir soal dan kunci jawaban (pilihan ganda atau esai) berdasarkan Taksonomi Bloom revisi! (C1 = 1 butir, C2 = 2 butir, C3 = 2 butir, C4 = 2 butir, C5 = 2 butir, dan C6 = 1 butir). Untuk butir soal C4, C5, dan C6 diberikan stimulus, 8) Tulislah rubrik holistik (butir pilihan ganda) atau analitik (butir esai), dan 9) lakukan validasi konten dengan rekan sejawat.

Skala pengukuran yang digunakan adalah 0 sampai dengan 2. Skor 0 jika tidak memahami, skor 1 jika belum memahami kurang dari 100%, dan skor 2 jika memahami 100%.

Teknik analisis data menggunakan statistika deskriptif. Skor kategori kemampuan mahasiswa dalam pengembangan instrumen tes menurut (Izzati & Febrian, 2021) dapat dilihat pada Tabel 1 berikut:

**Tabel 1.** Skor Kategori Kemampuan Mahasiswa dalam Pengembangan Instrumen Tes

| <i>Skor (x)</i>      | <i>Kategori</i> |
|----------------------|-----------------|
| $81 \leq x \leq 100$ | Sangat Baik     |
| $71 \leq x < 81$     | Baik            |
| $61 \leq x < 71$     | Cukup Baik      |
| $51 \leq x < 61$     | Kurang          |
| $x < 51$             | Sangat Kurang   |

## HASIL DAN PEMBAHASAN

### Hasil

Pengumpulan data dilakukan pada saat pertemuan ke-15. Selanjutnya, peneliti melakukan analisis data dengan menggunakan statistik deskriptif. Hasil analisis data penelitian dapat dilihat pada Tabel 2 berikut:

**Tabel 2.** Skor Kemampuan Mahasiswa dalam Pengembangan Instrumen Tes

| No. | Nama Mahasiswa              | Pendapat Ahli Skala |   | Konstruk Skala |   | Dimensi Skala |   | Indikator Skala |   | Kisi-Kisi Instrumen Skala |   | Menulis Butir Soal Skala |   | Rubrik PG Skala |   | Rubrik Essay Skala |   | Validasi Konten Skala |   | Skor | Skor (0-100) | Kategori    |
|-----|-----------------------------|---------------------|---|----------------|---|---------------|---|-----------------|---|---------------------------|---|--------------------------|---|-----------------|---|--------------------|---|-----------------------|---|------|--------------|-------------|
|     |                             | 1                   | 2 | 1              | 2 | 1             | 2 | 1               | 2 | 1                         | 2 | 1                        | 2 | 1               | 2 | 1                  | 2 | 1                     | 2 |      |              |             |
| 1.  | Muhammad Salem Saifulloh    | 1                   | - | 1              | - | 1             | - | 1               | - | -                         | 2 | -                        | 2 | -               | - | -                  | 2 | -                     | 2 | 12   | 75           | Baik        |
| 2.  | Apipudin                    | 1                   |   |                | 2 |               | 2 | 1               |   |                           | 2 | 1                        |   |                 | 2 |                    |   | 1                     |   | 12   | 75           | Baik        |
| 3.  | Anne Seftiany               |                     | 2 | 1              |   | 1             |   | 1               |   |                           | 2 | 1                        |   |                 | 2 |                    |   |                       | 2 | 12   | 75           | Baik        |
| 4.  | Qori Mawari                 |                     | 2 | 1              |   | 1             |   |                 | 2 |                           | 2 | 1                        |   |                 |   |                    | 2 |                       | 2 | 13   | 81           | Sangat Baik |
| 5.  | Adelia Nur Hapsari          |                     | 2 |                | 2 |               | 2 |                 | 2 |                           | 2 |                          | 2 |                 |   |                    | 2 | 1                     |   | 15   | 94           | Sangat Baik |
| 6.  | Tita Hanafiah               |                     | 2 | 1              |   |               | 2 |                 | 2 |                           | 2 |                          | 2 |                 |   | 1                  |   |                       | 2 | 14   | 88           | Sangat Baik |
| 7.  | Ahmad Khoirudin             |                     | 2 |                | 2 |               | 2 |                 | 2 |                           | 2 | 1                        |   |                 | 2 |                    |   |                       | 2 | 14   | 88           | Sangat Baik |
| 8.  | Dwi Apriani                 |                     | 2 | 1              |   | 1             |   |                 | 2 |                           | 2 |                          | 2 |                 | 2 |                    |   |                       | 2 | 14   | 88           | Sangat Baik |
| 9.  | Rohimah                     |                     | 2 |                | 2 |               | 2 |                 | 2 |                           | 2 |                          | 2 |                 | 2 |                    |   |                       | 2 | 16   | 100          | Sangat Baik |
| 10. | Imas Masyitoh               | 1                   |   | 1              |   | 1             |   | 1               |   |                           | 2 | 1                        |   |                 | 2 |                    |   |                       | 2 | 11   | 70           | Cukup Baik  |
| 11. | Mujahiddin Al Faruqul Adzim |                     |   | 2              |   | 2             | 1 |                 | 2 |                           | 2 |                          | 2 |                 |   |                    | 2 |                       | 2 | 15   | 94           | Sangat Baik |
| 12. | Astra Priya Leksana         |                     |   | 2              |   | 2             |   | 2               |   |                           | 2 |                          | 2 |                 |   | 1                  |   |                       | 2 | 15   | 94           | Sangat Baik |
| 13. | Firma Mustika               | 1                   |   |                | 2 |               | 2 |                 | 2 |                           | 2 | 1                        |   |                 | 2 |                    |   |                       | 2 | 14   | 88           | Sangat Baik |
| 14. | Ida Oktawati                |                     | 2 | 1              |   |               | 2 |                 | 2 |                           | 2 | 1                        |   |                 | 2 |                    |   |                       | 2 | 14   | 88           | Sangat Baik |
| 15. | Nida Nabilah Rahmah         |                     | 2 | 1              |   | 1             |   |                 | 2 |                           | 2 | 1                        |   |                 | 2 |                    |   |                       | 2 | 13   | 81           | Sangat Baik |
| 16. | Shiva Adelia Putri          |                     | 2 | 1              |   | 1             |   |                 | 2 |                           | 2 | 1                        |   |                 |   |                    | 2 |                       | 2 | 13   | 81           | Sangat Baik |
| 17. | Arina Sabilulhaq            |                     | 2 |                | 2 |               | 2 |                 | 2 |                           | 2 |                          | 2 |                 | 2 |                    |   |                       | 2 | 16   | 100          | Sangat Baik |
| 18. | Rif'atul Maula              | 1                   |   | 1              |   | 1             |   |                 | 2 |                           | 2 |                          | 2 |                 |   |                    | 2 | 0                     |   | 11   | 70           | Cukup Baik  |
| 19. | Erna Widi Hastuti           |                     | 2 | 1              |   | 1             |   |                 | 2 |                           | 2 | 1                        |   |                 |   |                    | 2 |                       | 2 | 13   | 81           | Sangat Baik |
| 20. | Marwan                      |                     | 2 | 1              |   | 1             |   |                 | 2 |                           | 2 | 1                        |   |                 |   |                    | 2 |                       | 2 | 13   | 81           | Sangat Baik |

| No. | Nama Mahasiswa             | Pendapat Ahli Skala |   | Konstruk Skala |   | Dimensi Skala |   | Indikator Skala |   | Kisi-Kisi Instrumen Skala |   | Menulis Butir Soal Skala |   | Rubrik PG Skala |   | Rubrik Essay Skala |   | Validasi Konten Skala |   | Skor | Skor (0-100) | Kategori    |
|-----|----------------------------|---------------------|---|----------------|---|---------------|---|-----------------|---|---------------------------|---|--------------------------|---|-----------------|---|--------------------|---|-----------------------|---|------|--------------|-------------|
|     |                            | 1                   | 2 | 1              | 2 | 1             | 2 | 1               | 2 | 1                         | 2 | 1                        | 2 | 1               | 2 | 1                  | 2 | 1                     | 2 |      |              |             |
| 21. | Herman Sugianto            | 1                   |   |                | 2 |               | 2 |                 | 2 |                           | 2 | 1                        |   |                 | 2 |                    |   | 1                     |   | 13   | 81           | Sangat Baik |
| 22. | Nilam S. Cahya             |                     | 2 |                | 2 |               | 2 |                 | 2 |                           | 2 | 1                        |   |                 |   | 2                  |   | 2                     |   | 15   | 94           | Sangat Baik |
| 23. | Ria Apri Kandari           |                     | 2 |                | 2 |               | 2 |                 | 2 |                           | 2 |                          | 2 |                 |   | 2                  |   | 2                     |   | 16   | 100          | Sangat Baik |
| 24. | Santi Nugraha              |                     | 2 |                | 2 |               | 2 |                 | 2 |                           | 2 |                          | 2 |                 |   | 2                  |   | 2                     |   | 16   | 100          | Sangat Baik |
| 25. | Devi Meita Susanti         |                     | 2 | 1              |   | 1             |   | 2               |   | 2                         | 1 |                          |   | 2               |   |                    |   | 2                     |   | 13   | 81           | Sangat Baik |
| 26. | Inntura Sidhi Purnama Dewi |                     | 2 | 1              |   | 1             |   | 2               |   | 2                         |   | 2                        |   | 2               |   |                    |   | 2                     |   | 14   | 88           | Sangat Baik |
| 27. | Eva Dewi Rahayu            |                     | 2 | 1              |   | 1             |   | 2               |   | 2                         | 1 |                          |   |                 |   | 2                  |   | 2                     |   | 13   | 81           | Sangat Baik |
| 28. | Ulfa Mariana               |                     | 2 | 1              |   | 1             |   | 2               |   | 2                         | 1 |                          |   |                 | 1 |                    |   | 2                     |   | 12   | 75           | Baik        |
| 29. | Taufik Hidayatullah Rerata | 1                   |   | 1              |   | 1             |   | 2               |   | 2                         |   | 2                        |   | 2               |   |                    |   | 2                     |   | 13   | 81           | Sangat Baik |
|     |                            |                     |   |                |   |               |   |                 |   |                           |   |                          |   |                 |   |                    |   |                       |   |      | 85,28        |             |

Berdasarkan Tabel 2 tersebut, dapat diinterpretasikan sebagai berikut:

- 1) Pada aspek mencari teori relevan (pendapat ahli), terdapat 7 mahasiswa atau 24% mahasiswa yang belum memahami dan 22 mahasiswa atau 76% yang sudah memahami.
- 2) Pada aspek mengonstruksi variabel, terdapat 17 mahasiswa atau 57% yang belum memahami dan 12 mahasiswa atau 43% yang sudah memahami.
- 3) Pada aspek menulis dimensi variabel, terdapat 16 mahasiswa atau 55% yang belum memahami dan 13 mahasiswa atau 45% yang sudah memahami.
- 4) Pada aspek menulis indikator variabel, terdapat 4 mahasiswa atau 18% yang belum memahami dan 25 mahasiswa atau 82% yang sudah memahami.
- 5) Pada aspek menulis kisi-kisi instrumen, terdapat 29 mahasiswa atau 100% yang sudah memahami.
- 6) Analisis kemampuan mahasiswa dalam menulis butir soal berbasis HOTS sebagai berikut:
  - a. Apipudin belum memahami cara butir C6.
  - b. Anne Seftiany belum memahami cara menulis butir C6.
  - c. Qori Mawari belum memahami cara menulis butir C4 pada butir 6 dan butir C5 pada butir 8 dan 9.
  - d. Ahmad Khoirudin belum memahami cara menulis butir C6.
  - e. Imas Masyitoh belum memahami cara menulis butir C6.
  - f. Firma Mustika belum memahami cara menulis butir C6.
  - g. Nida Nabilah Rahmah belum memahami cara menulis butir C6.
  - h. Shiva Adelia belum memahami cara menulis butir C6.
  - i. Erna Widi Hastuti belum memahami cara menulis butir C4.
  - j. Marwan belum memahami cara menulis butir C4 dan butir C6.
  - k. Herman Sugianto belum memahami cara menulis butir C6.
  - l. Devi Meita Susanti belum memahami cara menulis butir C4 dan butir C6.
  - m. Ulfa Mariana belum memahami cara menulis butir C6.
- 7) Pada aspek menulis rubrik pilihan ganda, terdapat 29 mahasiswa atau 100% yang sudah memahami. Pada rubrik esai, terdapat 3 mahasiswa atau 10% yang belum memahami, dan 26 mahasiswa atau 90% yang sudah memahami.
- 8) Pada aspek validasi konten, terdapat 1 mahasiswa atau 3% yang tidak memahami, 3 mahasiswa atau 10% yang belum memahami, dan 25 mahasiswa atau 87% yang sudah memahami.
- 9) Rerata kemampuan mahasiswa dalam pengembangan instrumen tes sebesar 85,28 dengan kategori sangat baik.
- 10) Kategori kemampuan mahasiswa dapat dilihat pada Tabel 3 berikut:

**Tabel 3.** Kategori Kemampuan Mahasiswa dalam Pengembangan Instrumen Tes

| Kategori    | Jumlah Mahasiswa | Persentase |
|-------------|------------------|------------|
| Sangat Baik | 24               | 83         |
| Baik        | 4                | 14         |
| Cukup Baik  | 1                | 3          |

Hasil penelitian tersebut menunjukkan kemampuan mahasiswa dalam pengembangan instrumen tes berdasarkan Taksonomi Bloom revisi berada pada kategori sangat baik dengan rerata skor sebesar 85,28. Mahasiswa telah memahami penulisan kisi-kisi, dimensi, penulisan indikator, penulisan butir soal dan penulisan rubrik penskoran secara sistematis. Temuan ini menunjukkan bahwa pembelajaran pada mata kuliah tes dan pengukuran dapat meningkatkan kompetensi mahasiswa dalam pengembangan instrumen tes yang sesuai dengan tujuan pembelajaran.

Penelitian ini juga menemukan bahwa sebanyak 13 mahasiswa atau 45% belum memahami penulisan butir soal berbasis HOTS pada level kognitif C4 (menganalisis), C5 (mengevaluasi), dan C6 (menciptakan). Kesulitan utama terlihat pada penulisan stimulus dan pengembangan pertanyaan yang menuntut kemampuan HOTS. Oleh karena itu, mahasiswa memerlukan pelatihan lebih intensif agar dapat mengembangkan penulisan butir soal berbasis HOTS.

### **Pembahasan**

Terdapat 13 mahasiswa atau 45% yang belum memahami penulisan butir soal berbasis HOTS, yaitu C4 (menganalisis), C5 (mengevaluasi), dan C6 (menciptakan). Temuan ini sejalan dengan (Stringer et al., 2021) bahwa pemahaman mahasiswa terkait pengembangan instrumen dengan Taksonomi Bloom berbasis HOTS belum optimal. Temuan ini berbeda dengan hasil penelitian (Izzati & Febrian, 2021) yang menyatakan bahwa bahwa skor rerata kemampuan 28 mahasiswa calon guru prodi Pendidikan Matematika Universitas Maritim Raja Ali Haji dalam pengembangan instrumen tes berbasis HOTS sebesar 76,2 dan (Maxnun et al., 2024; Saraswati et al., 2021) yang menyatakan bahwa pengembangan 10 butir soal esai berbasis HOTS valid dengan skor reliabilitas butir sebesar 0,863 dengan kategori tinggi (Yuliandini et al., 2019), tingkat kesukaran butir: 50% sedang, 50% sulit, dan daya pembeda dengan kategori baik. Temuan ini berbeda dengan hasil pengabdian kepada masyarakat (Adi et al., 2025; Riyadi et al., 2023) yang menyatakan bahwa pada pengembangan instrumen tes berbasis HOTS terjadi peningkatan dari 47% menjadi 66% dengan kategori baik sekali. Temuan ini sejalan dengan (Mawaddah & Fauziah, 2024; Putri et al., 2024) yang menyatakan bahwa implementasi di lapangan masih menghadapi beberapa kendala, yaitu keterbatasan pemahaman guru terhadap konsep HOTS, rendahnya partisipasi aktif siswa dalam proses berpikir kritis, dan kurangnya fasilitas dan sistem evaluasi yang mendukung pembelajaran berbasis analitis dan kreatif. Temuan ini sejalan dengan (Adhitya, 2024; Rahman Rahim & Sukri Syamsuri, 2021) yang menyatakan bahwa penulisan butir soal guru mata pelajaran Bahasa Indonesia berada pada level kognitif LOTS: mengingat, memahami, dan menerapkan. Temuan ini sejalan dengan (Indriani et al., 2023) bahwa guru kesulitan dalam pengembangan instrumen dengan Taksonomi Bloom revisi yang meliputi: 1) pemahaman guru dalam menulis kata kerja operasional, 2) menulis buku ajar yang sesuai dengan Taksonomi Bloom revisi, 3) menulis kisi-kisi instrumen, dan 4) validasi instrumen dengan teman sejawat. Temuan ini juga sejalan dengan (Bekti Mulatsih, 2021) bahwa guru dalam pengembangan instrumen dengan Taksonomi Bloom revisi harus memahami secara mendalam kata kerja operasional yang digunakan dalam menulis butir soal.

Berdasarkan hasil analisis data, temuan penelitian menunjukkan bahwa kemampuan mahasiswa pascasarjana dalam pengembangan instrumen tes berdasarkan Taksonomi Bloom revisi sebanyak 24 mahasiswa atau 84% dengan kategori sangat baik dan sebanyak 13 mahasiswa belum memahami penulisan butir soal berbasis HOTS. Kendala utama terlihat pada penulisan stimulus dan pengembangan pertanyaan yang menuntut kemampuan HOTS. Mahasiswa

terkendala terkait penggunaan kata kerja operasional. Kata kerja operasional dalam butir soal HOTS harus dilihat dari aktivitas kognitif yang benar-benar diperlukan peserta didik untuk merespons butir soal, bukan dari kata kerja operasional yang digunakan dalam menulis butir soal tersebut. Temuan ini sesuai dengan (Miladanta et al., 2024).

## KESIMPULAN

Tujuan dari penelitian ini adalah menganalisis kemampuan mahasiswa dalam pengembangan instrumen tes berdasarkan Taksonomi Bloom revisi. Berdasarkan hasil penelitian, dapat disimpulkan: 1) kemampuan mahasiswa dalam pengembangan instrumen tes dengan Taksonomi Bloom revisi pada kategori sangat baik sebesar 83%, 2) rerata kemampuan mahasiswa dalam pengembangan instrumen tes sebesar 85,28 dengan kategori sangat baik, dan 3) terdapat 13 mahasiswa atau 45% yang belum memahami penulisan butir soal berbasis HOTS, yaitu C4 (menganalisis), C5 (mengevaluasi), dan C6 (menciptakan). Temuan ini menunjukkan bahwa kemampuan mahasiswa dalam menulis butir soal belum optimal, terutama dalam penulisan butir soal berbasis HOTS. Keterbatasan penelitian ini adalah terkait waktu, yaitu kurangnya komunikasi antara mahasiswa dan dosen, terutama dalam penulisan butir soal berbasis HOTS dan kurangnya implementasi mahasiswa dalam pembelajaran berbasis HOTS di kelas. Kurangnya komunikasi antara mahasiswa dan dosen terjadi karena pembelajaran 80% dilakukan secara online. Kurangnya implementasi pembelajaran berbasis HOTS di kelas merupakan indikator bahwa guru belum mengembangkan instrumen tes berbasis HOTS (Wahyudin et al., 2026). Solusi yang dapat dilakukan bagi mahasiswa atau guru dalam pengembangan butir soal berbasis HOTS adalah membiasakan menulis butir soal berbasis (HOTS) dalam kegiatan ulangan harian, ujian tengah semester, dan ujian akhir semester, sehingga pengalaman dan kemampuan guru dalam menulis butir soal dapat meningkat. Hal ini sesuai dengan (Hamdi et al., 2022).

## DAFTAR PUSTAKA

- Adhitya, Y. (2024). Analysis of Prospective Teacher Students' Errors in Making Math Test Instruments Based on Bloom's Taxonomy Revised. *Eduscape : Journal of Education Insight*, 2(1), 26–36. <https://doi.org/10.61978/eduscape.v2i1>
- Adi, K. R., Nurul Ratnawati, Kurniawan, B., Nareswari, D. A., & Alfania, S. (2025). Peningkatan Kompetensi Guru dalam Menyusun Instrumen Evaluasi Berbasis Literasi Numerasi melalui Elaborasi Taksonomi Bloom, Marzano, dan Solo di SMP Negeri 22 Malang. *Jurnal Masyarakat Madani Indonesia*, 4(4), 960–969. <https://doi.org/10.59025/zwvbm543>
- Afifah Aulia Zayrin, Hayatun Nupus, Khalista Khansa Maizia, Siska Marsela, Rully Hidayatullah, & Harmonedi, H. (2025). Analisis Instrumen Penelitian Pendidikan (Uji Validitas Dan Reliabilitas Instrumen Penelitian). *Jurnal QOSIM Jurnal Pendidikan Sosial & Humaniora*, 3(2), 780–789. <https://doi.org/10.61104/jq.v3i2.1070>
- Agusdianita, N. (2023). Model Pembelajaran PJB L Untuk Meningkatkan Hasil Belajar Mahasiswa pada Perkuliahan Pengembangan Pembelajaran Tematik. *Seminar Nasional Inovasi Pendidikan*, 6(3), 160–166.
- Akinboboye, J. T., Musa, &, & Ayanwale, A. (2021). Bloom Taxonomy Usage And Psychometric Analysis Of Classroom Teacher-Made Test. *African Multidisciplinary Journal Of Development (AMJD)*, 10(1), 2021.
- Anggermawan, Y. P., Sukmo Wardhono, W., & Suharsono, A. (2017). Pengembangan E-Modul Mata Pelajaran Informatika Materi Sistem Komputer dan Komponen Penyusunnya dengan Model Pengembangan BORG & GALL (STUDI KASUS: SMA). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(1), 2548–2964.
- Ayu, G., Juniari, O., Gusti, I., & Oka Juniari, A. (2021). Upaya Meningkatkan Semangat Belajar Siswa Melalui Media Pembelajaran Multimedia Interaktif pada Pelajaran IPA Kelas V Sekolah Dasar. *Jurnal Edutech Undiksha*, 8(1), 140–148.

- Bekti Mulatsih. (2021). Implementation of Revised Bloom Taxonomy in Developing Chemistry Questions in the Domain of Knowledge. *Ideguru: Jurnal Karya Ilmiah Guru*, 6(1), 1–10. <https://doi.org/10.51169/ideguru.v6i1.158>
- Billa, S., Purnamasari, R., & Suchyadi, Y. (2024). PENGEMBANGAN INSTRUMEN TES PILIHAN GANDA BERBASIS HOTS MENGGUNAKAN APLIKASI QUIZZZ DAN QR-CODE PADA PEMBELAJARAN MATEMATIKA MATERI PECAHAN. *Didaktik: Jurnal Ilmiah PGSD FKIP Universitas Mandiri*, 10(4), 335–343.
- Chasanah, N., Widodo, W., & Suprpto, N. (2022). Pengembangan Instrumen Asesmen Literasi Sains Untuk Mendeskripsikan Profil Peserta Didik. *PENDIPA Journal of Science Education*, 6(2), 474–483. <https://doi.org/10.33369/pendipa.6.2.474-483>
- Choir, M. N. A. C. (2025, December 24). Rerata Nilai TKA Jeblok: Alarm Bahaya Pendidikan. 1–2.
- Desiriah, E., Setyarsih, W., Fisika, J., Matematika, F., Ilmu, D., & Alam, P. (2021). Tinjauan Literatur Pengembangan Instrumen Penilaian Kemampuan Berpikir Tingkat Tinggi (HOTS) FISIKA di SMA. 7(1).
- Eliza, R., Utami, N. P., & Warahma, S. (2022). Higher Order Thinking Skills of Mathematics Education Students at State Islamic Higher Education in Central Sumatera. *Edumatika: Jurnal Riset Pendidikan Matematika*, 5(1). <https://doi.org/10.32939/ejrpm.v5i1.1273>
- Fadzillah Karira, N., Sunarti, T., Niswati, M., & Setyasih, W. (2023). Validitas Instrumen Tes Berbasis Literasi Sains untuk Mengukur Keterampilan Berpikir Kritis Siswa SMA pada Materi Energi Terbarukan. *Inovasi Pendidikan Fisika*, 12(2), 118–125.
- Fitrianty, F., Yunita, A., & Juwita, R. (2022). Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Matematika di SMP Negeri 12 Padang. *Lattice Journal: Journal of Mathematics Education and Applied*, 2(1), 91. <https://doi.org/10.30983/lattice.v2i1.5337>
- Hadiyanti, H., & Susongko, P. (2024). Pengembangan Instrumen Higher Order Thinking Skill Mata Pelajaran Matematika Dengan Rasch Model. *Journal of Education Research*, 5(1), 399–407. <https://doi.org/https://doi.org/10.37985/jer.v5i1.765>
- Hamdi, S., Kartowagiran, B., Istiyono, E., & Rahman, H. N. (2022). The ability of Junior High School mathematics teachers to write Higher Order Thinking Skills (HOTS) questions. *Jurnal Riset Pendidikan Matematika*, 9(2), 115–123. <https://doi.org/10.21831/jrpm.v9i2.55191>
- Handoko, Rohman, F., Isnir Rahman, S., Rahmawati, D., & Ayu Sucitra, D. (2025). Pengembangan Instrumen Penilaian Proses Menggunakan Model ADDIE. *Arus Jurnal Pendidikan (AJP)*, 5(3), 287–293. <https://doi.org/https://doi.org/10.57250/ajup.v5i3.1719>
- Hanik, A., Ashari, & Ngazizah, N. (2021). Pengembangan Instrumen Tes Berbasis Higher Order Thinking Skills (Hots) Kelas V Madrasah Ibtidaiyah. *Jurnal Pendidikan Dasar.*, 2(10), 74–84.
- Hulwani, A. Z., & Yuhana, Y. (2024). Pengembangan Instrumen Tes Matematika Berbasis Higher Order Thinking Skills (HOTS) Untuk Siswa SMP. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 5(2), 1074–1081. <https://doi.org/10.46306/lb.v5i2>
- Imania, K. A. N., & Bariah, S. K. (2019). Rancangan Pengembangan Instrumen Penilaian Pembelajaran Berbasis Daring. *Jurnal PETIK*, 5(1), 31–47. <https://doi.org/https://doi.org/10.31980/petik.v5i1.1534>
- Indriani, N., Wahyuni, N., Putri Maharani, A., Afidatul Afi, D., Faradisa Bahri, F., & Nur Rahma, S. (2023). Analisis Kesulitan Guru Dalam Mengembangkan Instrumen Asesmen Autentik Higher Order Thinking Skill di MI Sidoarjo. *Afeksi: Jurnal Penelitian Dan Evaluasi Pendidikan*, 4(6), 596–607.
- Izzati, N., & Febrian, F. (2021). Kemampuan Mahasiswa Calon Guru Dalam Mengembangkan Instrumen Tes Hasil Belajar Kategori Higher Order Thinking Skill. *Jurnal Gantang*, 6(1), 75–82. <https://doi.org/10.31629/jg.v6i1.2672>
- Kristiana Wabang, L., Marwina Almet, S., & Panjaitan, R. (2026). Efektivitas Penggunaan Rubrik Analitik dalam Meningkatkan Objektivitas Penilaian Tes Uraian. *Educational Journal*, 1(4). <https://doi.org/10.63822/v0n7km25>
- Kurniawati, A. (2020). Pengembangan Instrumen Penilaian Tes Berbasis Mobile Online Pada Prodi Pendidikan Matematika. *Jurnal Pendidikan Matematika*, 4(1), 46–63.

- Lara Juni Ulpa, M. Naswir, & Fuldariatman. (2025). Pengembangan Instrumen Tes untuk Mengukur HOTS Siswa pada Materi Asam Basa. *Jurnal Pengabdian Masyarakat Dan Riset Pendidikan*, 3(4), 3369–3379. <https://doi.org/10.31004/jerkin.v3i4.1018>
- Litna, K. O., Mertasari, N. M. S., & Sudirtha, G. (2021). Pengembangan Instrumen Tes Higher Order Thinking Skills (Hots) Matematika SMA Kelas X. *Jurnal Penelitian Dan Evaluasi Pendidikan Indonesia*, 11(1).
- Magdalena, I., Hifziyah, M., Aeni, V. N., & Rahayu, R. P. (2020). Pengembangan Instrumen Tes Siswa Tingkat Sekolah Dasar Kabupaten Tangerang. *Jurnal Pendidikan Dan Ilmu Sosial*, 2(2), 227–237.
- Mawaddah, R., & Fauziah, N. (2024). Analysis of Higher Order Thinking Skills (HOTS) of Class XI MIPA Students in the Coordination System Material. *Perspektif Pendidikan dan Keguruan*, 15(1), 54–66. [https://doi.org/10.25299/perspektif.2024.vol15\(1\).16462](https://doi.org/10.25299/perspektif.2024.vol15(1).16462)
- Maxnun, L., Kristiani, K., & Sulistyaningrum, C. D. (2024). Development of HOTS-based cognitive assessment instruments: ADDIE model. *Journal of Education and Learning*, 18(2), 489–498. <https://doi.org/10.11591/edulearn.v18i2.21079>
- Miladanta, A. N., Nuryantini, A. Y., Farida, I., & Cahyanto, T. (2024). Pengembangan Instrumen Higher Order Thinking Skills (HOTS) Materi Alat Optik. 4(2), 148–158. <https://doi.org/https://doi.org/10.23971/jpsp.v4i2.7828>
- Nadzir, H. N. (2023). Pengembangan E-Modul Menggunakan Model Hannafin and Peck pada Mata Pelajaran Seni Budaya. *Jurnal Media Dan Teknologi Pendidikan*, 3(1), 47–55. <https://doi.org/10.23887/jmt.v3i1.58570>
- Nasrum, A. (2020). Pengembangan Instrumen Evaluasi Pemahaman Konsep Kalkulus Berbasis Komputer. *HISTOGRAM: Jurnal Pendidikan Matematika*, 4(1), 78–91. <https://doi.org/10.31100/histogram.v4i1.540>
- Ningrum, A. S., Muslim, S., & Siregar, E. (2022). Multimedia Simulation Model on Basic Electrical and Electronics Subjects for Vocational Secondary School. *Journal of Education Research and Evaluation*, 6(1), 72–79. <https://doi.org/10.23887/jere.v6i1.38783>
- Nugrahnastiti, N., & Kamaludin, A. (2024). Development of a Higher-Order Thinking Skills (HOTS) Test Instrument on Electrochemical Material for High School Students. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran*, 10(2), 451. <https://doi.org/10.33394/jk.v10i2.11370>
- Nurmaliati, Ilwandri, & Apra Santosa, T. (2024). Development Of Instruments For Assessing Higher-Order Thinking Skills (HOTS) in Physics in High School (Literature Review). *Indonesia Journal of Engineering and Education Technology (IJEET)*, 2, 6–17.
- Nurmaya, E., Rusilowati, A., & Astuti, B. (2022). Development of Cognitive Ability Test Instrument Based on Revision Bloom Taxonomy on Dynamic Electricity Materials For Students of Senior High School. *Phys. Comm*, 6(2), 50–60.
- OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*, PISA, OECD Publishing, Paris.
- Phafiandita, A. N., Permadani, A., Pradani, A. S., & Wahyudi, M. I. (2022). Urgensi Evaluasi Pembelajaran di Kelas. *JIRA: Jurnal Inovasi dan Riset Akademik*, 3(2), 111–121. <https://doi.org/10.47387/jira.v3i2.262>
- Prima Kurniawati, R., & Rahmantika Hadi, F. (2021). Pelatihan Pengembangan Instrumen Evaluasi Berbasis HOTS untuk Guru Sekolah Dasar. *Pengabdian Kepada Masyarakat*, 1(4), 267–276. <https://doi.org/10.25008/altifani.v1i3.182>
- Pujiastuti, R., & Kulup, L. I. (2021). Penyusunan Instrumen Penilaian Kognitif Berbasis HOTS Melalui Problem-Based Learning dan Peer Assessment. *Indonesian Language Education and Literature*, 7(1), 88. <https://doi.org/10.24235/ileal.v7i1.9058>
- Putri, R. N., Wardhani, N., & Dayusman, E. A. (2024). Analisis Pelaksanaan Pembelajaran Pendidikan Agama Islam Berbasis Higher Order Thinking Skills (HOTS) di SMA Negeri 1 Gandapura. *Journal of Contemporary Islamic Education*, 3(2), 145–160. <https://doi.org/10.47766/pase.v3i2.5212>

- Putu Agus Adi Saputra, I., Nyoman Jampel, I., & Wayan Suwatra, I. I. (2021). Pengembangan Instrumen Penilaian Kompetensi Pengetahuan IPA Siswa Sekolah Dasar. *Journal for Lesson and Learning Studies*, 4(1), 13–19.
- Rahman Rahim, A., & Sukri Syamsuri, A. (2021). Analisis Instrumen Tes Hasil Belajar Berbasis Higher Order Thinking Skill (HOTS). *JRIP: Jurnal Riset dan Inovasi Pembelajaran*, 1(3), 88–101.
- Rahmawati, N. D., Komarudin, K., & Suherman, S. (2022). Pengembangan Instrumen Penilaian Matematika Berbasis Hots Pada Calon Guru Sekolah Dasar. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 11(2), 860–871. <https://doi.org/10.24127/ajpm.v11i2.4893>
- Retnawati, H. (2017). *Validitas Reliabilitas dan Karakteristik Butir* (2nd ed.). Parama Publishing.
- Rivaldi, H. R., Wahyudi, A., & Sari, R. A. N. I. (2025). Analisis Kualitas Butir Soal Asesmen Sumatif Geografi pada Materi Peta, Penginderaan Jauh, dan Sistem Informasi Geografis. 4(1), 102–113. <https://doi.org/https://doi.org/10.58706/jipp.v4n1.p102-113>.
- Riyadi, Mahir I., & Margono, G. (2023). Peningkatan Kemampuan Guru Dalam Mengembangkan Instrumen Penilaian Berbasis Higher Order Thinking Skills. *Prosiding Seminar Nasional Pengabdian Kepada Masyarakat*, 2023, 2023.
- Rukmi Octaviana, D., Sutomo, M., & Mashudi. (2022). Model Pembelajaran Dick and Carey Serta Implementasinya Dalam Pembelajaran PAI. *Jurnal Tawadhu*, 6(2), 14–126. <https://doi.org/https://doi.org/10.23887/jpepi.v11i1.278>
- Rumtini, Kasimin, & Ari Setiawan. (2022). Pengembangan Instrumen Penilaian Afektif Kemampuan Bernalar Kritis Peserta Didik Sekolah Menengah Atas. *JUPEIS: Jurnal Pendidikan dan Ilmu Sosial*, 1(2), 138–147.
- Safitri, F., Rusdiana, D., Samsudin, A., & Widiyatmoko, A. (2025). Development of Understanding Test Instruments for Grade 7 Junior High School Students on Temperature, Heat, and Expansion Topics. *Jurnal Penelitian Pendidikan IPA*, 11(1), 395–404. <https://doi.org/10.29303/jppipa.v11i1.9271>
- Saraswati, S., Rodliyah, I., & Rahmawati, N. D. (2021). Analisis Instrumen Penilaian Berbasis Higher Order Thinking Skills pada Mata Kuliah Matematika Lanjut. *Inovasi Matematika (Inomatika)*, 3(2), 138–151. <https://doi.org/10.35438/inomatika>
- Siagian, Q. A., & Sagala, P. N. (2021). Development of Test Instruments to Measure Higher-Order Thinking Skills (HOTS) Mathematics of Students in MTs 2 State Of Medan City. *Mathline : Jurnal Matematika Dan Pendidikan Matematika*, 6(2), 154–174. <https://doi.org/10.31943/mathline.v6i2.222>
- Stringer, J. K., Santen, S. A., Lee, E., Rawls, M., Bailey, J., Richards, A., Perera, R. A., & Biskobing, D. (2021). Examining Bloom's Taxonomy in Multiple Choice Questions: Students' Approach to Questions. *Medical Science Educator*, 31(4), 1311–1317. <https://doi.org/10.1007/s40670-021-01305-y>
- Sukamto, A., Bahrani, B., & Nurvayanti, N. (2026). Tes Kemampuan Akademik (TKA) sebagai Inovasi Evaluasi Pendidikan: Sebuah Tinjauan Naratif. *Journal of Instructional and Development Researches*, 6(1), 54–69. <https://doi.org/10.53621/jider.v6i1.695>
- Wahyudin, Ruslan, & Nasaruddin. (2026). Pengaruh Pembelajaran Berbasis Higher Order Thinking Skills (HOTS) terhadap Kemampuan Berpikir Kritis di SMP Negeri 4 Sanggar. *TAJDID: Jurnal Pemikiran Keislaman dan Kemanusiaan*, 10(1), 816–826. <https://doi.org/10.52266/tadjid.v10i1.5862>
- Yuliandini, N., Hamdu, G., & Respati, R. (2019). Pengembangan Soal Tes Berbasis Higher Order Thinking Skill (Hots) Taksonomi Bloom Revisi di Sekolah Dasar. *All Rights Reserved*, 6(1), 37–46.