



ABDI CENDEKIA:

JURNAL PENGABDIAN MASYARAKAT

Volume 5 Nomor 2 Tahun 2026 Halaman 1428-1436

<https://zia-research.com/index.php/abdicendekia/index>



Two-Tier Multiple Choice Sebagai Inovasi Instrumen TKA Pada Materi Bilangan

Fadiyah Syafigah¹, Arfin Efriani², Retni Paradesa³

^{1,2,3} Universitas Islam Negeri Raden Fatah Palembang, Indonesia

Email: fadiisyaf2311@gmail.com¹; arvinefriani_uin@radenfatah.ac.id^{2*};

retniparadesa_uin@radenfatah.ac.id³

Abstrak

Penelitian ini dirancang untuk menghasilkan instrumen evaluasi berbasis kemampuan berpikir tingkat tinggi (HOTS) dalam format soal *Two-Tier Multiple Choice* pada materi bilangan untuk peserta didik SMP kelas VIII, serta untuk mengetahui tingkat kevalidan instrumen yang dikembangkan. Latar belakang penelitian ini adalah rendahnya capaian nilai TKA matematika di tingkat nasional maupun regional, yang mengindikasikan masih banyak siswa yang belum mampu memahami konsep secara mendalam maupun menyelesaikan soal bertipe HOTS. Penelitian dilaksanakan di SMP Negeri 1 Pemulutan Barat dengan subjek 10 siswa. Pendekatan yang digunakan adalah *Research and Development* (R&D) dengan model ADDIE yang melewati lima tahapan: *Analysis, Design, Development, Implementation*, dan *Evaluation*. Pengumpulan data dilakukan melalui lembar validasi ahli, hasil pengerjaan soal oleh siswa, serta dokumentasi. Data dianalisis secara kualitatif berdasarkan masukan validator dan secara kuantitatif melalui analisis *deskriptif persentase*. Pengujian validitas butir soal menggunakan *korelasi Pearson product-moment* pada *SPSS Statistics 16*. Instrumen mencapai persentase validasi 82% (sangat valid). Hasil uji validitas menunjukkan bahwa seluruh butir soal dinyatakan valid dan uji reliabilitas memperoleh nilai 0,986 yang menunjukkan instrumen memiliki reliabilitas sangat tinggi. Dengan demikian, instrumen layak digunakan untuk mengukur kemampuan berpikir kritis siswa pada materi bilangan.

Kata kunci: ADDIE, Two-Tier Multiple Choice, HOTS, Tes Kemampuan Akademik, Materi Bilangan, Pengembangan Instrumen.

Two-Tier Multiple Choice As A TKA Instrument Innovation On Numbers

Abstract

This study was designed to produce a Higher Order Thinking Skills (HOTS)-based evaluation instrument in the format of Two-Tier Multiple Choice questions on number material for eighth-grade junior high school students, as well as to determine the validity level of the developed instrument. The background of this research is the low achievement of national and regional

mathematics TKA scores, indicating that many students have not been able to understand concepts deeply or solve HOTS-type problems. The study was conducted at SMP Negeri 1 Pemulutan Barat with 10 students as subjects. The approach used was Research and Development (R&D) with the ADDIE model, which consists of five stages: Analysis, Design, Development, Implementation, and Evaluation. Data collection was carried out through expert validation sheets, student answer sheets, and documentation. Data were analyzed qualitatively based on validator feedback and quantitatively through descriptive percentage analysis. Item validity was tested using Pearson product-moment correlation in SPSS Statistics 16. The instrument achieved a validation percentage of 82% (very valid). The results of the validity test showed that all items were declared valid, and the reliability test obtained a value of 0.986, indicating that the instrument has very high reliability. Therefore, the instrument is suitable for measuring students' critical thinking skills on number material.

Keywords: ADDIE, Two-Tier Multiple Choice, HOTS, Academic Ability Test, Number Material, Instrument Development.

PENDAHULUAN

Pembaruan sistem penilaian dalam dunia pendidikan Indonesia merupakan bagian dari upaya sistematis pemerintah dalam meningkatkan mutu pembelajaran secara menyeluruh. Sistem evaluasi yang baik akan memberikan gambaran mengenai penguasaan kompetensi yang dimiliki peserta didik. Salah satu evaluasi yang digunakan saat ini adalah Tes Kemampuan Akademik, yang keberadaannya diatur melalui Permendiknasmen Nomor 9 Tahun 2025. Melalui TKA, pemerintah dapat memperoleh data capaian belajar siswa di berbagai jenjang sekaligus menjadikannya dasar perumusan kebijakan peningkatan mutu pendidikan nasional (Widana, 2021). Pelaksanaan TKA secara nasional menunjukkan bahwa kemampuan akademik siswa dalam mata pelajaran matematika tergolong rendah. Data yang dirilis oleh Kementerian Pendidikan Dasar dan Menengah melalui Pusat Pendidikan tahun 2025 secara nasional dari 43.441 sekolah di seluruh Indonesia, sebanyak 44,7% siswa berkategori rendah. Temuan serupa juga ditemukan pada Provinsi Sumatra Selatan, sebanyak 1.228 sekolah, sebanyak 48,5% siswa berkategori rendah.

Fakta tersebut mengisyaratkan bahwa sebagian besar siswa masih menghadapi kendala serius dalam memahami konsep-konsep matematika, terutama dalam mengerjakan soal yang menuntut kemampuan analitis dan penalaran tingkat tinggi. Sebagai mata pelajaran yang berkontribusi besar dalam membentuk pola pikir logis, kritis, dan sistematis, matematika memegang posisi strategis dalam kurikulum pendidikan. (Widana, 2021). Salah satu materi penting dalam pembelajaran matematika di jenjang SMP adalah materi bilangan yang menjadi fondasi bagi berbagai konsep lanjutan sekaligus berhubungan erat dengan penyelesaian masalah dalam kehidupan nyata. Meskipun demikian, berbagai penelitian menunjukkan bahwa siswa masih mengalami terkendala dalam memahami konsep bilangan secara mendalam (Hidayat, 2021). sehingga diperlukan instrumen penilaian yang mampu mengidentifikasi dan memetakan pemahaman konseptual siswa secara lebih mendalam.

Pengembangan Instrumen penilaian yang baik harus mampu mengukur kemampuan siswa secara valid, reliabel, dan sesuai dengan tujuan pembelajaran (Rahman, 2022). Pengembangan soal bertipe HOTS terbukti mampu mendorong peningkatan kemampuan

berpikir kritis peserta didik secara signifikan (Rohana, 2021) Lebih lanjut, pengembangan alat asesmen harus menghasilkan instrumen yang valid untuk mengukur keterampilan tingkat tinggi siswa (Kesorn, et al., 2020). Salah satu format instrumen penilaian yang relevan *adalah two-tier multiple choice*, yakni soal berlapis dua tingkat. Pada tingkat pertama, siswa dihadapkan pada pertanyaan utama yang menguji pemahaman konsep, sementara tingkat kedua menuntut siswa untuk mengemukakan alasan yang mendasari jawaban yang telah dipilih. (Rahmawati & Widodo, 2021).

Penelitian Rahmawati dan Widodo (2021) menunjukkan bahwa instrumen *two-tier* mampu mengidentifikasi pemahaman konsep siswa secara lebih akurat dibandingkan dengan soal pilihan ganda biasa. Selain itu, penelitian (Suryani & Widodo, 2023) juga membuktikan bahwa format soal ini relevan untuk mengukur kemampuan berpikir tingkat tinggi dalam konteks pembelajaran. Temuan ini selaras dengan pengembangan soal HOTS berbasis *literasi-numerasi* yang mendukung evaluasi komprehensif pada materi dasar matematika (Rohana, 2021). Meski demikian, penelitian yang secara spesifik mengaitkan format *two-tier* dengan konteks TKA pada materi bilangan di jenjang SMP masih sangat terbatas, sehingga pengembangan instrumen ini menjadi penting dan relevan untuk dilakukan.

Berdasarkan pemaparan di atas, penelitian ini difokuskan pada pengembangan instrumen TKA dalam format *two-tier multiple choice* untuk mengukur pemahaman siswa pada materi bilangan. Instrumen ini diharapkan sekali untuk menghasilkan alat evaluasi yang valid dan mampu mengidentifikasi pemahaman konsep siswa secara lebih mendalam, sehingga mendukung peningkatan kualitas evaluasi pembelajaran matematika di tingkat SMP.

METODE

Jenis penelitian yang digunakan adalah *Research and Development* (R&D) yang bertujuan untuk menghasilkan instrumen soal *Two-Tier Multiple Choice* berbasis kemampuan berpikir tingkat tinggi pada materi bilangan siswa SMP kelas VIII serta untuk mengetahui tingkat kevalidan instrumen yang dikembangkan. Model ADDIE dipilih sebagai kerangka pengembangan karena mencakup tahapan yang mendalam mulai dari *analysis, design, development, Implementation*, hingga *evaluation*, serta dikenal adaptif terhadap berbagai kebutuhan dalam pengembangan produk pembelajaran. Penelitian ini dilaksanakan di SMP Negeri 1 Pemulutan Barat dengan subjek penelitian siswa kelas VIII. Uji coba instrumen dilakukan secara terbatas pada kelompok kecil yang terdiri dari 10 siswa, serta melibatkan dosen ahli pendidikan matematika sebagai validator. Tahap *analysis* dilakukan untuk mengkaji kebutuhan pengembangan instrumen melalui analisis kurikulum yang berlaku. Analisis kebutuhan menunjukkan bahwa siswa belum mampu secara optimal dalam menghadapi soal-soal yang menuntut kemampuan berpikir tingkat tinggi. Tahap *design* dilakukan dengan menyusun draf soal berbasis kontekstual yang dirancang untuk mengukur kemampuan analisis dan evaluasi siswa.

Tahap *development* mencakup proses validasi oleh pakar untuk menilai kevalidan instrumen, revisi soal sesuai masukan validator, serta uji coba terbatas untuk memperoleh data kualitas instrumen dari sisi validitas dan reliabilitas. Aspek yang dievaluasi pada tahap ini meliputi kesesuaian materi, konstruksi soal, kebahasaan, dan kelayakan instrumen secara keseluruhan. Tahap *Implementation* dilakukan dengan mengidentifikasi indikator kemampuan berpikir tingkat tinggi pada level kognitif C3 (mengaplikasikan) dan C4 (menganalisis), serta

menyusun kisi-kisi dan kartu soal. Instrumen yang dikembangkan berbentuk *two-tier multiple choice*, yaitu terdiri dari dua tingkat jawaban, di mana tingkat pertama berupa pilihan jawaban dan tingkat kedua berupa alasan terhadap jawaban yang dipilih. Tahap *evaluation* dilakukan untuk mengetahui kualitas instrumen yang dikembangkan, meliputi aspek kevalidan dan reliabilitas instrumen. Data penelitian dikumpulkan melalui lembar validasi ahli, hasil tes siswa, dan dokumentasi. Hasil validasi ahli digunakan untuk menentukan tingkat kevalidan instrumen berdasarkan kriteria yang disajikan pada Tabel 1.

Tabel 1. Kriteria Analisis Kevalidan

Persentase	Kategori
$81\% \leq R < 100\%$	Sangat Valid
$61\% \leq R < 80\%$	Valid
$41\% \leq R < 60\%$	Cukup Valid
$21\% \leq R < 40\%$	Kurang Valid
$\%R < 20\%$	Tidak Valid

Data hasil tes siswa selanjutnya digunakan untuk menganalisis validitas dan reliabilitas butir soal. Data penelitian diolah secara kualitatif melalui telaah masukan validator dan secara kuantitatif menggunakan teknik deskriptif persentase terhadap penilaian ahli untuk menentukan tingkat kelayakan instrumen yang dikembangkan.

HASIL DAN PEMBAHASAN

HASIL

Jenis penelitian yang digunakan adalah *Research and Development* (R&D) untuk menghasilkan instrumen soal *Two-Tier Multiple Choice* berbasis kemampuan berpikir tingkat tinggi pada materi bilangan siswa SMP kelas VIII. Model ADDIE dipilih sebagai kerangka pengembangan karena mencakup tahapan yang mendalam mulai dari *analysis, design, development, Implementation*, hingga *evaluation*, serta dikenal adaptif terhadap berbagai kebutuhan dalam pengembangan produk pembelajaran (Rustandi & Rismayanti, 2021).

1. Prosedur Pengembangan

a. Tahap Analisis (*Analysis*)

Pada tahap ini dilakukan identifikasi permasalahan pembelajaran yang terjadi di lapangan. Temuan menunjukkan bahwa kemampuan berpikir kritis siswa masih sangat rendah, sebagaimana tercermin dari capaian TKA tahun 2025 yang jauh dari memuaskan, terutama pada soal-soal yang menuntut penalaran tingkat tinggi. Selain itu, instrumen asesmen yang digunakan masih didominasi oleh soal tingkat rendah sehingga belum mampu mengukur kemampuan berpikir kritis siswa secara optimal. Menurut Andi Rustandi dan Rismayanti (2021), tahap analisis dalam model ADDIE dilakukan untuk mengidentifikasi kebutuhan, permasalahan, dan kondisi pembelajaran sebagai dasar dalam pengembangan produk yang sesuai dengan tujuan pembelajaran.

b. Tahap Desain (*Design*)

Pada tahap desain dilakukan perancangan instrumen yang akan dikembangkan. Kegiatan yang dilakukan meliputi penyusunan kisi-kisi soal berbasis HOTS, kartu soal,

penentuan indikator pembelajaran, serta perancangan bentuk soal *two-tier multiple choice* yang terdiri atas jawaban dan alasan. Desain ini disesuaikan dengan karakteristik siswa. Soal dirancang dalam bentuk *two-tier multiple choice* agar mampu mengukur kemampuan berpikir kritis siswa. Menurut Andi Rustandi dan Rismayanti (2021), tahap desain berfungsi untuk merancang spesifikasi produk dan menyusun komponen pembelajaran secara sistematis sebagai dasar pengembangan produk yang sesuai dengan tujuan pembelajaran (Sugiyono, 2021).

c. Tahap Pengembangan (*Development*)

Tahap pengembangan dilakukan dengan membuat produk awal instrumen berdasarkan materi bilangan. Produk yang telah dikembangkan kemudian akan divalidasi oleh dosen Pendidikan Matematika untuk menilai kesesuaian materi, konstruksi soal, bahasa dan kelayakan instrumen sebelum diuji coba kepada siswa. Sugiyono (2021), validasi ahli menjadi tahapan krusial dalam penelitian pengembangan guna memastikan kelayakan produk sebelum diimplementasikan di lapangan.

d. Tahap Implementasi (*Implementation*)

Tahap implementasi dilakukan dengan uji coba kepada siswa kelas VIII berjumlah 10 orang. Dengan mengerjakan soal yang telah dikembangkan, kegiatan ini bertujuan untuk memperoleh data empiris terkait validitas dan reliabilitas.

e. Tahap Evaluasi (*Evaluation*)

Tahap evaluasi dilakukan untuk menilai keseluruhan proses dan hasil pengembangan instrumen. Evaluasi dilakukan secara formatif pada setiap tahap untuk memperbaiki kekurangan produk.

2. Validitas Instrumen

Hasil validasi menunjukkan bahwa instrumen yang dikembangkan telah memenuhi kriteria kelayakan, dengan rincian pada tabel 2.

Tabel 2. Hasil Validasi Ahli

Aspek penilaian (%)				Jumlah	%	Kriteria
Kesesuaian Materi	Konstruksi soal	Bahasa	Kelayakan Instrumen			
84	84	72	88	328	82	Sangat Valid

Selain itu uji validitas butir soal menggunakan *IBM SPSS Statistics 16*, menunjukkan bahwa 10 butir soal memiliki nilai $r_{hitung} > r_{tabel}(0,632)$, sehingga butir soal dinyatakan valid. sehingga seluruh butir soal dinyatakan valid.

3. Reabilitas Instrumen

Uji reliabilitas instrumen dilakukan menggunakan *IBM SPSS Statistics 16* dengan metode *Cronbach's Alpha*. Hasil uji reliabilitas pada Gambar 1.

Reliability Statistics	
Cronbach's Alpha	N of Items
.968	10

Gambar 1. Hasil Uji Reabilitas Instrumen

Berdasarkan hasil analisis pada Gambar 1, nilai koefisien *Cronbach's Alpha* instrumen sebesar 0,968 dari 10 butir soal. Angka ini melampaui batas minimum yang disyaratkan, yaitu 0,60, sehingga instrumen dapat dikategorikan memiliki keandalan yang sangat tinggi dan dinyatakan layak sebagai alat evaluasi yang dapat dipercaya.

PEMBAHASAN

Hasil penelitian menunjukkan bahwa instrumen berbasis HOTS dalam bentuk soal *two-tier multiple choice* pada materi bilangan yang telah dikembangkan telah memenuhi kriteria valid dan reliabel. Hal ini menunjukkan bahwa proses pengembangan dengan model ADDIE terbukti efektif dalam menghasilkan produk yang layak digunakan. Pada tahap analisis, ditemukan adanya kesenjangan dalam hasil asesmen siswa yang terlihat dari capaian TKA tahun 2025 yang sangat rendah terutama pada soal yang menuntut kemampuan analisis dan mengaplikasikan konsep. Oleh karena itu pengembangan instrumen berbasis HOTS menjadi Solusi mengukur dan melatih kemampuan berpikir kritis siswa (Prasetya, 2025). Validitas instrumen yang tinggi menunjukkan bahwa isi, bahasa, dan tampilan media telah sesuai dengan kebutuhan pembelajaran. Instrumen tidak hanya mengukur jawaban benar atau salah, tetapi juga alasan di balik jawaban tersebut dipilih. Sehingga mampu mengukur tingkat berpikir kritis siswa. Untuk memvalidasi instrumen tes, dilakukan validasi kepada dosen Pendidikan Matematika sebagai ahli.

Berdasarkan Tabel 1, persentase validasi ahli sebesar 82% menempatkan instrumen pada kategori sangat valid. Dari hasil penilaian tersebut, Meskipun demikian, validator memberikan catatan pada aspek konstruksi soal yang belum sepenuhnya mencerminkan karakteristik HOTS. Hal ini sejalan dengan Lorin W. Anderson dalam Widana (2021) soal hots harus mampu membuat siswa menganalisis, mengevaluasi, dan menciptakan. Selain itu, siswa juga harus mampu mengelompokkan, memanipulasi, menghasilkan bentuk aktual dengan kreatif serta menjadikannya untuk memecahkan permasalahan baru. Pada butir 6, didapatkan masukan dan saran, yaitu soal terdapat kesalahan konsep, sehingga akan menimbulkan kekeliruan pada siswa.

NO	SEBELUM DIREVISI	SETELAH DIREVISI
6.	Pada bulan Ramadhan panitia zakat akan membagikan 3 jenis bahan makanan kepada Masyarakat yang membutuhkan: 1. Beras sebanyak hasil dari 256 – 224 bungkus 2. Gula sebanyak hasil dari 180+60 bungkus 3. Minyak goreng sebanyak hasil dari 96×5 bungkus Panitia ingin membagikan semua bahan kedalam beberapa paket sembako dengan jumlah yang sama banyak dan tanpa sisa. Seorang panitia menyatakan "Jumlah paket maksimal yang akan didapati Adalah $2^4 \times 3 \times 5$. Apakah pernyataan tersebut benar? A. Benar B. Salah Alasannya: A. Karena jumlah paket maksimal diperoleh dari FPB dengan mengambil pangkat terkecil dari faktor prima yang sama. B. Karena semua faktor prima yang muncul pada ketiga bilangan harus dikalikan agar hasil maksimal. C. Karena untuk mencari jumlah paket maksimal, kita mengambil pangkat terbesar dari setiap faktor prima. D. Karena jika bilangan terlihat besar dan sama-sama memiliki faktor 2, 3, dan 5, maka hasil perkalian semuanya pasti benar.	6. Pada bulan Ramadhan panitia zakat akan membagikan 3 jenis bahan makanan kepada Masyarakat yang membutuhkan: 1. Beras sebanyak hasil dari 256 – 224 bungkus 2. Gula sebanyak hasil dari 180+60 bungkus 3. Minyak goreng sebanyak hasil dari 96×5 bungkus Panitia ingin membagikan semua bahan kedalam beberapa paket sembako dengan jumlah yang sama banyak dan tanpa sisa. Seorang panitia menyatakan "Jumlah paket maksimal yang akan didapati Adalah $2^4 = 16$. Apakah jumlah paket sembako tersebut benar? A. Benar B. Salah Alasannya: A. Karena jumlah paket maksimal diperoleh dari FPB dengan mengambil pangkat terkecil dari faktor prima yang sama. B. Karena semua faktor prima dari ketiga bilangan harus dikalikan agar hasil maksimal. C. Karena untuk mencari jumlah paket maksimal, kita mengambil pangkat terbesar dari setiap faktor prima. D. Karena jika bilangan terlihat besar dan sama-sama memiliki faktor 2, 3, dan 5, maka hasil perkalian semuanya pasti benar.

Gambar 2. Soal Setelah dan Sebelum Revisi

Hasil butir soal yang sudah direvisi pada Gambar 1 dilakukan dengan memperbaiki konsep FPB yang masih keliru agar sesuai dengan konsep bilangan yang benar. sehingga dilakukan perbaikan pada konsep FPB agar sesuai dengan pemahaman bilangan yang tepat. Dengan revisi ini, siswa dapat memahami permasalahan dengan lebih tepat sehingga menunjukkan bahwa instrumen *two-tier multiple choice* memiliki kualitas yang memadai untuk digunakan dalam asesmen kemampuan berpikir tingkat tinggi siswa.

Untuk memberikan Gambaran yang lebih konkret mengenai proses revisi yang dilakukan, berikut disajikan perbandingan soal sebelum dan sesudah dikembangkan menjadi format *two tier multiple choice* berbasis HOTS.

NO	SEBELUM DIKEMBANGKAN	SETELAH DIKEMBANGKAN
6.	11. Terdapat tiga bilangan yang dinyatakan dengan $(332 - 32)$, $(82 + 296)$, dan (36×35). Bilangan apa saja yang merupakan faktor persekutuan ketiga bilangan tersebut? Pilih lebih dari satu jawaban! a. $2^2 \times 3^2 \times 5$ b. $2^2 \times 3^2 \times 7$ c. $2 \times 3^2 \times 5$ d. $2^3 \times 3^2 \times 7$ kunci jawaban: a. $2^2 \times 3^2 \times 5$ dan c. $2 \times 3^2 \times 5$	Pada bulan Ramadhan panitia zakat akan membagikan 3 jenis bahan makanan kepada Masyarakat yang membutuhkan: 1. Beras sebanyak hasil dari 256 – 224 bungkus 2. Gula sebanyak hasil dari 180+60 bungkus 3. Minyak goreng sebanyak hasil dari 96×5 bungkus Panitia ingin membagikan semua bahan kedalam beberapa paket sembako dengan jumlah yang sama banyak dan tanpa sisa. Seorang panitia menyatakan "Jumlah paket maksimal yang akan didapati Adalah $2^4 \times 3 \times 5$. Apakah pernyataan tersebut benar? A. Benar B. Salah Alasannya: A. Karena jumlah paket maksimal diperoleh dari FPB dengan mengambil pangkat terkecil dari faktor prima yang sama. B. Karena semua faktor prima yang muncul pada ketiga bilangan harus dikalikan agar hasil maksimal. C. Karena untuk mencari jumlah paket maksimal, kita mengambil pangkat terbesar dari setiap faktor prima. D. Karena jika bilangan terlihat besar dan sama-sama memiliki faktor 2, 3, dan 5, maka hasil perkalian semuanya pasti benar.

Gambar 3. Soal Sebelum dan Setelah dikembangkan

Berdasarkan Gambar 3, terlihat bahwa soal konvensional yang semula hanya mengukur jawaban akhir (*single-tier*) berhasil dikembangkan menjadi soal *two-tier* yang memuat konteks nyata dan menuntut siswa untuk menyertakan alasan atas jawaban yang dipilih. Transformasi ini menjadikan soal lebih mampu mengidentifikasi kualitas pemahaman konseptual siswa ini

sejalan dengan pendapat Andi Rustandi dan Rismayanti (2021) instrumen yang baik harus memenuhi aspek ketepatan materi dan kejelasan konsep.

Uji validitas butir soal menghasilkan nilai sebesar 0,623. Reliabilitas yang tinggi, yaitu 0,986, menunjukkan bahwa instrumen dapat memberikan hasil yang stabil jika digunakan sehingga layak dijadikan alat evaluasi pembelajaran. Secara keseluruhan, temuan penelitian ini menegaskan bahwa instrumen formatif *two-tier multiple choice* berbasis HOTS merupakan inovasi evaluasi yang efektif untuk mengukur dan mengembangkan kemampuan berpikir kritis siswa.

SIMPULAN

Penelitian ini berhasil menghasilkan produk berupa instrumen berbasis HOTS dalam format soal *two-tier multiple choice* pada materi bilangan untuk siswa SMP kelas VIII melalui pendekatan *Research and Development* (R&D) dengan model ADDIE. Berdasarkan hasil uji validitas dan reliabilitas yang telah dilakukan, instrumen ini dinyatakan layak digunakan sebagai alat evaluasi dalam konteks Tes Kemampuan Akademik (TKA), khususnya untuk mengukur kemampuan berpikir tingkat tinggi siswa pada materi bilangan. Hasil validasi ahli menunjukkan persentase 82% (sangat valid), seluruh butir soal valid dengan $r_{hitung} > r_{tabel}$ 0,632, dan koefisien reliabilitas 0,986 yang tergolong sangat tinggi. Dengan demikian, instrumen ini layak direkomendasikan sebagai salah satu alternatif evaluasi yang mampu mengidentifikasi kualitas pemahaman konseptual dan kemampuan berpikir kritis siswa secara lebih mendalam.

SARAN

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran dikemukakan sebagai berikut. Bagi guru matematika, instrumen *two-tier multiple choice* berbasis HOTS yang telah dikembangkan ini dapat dimanfaatkan sebagai alternatif alat evaluasi yang mampu mengungkap tidak hanya ketepatan jawaban, tetapi juga kualitas pemahaman konseptual dan kemampuan berpikir kritis siswa pada materi bilangan.

Bagi peneliti selanjutnya, mengingat uji coba dalam penelitian ini masih terbatas pada 10 siswa di satu sekolah, disarankan untuk memperluas cakupan penelitian dengan melibatkan subjek yang lebih besar dan beragam dari berbagai wilayah, sekaligus mengembangkan instrumen serupa pada materi matematika lainnya seperti aljabar, geometri, maupun statistika.

Bagi pemangku kebijakan pendidikan, perlu dipertimbangkan integrasi format soal *two-tier multiple choice* berbasis HOTS ke dalam kerangka instrumen TKA secara nasional, karena format ini dinilai lebih akurat dalam mengidentifikasi kesenjangan pemahaman konseptual siswa dibandingkan soal pilihan ganda konvensional.

Terakhir, bagi pihak sekolah, kepala sekolah dan tim kurikulum disarankan untuk mendorong guru dalam mengembangkan instrumen evaluasi berbasis HOTS secara mandiri melalui pelatihan dan penyediaan akses referensi yang relevan demi meningkatkan kualitas evaluasi pembelajaran secara berkelanjutan.

DAFTAR PUSTAKA

Aldi, S., Adnan, Ismail, & Dzulqarnain, A. F. (2022). Uji Kepraktisan Lembar Kerja Peserta Didik Elektronik Berbasis Keterampilan . *Bioedusiana: Jurnal Pendidikan Biologi*, 7(1), 128-143.

- Hidayat, m. (2021). Analisis miskonsepsi siswa SMP pada materi bilangan: Penyebab hafalan prosedur tanpa pemahaman konsep. *Jurnal Pendidikan Matematika Indonesia*, 15(1), 45-58.
- Kesorn, N., Junpeng, P., Marwing, M., Pongboriboon, K., Tang, K., Bathia, S., & Wilson, M. (2020). Development of an assessment tool for mathematical reading, analytical thinking and mathematical writing. *International Journal of Evaluation and Research in Education (IJERE)*, 9(4), 955-962.
- Maudina, A. V., & Zuhrie, M. S. (2020). Analisis validitas pengembangan media pembelajaran mobile application pada mata pelajaran sistem komputer kelas x. *Jurnal pendidikan teknik elektro*.
- Nafi'ah, Z., Mustikasari, V. R., & Pratiwi, N. (2020). Pengembangan instrumen tes pilihan ganda dua tingkat untuk mengukur kemampuan berpikir tingkat tinggi peserta didik SMP pada materi suhu dan kalor. *JIPVA(Jurnal pendidikan IPA Veteran)*, 3(2), 116-126.
- Prasetya, R. D. (2025). Pengembangan instrumen penilaian berbasis hots (Higher Order Thinking Skills). *Jurnal pendidikan teknik mesin*, 14(1).
- Putri Hapsari, Y. H., & Gularso, D. (2024). Pengembangan Kepraktisan Instrumen Penilaian Berbasis Keterampilan Berpikir Kritis Siswa Kelas IV SD Negeri Wirosaban. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 8(2), 1351-1357.
- Rahayu, A., Destrineli, & Eka Putri, A. G. (2026). Pengembangan instrumen formatif two-tier multiple choice. *Jurnal Ilmiah Pendidikan Dasar*, 11(2), 240-251.
- Rahman, A. (2022). Pengembangan soal berkualitas untuk evaluasi pembelajaran matematika: Validitas, reliabilitas, dan objektivitas. *Jurnal Evaluasi Pendidikan*, 12(2), 78-92.
- Rahmaniasan, V., Berlian, L., & Suryani, D. I. (2022). Pengembangan Instrumen Tes Two-Tier Multiple Choice Berbasis HOTS Tema Pemanfaatan Gelombang untuk Menumbuhkan Kemampuan Berpikir Tingkat Tinggi Siswa. *Jurnal Pendidikan MIPA*, 12(3), 929-935.
- Rahmawati, A., & Widodo, S. (2021). Pengembangan instrumen two tier multiple choice untuk mengidentifikasi pemahaman konsep dan miskonsepsi siswa matematika SMP. *Jurnal Pendidikan Matematika Indonesia*, 15(2), 123-135.
- Rohana. (2021). Pengembangan soal HOTS untuk meningkatkan kemampuan berpikir kritis peserta didik. *Bioilmi: Jurnal Pendidikan*, 7(1), 39-45.
- Rustandi, A., & Rismayanti. (2021). Penerapan Model ADDIE dalam pengembangan media pembelajaran di SMPN 22 Kota Samarinda. *Jurnal Fasion*, 12(1), 45-53.
- Sugiyono. (2021). *Metode penelitian dan pengembangan (Reserch and Development/R&D)*. Bandung: Alfabeta.
- Suhardi, I. (2022). Perangkat Instrumen Pengembangan Paket Soal Jenis Pilihan Ganda Menggunakan Pengukuran Validitas Konten Formula Aiken's V. *Jurnal Pendidikan Tambusai*, 6(1), 4158-4170.
- Suryani, & Widodo, S. (2023). Pengembangan instrumen two-tier multiple choice berbasis HOTS untuk mengukur kemampuan berpikir tingkat tinggi siswa kelas 8 SMP materi bilangan. *Jurnal Pendidikan dan Aplikasi Kimia*, 5(2), 15895-15905.
- Widana, I. W. (2021). Pengembangan model asesmen berbasis kompetensi untuk pengukuran capaian belajar siswa pada era kurikulum merdeka. *Jurnal Penelitian dan Pengembangan Pendidikan*, 6(1), 1-15.