

Pengembangan Instrumen Diagnostik Berformat *Four Tier Multiple Choice* pada Materi Operasi Aljabar

Afina Fidaroin Naja*, Mutiara Arlisyah Putri Utami

Program Studi Tadris Matematika, Universitas Islam Negeri Maulana Malik Ibrahim Malang
e-mail: *fidaafina23@gmail.com

ABSTRACT. This study aims to develop a diagnostic test instrument in the form of a four-tier multiple-choice algebraic operations material for grade VII SMP/MTs that is feasible, practical, and can also determine students' conceptions. This study was conducted at MTs Almaarif 01 Singosari using the Research and Development (R&D) method, as outlined in the Borg and Gall model. The stages of implementing the research consisted of eight key steps, namely: potential and problems, data collection, product design, design validation, design revision, small-scale trials involving six students from Grade VII A, and product revision and usage tests involving 35 students from Grade VII B. The study yielded 13 feasible and practical questions. The results of the feasibility test, as validated by instrument experts, materials, and practitioner validators, were 85%, 81%, and 95%, respectively, falling within the very feasible category. Regarding the results of the practicality test, the average response from the student questionnaire was 80.75% in the practical category. The results of the diagnostic test in the trial usage showed that the percentage of students who understood the concept was 42%, those who did not understand the idea was 33%, those with misconceptions were 20%, and those with errors were 5%. In general, the study's results indicate that the diagnostic test instrument, presented in a four-tier multiple-choice format, is feasible and practical for use in learning algebraic operations material for grade VII students in junior high school/Islamic junior high school. This instrument can effectively determine students' conceptions.

Keywords: algebraic operations; diagnostic instruments; four-tier multiple-choice

ABSTRAK. Penelitian ini bertujuan untuk mengembangkan instrumen tes diagnostik berformat *four tier multiple choice* pada materi operasi aljabar kelas VII SMP/MTs yang layak, praktis dan juga dapat mengetahui konsepsi peserta didik. Penelitian ini dilaksanakan di MTs Almaarif 01 Singosari dengan metode *Research and Development* (R&D) model Borg and Gall. Tahapan pelaksanaan penelitian melalui delapan tahapan, yaitu: potensi dan masalah, pengumpulan data, desain produk, validasi desain, revisi desain, uji coba skala kecil dengan melibatkan 6 peserta didik kelas VII A, revisi produk dan uji pemakaian dengan melibatkan 35 peserta didik kelas VII B. Hasil dari penelitian adalah diperolehnya 13 butir soal yang layak dan praktis. Hasil uji kelayakan dari validasi ahli instrumen, materi dan validator praktisi masing-masing sebesar 85%, 81 % dan 95% dengan kategori sangat layak. Sedangkan untuk hasil uji kepraktisan, hasil rata-rata angket respons peserta didik sebesar 80,75% dengan kategori praktis. Hasil tes diagnostik pada uji coba pemakaian diperoleh persentase peserta didik yang paham konsep sebesar 42%, tidak paham konsep sebesar 33%, miskonsepsi sebesar 20% dan error sebesar 5%. Secara umum, hasil penelitian menunjukkan bahwa instrumen tes diagnostik berformat *four tier multiple choice* pada materi operasi aljabar kelas VII SMP/MTs layak dan praktis untuk digunakan dalam pembelajaran dan dapat mengetahui konsepsi peserta didik.

Kata Kunci: *four tier multiple choice*; instrumen diagnostic; operasi aljabar

PENDAHULUAN

Tes diagnostik merupakan alat yang sangat penting untuk mengidentifikasi kesalahan konsep peserta didik dalam pembelajaran, termasuk dalam memahami materi aljabar. Menurut Departemen

Pendidikan Nasional (2007), tes diagnostik digunakan untuk kemampuan peserta didik. Elvia dkk. (2021) dan Sriyanti dkk. (2019) mengemukakan bahwa tes diagnostik secara efektif dapat mendeteksi miskonsepsi yang sering terjadi, terutama dalam pembelajaran matematika. Selain itu menurut Mardapi (2012), tes diagnostik sangat penting dalam mengungkap kesalahan konsep peserta didik dan dapat membantu guru untuk merencanakan strategi pembelajaran yang lebih baik.

Salah satu bentuk tes diagnostik yang dianggap efektif adalah *four-tier multiple choice*. Tes ini adalah pengembangan dari tes diagnostik *three-tier* dengan menambahkan komponen tingkat keyakinan peserta didik terhadap jawaban dan alasan yang mereka pilih (Rusilowati, 2015). Menurut Amin dkk. (2016), format *four-tier multiple choice* dapat memberikan wawasan lebih mendalam tentang pemahaman peserta didik, membantu mengidentifikasi kesalahan konsep, serta mengarahkan pembelajaran pada materi yang memerlukan perhatian khusus.

Materi operasi aljabar memiliki peran penting dalam kurikulum matematika, karena merupakan dasar untuk memahami konsep matematika lanjutan (Sari & Afriansyah, 2020). Peserta didik sulit dalam memahami konsep-konsep aljabar terutama yang berkaitan dengan operasi bilangan bulat dan manipulasi bentuk aljabar. Kesalahan umum yang ditemukan meliputi penjumlahan suku tak sejenis, pengalihan konstanta dengan variabel berbeda tanda, dan pengalihan konstanta pada bentuk aljabar binomial (Wahid dkk., 2015). Menurut Permatasari dkk. (2015), kurangnya pemahaman konsep aljabar menyebabkan 61,59% peserta didik mengalami kesulitan memahami konsep tersebut.

Observasi awal di MTs Al-Maarif 01 Singosari menunjukkan bahwa peserta didik kelas VII sering mengalami kesulitan dalam memahami materi aljabar, karena lebih sering menghafal materi daripada memahami konsepnya secara menyeluruh. Selain itu, sekolah ini belum memiliki instrumen tes diagnostik untuk membantu mengidentifikasi kesalahan konsep. Guru hanya menggunakan penilaian formatif dan sumatif, sehingga kurang mampu mengidentifikasi miskonsepsi secara mendalam.

Penelitian sebelumnya telah mengembangkan berbagai instrumen diagnostik dalam pembelajaran matematika, seperti yang dilakukan oleh Nabilah dkk. (2019), Putranto dkk. (2020) dan Agustin dkk. (2022). Namun, penelitian tersebut memiliki fokus berbeda dalam aspek materi, model pengembangan, dan tujuan instrumen. Nabilah dkk. (2019) menyarankan pengembangan instrumen diagnostik untuk materi matematika lainnya guna mengeksplorasi miskonsepsi yang belum teridentifikasi.

Berdasarkan latar belakang tersebut, pengembangan instrumen diagnostik *four-tier multiple choice* pada materi operasi aljabar untuk kelas VII SMP/MTs merupakan tujuan dari penelitian ini. Diharapkan instrumen ini dapat mendeskripsikan proses pengembangan, mengukur kelayakan dan kepraktisannya, serta membantu mengidentifikasi konsepsi peserta didik secara lebih rinci.

METODE

Penelitian ini menggunakan metode *Research and Development* (R&D) yang membantu guru membuat produk yang digunakan peserta didik (Fatchiyah & Utami, 2022). Prosedur R&D menggunakan model Borg and Gall mencakup delapan langkah berikut: (1) potensi dan masalah, (2) pengumpulan data, (3) desain produk, (4) validasi desain, (5) revisi desain, (6) uji coba produk, (7) revisi produk yang dikembangkan, serta (8) uji coba pemakaian. Subjek penelitian melibatkan 6 peserta didik kelas VII A untuk uji skala kecil dan 35 peserta didik kelas VII B pada uji coba pemakaian. Instrumen pengumpulan data yang digunakan adalah angket penilaian yang terdiri dari angket kelayakan dan kepraktisan. Angket kelayakan instrumen yang digunakan berdasarkan 3 aspek yaitu materi, konstruksi dan bahasa. Beberapa indikator yang diadopsi dari penelitian Harahap & Novita (2020), sebagaimana tercantum dalam Tabel 1. berikut.

Tabel 1. Kisi-kisi Angket Kelayakan Instrumen

Aspek yang dinilai	Indikator
Aspek Materi	Kesesuaian antara butir soal dan sub materi Butir soal disusun berdasarkan indikator soal Setiap soal konsisten dengan urutan materi pembelajaran Pertanyaan memiliki batasan yang jelas, didukung oleh jawaban dan respon yang lengkap.
Aspek Konstruksi	Petunjuk pengerjaan tes diagnostik jelas dan terperinci Kesesuaian antara butir soal, indikator soal dan capaian pembelajaran Miskonsepsi peserta didik dapat diidentifikasi pada butir soal tes diagnostik Pilihan alasan sebagai pengecoh bersifat rasional Pengecoh pada pilihan alasan bersifat rasional dan homogen dengan jawaban pertama Kesesuaian tabel, grafik, gambar, dan sejenisnya berdasarkan konteks masalah yang diberikan
Aspek Bahasa	Kalimat yang digunakan dalam soal menggunakan kaidah yang baik dan benar Penafsiran ganda tidak ditimbulkan dalam pertanyaan Pertanyaan-pertanyaan dalam butir tes ditulis dengan jelas dan komunikatif

Setelah validator menyatakan instrumen layak untuk digunakan, langkah berikutnya adalah melakukan uji coba produk kepada peserta didik dengan memberikan angket respon peserta didik terhadap kepraktisan instrumen. Angket ini diadopsi dari penelitian yang dilakukan Nai (2019) sebagaimana yang dicantumkan pada Tabel 2. berikut.

Tabel 2. Kisi-kisi Angket Kepraktisan Instrumen

Aspek yang dinilai	Indikator
Aspek Perencanaan	Petunjuk tes diagnostik dituliskan dengan jelas. Pengidentifikasian kesalahan konsep peserta didik dapat menggunakan tes diagnostik. Tes diagnostik memenuhi kriteria indikator pembelajaran. Pelaksanaan tes diagnostik tidak memerlukan pendampingan ahli.
Aspek Efisiensi	Alokasi waktu dalam pengerjaan tes diagnostik sesuai dengan petunjuk soal. Penggunaan instrumen tes tidak memerlukan biaya besar. Pemanfaatan sarana di sekolah dalam penggunaan instrumen tes diagnostik. Pelaksanaan tes diagnostik tidak membutuhkan tempat khusus.

Pengumpulan data menggunakan pendekatan kualitatif dan kuantitatif. Proses untuk menganalisis data mencakup analisis kualitatif dan kuantitatif. Teknik analisis data yang diimplementasikan adalah analisis deskriptif yang dilakukan dengan cara berikut:

Analisis Angket Kelayakan Instrumen

Para ahli melengkapi pengisian lembar validasi dalam format tabel kelayakan produk sebagai acuan dalam melakukan revisi pada setiap komponen. Lembar validasi disusun dan dianalisis guna mengetahui kelayakan dari produk yang dikembangkan. Hasil angket validasi dapat dianalisis menggunakan persentase kelayakan, dengan rumus sebagai berikut:

$$P = \frac{f}{N} \times 100\%$$

Keterangan:

- P = Hasil persentase data angket
 f = Jumlah skor yang diperoleh
 N = Skor Maksimal

Tahap akhir yaitu menyimpulkan hasil perhitungan berdasarkan aspek dengan berpedoman pada Tabel 3. berikut.

Tabel 3. Kriteria Kelayakan Instrumen

Persentase %	Kriteria Interpretasi Kelayakan
$81\% \leq x \leq 100\%$	Sangat Layak
$61\% \leq x \leq 80\%$	Layak
$41\% \leq x \leq 60\%$	Cukup Layak
$21\% \leq x \leq 40\%$	Tidak layak

Sumber: Saputri dkk. (2022)

Analisis Angket Kepraktisan Instrumen

Angket kepraktisan bertujuan untuk uji respon peserta didik terhadap tes yang diberikan. Langkah ini melibatkan pemberian kuesioner kepada setiap peserta didik dengan menggunakan skala Likert. Selanjutnya, hasil kuesioner dihitung setiap item pernyataan dengan menggunakan rumus yang telah ditentukan:

$$P = \frac{f}{N} \times 100\%$$

Keterangan:

 P = Hasil persentase data angket f = Jumlah skor yang diperoleh N = Skor Maksimal

Langkah selanjutnya yaitu menyimpulkan hasil perhitungan untuk mengetahui kepraktisan instrument. Kriteria kepraktisan mengacu pada Tabel 4. berikut.

Tabel 4. Kriteria Kepraktisan Instrumen

Persentase %	Kriteria Interpretasi Kepraktisan
$81\% \leq x \leq 100\%$	Sangat Praktis
$61\% \leq x \leq 80\%$	Praktis
$41\% \leq x \leq 60\%$	Cukup Praktis
$21\% \leq x \leq 40\%$	Tidak Praktis

Sumber: Saputri dkk. (2022)

Analisis Hasil Tes

Data yang digunakan berupa data kuantitatif yang diperoleh dari hasil tes diagnostik berbentuk *four-tier multiple choice*. Tes ini disusun untuk mengukur tidak hanya kemampuan siswa dalam memilih jawaban yang benar, tetapi juga untuk menilai alasan yang mendasari jawaban tersebut serta tingkat keyakinan siswa terhadap pilihannya. Dengan pendekatan ini, analisis tidak hanya berfokus pada hasil akhir, tetapi juga memperhatikan proses berpikir siswa dalam menjawab soal.

Hasil tes dianalisis untuk menilai sejauh mana siswa memahami konsep yang telah dipelajari. Analisis dilakukan dengan mencocokkan pilihan jawaban, alasan yang diberikan, serta keyakinan siswa terhadap jawabannya untuk mengidentifikasi tingkat pemahaman yang dimiliki. Pendekatan ini memungkinkan pendidik untuk mendeteksi adanya miskonsepsi atau ketidaktahuan terhadap suatu konsep secara lebih mendalam dan sistematis.

Pemahaman konsep siswa digolongkan ke dalam empat kategori, yaitu: (1) pemahaman konsep yang baik, (2) tidak memahami konsep, (3) miskonsepsi, serta (4) error. Klasifikasi ini bertujuan untuk memberikan gambaran yang lebih komprehensif mengenai profil pemahaman siswa, sehingga dapat menjadi dasar bagi perbaikan strategi pembelajaran. Dengan mengidentifikasi kategori pemahaman secara tepat, guru dapat merancang intervensi yang sesuai guna meningkatkan kualitas pembelajaran serta mendorong pencapaian hasil belajar yang lebih optimal. Adapun kombinasi jawaban *four-tier multiple choice* yang digunakan pada penelitian sebagaimana yang tercantum dalam Tabel 5 berikut.

Tabel 5. Kombinasi Jawaban *Four-Tier Multiple Choice*

Kategori	Kombinasi Jawaban			
	Jawaban	Tingkat Keyakinan Jawaban	Alasan	Tingkat Keyakinan alasan
Paham Konsep	Benar	Yakin	Benar	Yakin
Tidak Paham Konsep	Benar	Yakin	Benar	Tidak
	Benar	Yakin	Salah	Tidak
	Benar	Tidak	Benar	Tidak
	Benar	Tidak	Benar	Tidak
	Benar	Tidak	Salah	Tidak
	Salah	Yakin	Benar	Tidak
	Salah	Yakin	Salah	Tidak
	Salah	Tidak	Benar	Tidak
Miskonsepsi	Salah	Tidak	Salah	Tidak
	Benar	Yakin	Salah	Yakin
	Benar	Tidak	Salah	Yakin
	Salah	Yakin	Salah	Yakin
Error	Salah	Yakin	Benar	Yakin
	Salah	Tidak	Benar	Yakin

Sumber: Ismail dkk. (2015)

Jawaban peserta didik dianalisis untuk menghitung persentase mereka dalam setiap kategori (memahami konsep, tidak memahami konsep, mengalami miskonsepsi, dan *error*) dengan menggunakan rumus:

$$P = \frac{f}{N} \times 100\%$$

Keterangan:

- P = Angka persentase kelompok
 f = Jumlah skor yang diperoleh pada setiap kelompok
 N = Jumlah seluruh peserta didik

Kemudian deskripsi data pemahaman konsep peserta didik mengacu pada Tabel 6. berikut.

Tabel 6. Pendeskripsian Data Tingkat Pemahaman Konsep

Persentase %	Kriteria Interpretasi
80 – 100	Sangat Baik
66 – 79	Baik
56 – 65	Cukup
46 – 55	Kurang
30 – 45	Gagal

Sumber : Yakubi dkk. (2017)

Hasil jawaban peserta didik dengan melalui instrumen diagnostik *four-tier* mengidentifikasi peserta didik yang memahami konsep, tidak memahami konsep, mengalami miskonsepsi, serta *error* pada materi operasi aljabar kelas VII SMP/MTs.

HASIL DAN PEMBAHASAN

Pengembangan instrumen dilakukan pada bulan Februari 2024 di MTs Almaarif 01 Singosari Malang. Proses ini menggunakan metode *Research and Development* dengan model Borg and Gall sebagaimana dijelaskan oleh Sugiyono. Pengembangan ini terdiri dari delapan tahapan berikut: (1) potensi dan masalah, (2) pengumpulan data, (3) desain produk, (4) validasi desain, (5) revisi

desain, (6) uji coba produk, (7) revisi produk yang dikembangkan, dan (8) uji coba pemakaian. Beberapa tahapan yang telah dilaksanakan sebagai berikut:

Potensi dan Masalah

Hasil observasi di MTs Almaarif 01 Singosari menunjukkan beberapa bahwa guru hanya menggunakan penilaian formatif dan sumatif berformat pilihan ganda *one tier* dan esai untuk menilai hasil belajar peserta didik saja. Guru belum pernah mengidentifikasi bagaimana pemahaman konsep peserta didik, sehingga guru membutuhkan instrumen tes yang dapat membantu menganalisis pemahaman konsep tersebut. Masalah ini menyebabkan peserta didik sering mengalami kesulitan dalam memahami materi aljabar. Peserta didik seringkali kebingungan saat menghadapi soal-soal aljabar yang lebih kompleks, meskipun mereka mampu memahami penjelasan konsep. Salah satu dampaknya yaitu menyebabkan peserta didik tidak memahami materi aljabar dengan kecenderungan peserta didik hanya menghafal materi tanpa memahami konsep dasar. Hal ini ditunjukkan dengan rata-rata nilai materi operasi aljabar kelas VII hanya mencapai 56, di bawah standar kriteria ketuntasan minimum.

Pengumpulan Data

Tahap ini dilaksanakan dengan tujuan mengatasi potensi dan masalah yang ditemukan dari hasil observasi. Langkah ini mencakup studi literatur melalui pembacaan buku dan jurnal terkait pengembangan instrumen diagnostik *four-tier multiple choice* serta referensi pendukung lainnya. Data tambahan diperoleh melalui wawancara dengan guru matematika kelas VII di MTs Almaarif 01 Singosari. Dari hasil wawancara, ditemukan bahwa: 1) Hasil belajar rendah dikarenakan kurangnya minat peserta didik dalam belajar serta kecenderungan menghafal materi tanpa memahami konsep, 2) Guru hanya memanfaatkan tes formatif dan sumatif sebagai alat untuk menilai pencapaian hasil belajar peserta didik, 3) Guru membutuhkan instrumen tes yang dapat mengetahui kesalahan konsep peserta didik, 4) Materi operasi aljabar memiliki nilai rata-rata rendah yang menunjukkan kurangnya pemahaman konsep. Hal ini selaras dengan penelitian yang dilakukan oleh Sriyanti dkk. (2019), bahwa tes diagnostik adalah alat untuk mengidentifikasi kesalahan konsep. Di lain pihak, penting bagi peserta didik untuk mempelajari materi operasi aljabar (Abidin dkk., 2019; Yuliyani, 2016). Sementara itu, diketahui bahwa ditemui adanya tingkat kesulitan yang tinggi pada materi aljabar akibat kurangnya pemahaman konsep peserta didik (Permatasari dkk., 2015). Berdasarkan hal tersebut maka dikembangkan instrumen diagnostik berformat *four tier multiple choice* pada materi operasi aljabar.

Desain Produk

Perancangan soal tes diagnostik *four-tier multiple choice* dilakukan secara sistematis dengan beberapa tahap, yaitu pembuatan kisi-kisi soal, pembuatan petunjuk pengerjaan, penulisan butir soal, hingga pembuatan kunci jawaban. Kisi-kisi soal dirancang untuk menjamin validitas isi dan relevansi dengan kurikulum serta kompetensi yang diukur dalam tes tersebut (Evendi, 2020). Kisi-kisi ini menjadi dasar dalam menentukan cakupan materi dan indikator ketercapaian pembelajaran yang ingin diukur.

Petunjuk soal disusun dengan jelas dan mudah dipahami oleh peserta didik, sehingga dapat mengikuti prosedur pengerjaan dengan baik. Tes diagnostik terdiri dari empat tingkatan, yaitu: (1) Pilihan ganda untuk menjawab soal utama, (2) Pilihan alasan yang mendukung jawaban utama, (3) Penilaian tingkat keyakinan terhadap jawaban utama, serta (4) Penilaian tingkat keyakinan terhadap alasan yang dipilih.

Capaian pembelajaran materi operasi aljabar kelas VII SMP/MTs menjadi dasar dalam penyusunan soal. Indikator ketercapaian meliputi kemampuan peserta didik dalam: menunjukkan unsur-unsur bentuk aljabar (konstanta, koefisien, variabel dan suku), menentukan bentuk aljabar,

menyelesaikan operasi penjumlahan dan pengurangan, perkalian dan pembagian bentuk aljabar, dan memodelkan dan menyelesaikan permasalahan sehari-hari yang berkaitan dengan aljabar.

Opsi pengecoh dalam soal dirancang berdasarkan analisis kesalahan konsep yang umum terjadi pada peserta didik, seperti kesalahan dalam menjumlahkan suku tak sejenis atau mengalikan konstanta dengan variabel berbeda tanda. Opsi alasan juga disusun untuk mengungkap pola berpikir peserta didik dan mengidentifikasi miskonsepsi yang mungkin terjadi. Tingkat keyakinan pada jawaban utama dan alasan dipertimbangkan sebagai komponen penting dalam instrumen ini, karena memberikan wawasan lebih mendalam tentang sejauh mana peserta didik memahami konsep yang diuji. Dengan demikian, instrumen ini tidak hanya mengukur hasil akhir berupa jawaban benar atau salah, tetapi juga mengidentifikasi keyakinan peserta didik terhadap pemahaman konsep mereka. Kunci jawaban dibuat berdasarkan analisis terhadap indikator pembelajaran dan materi yang diujikan, sehingga memastikan kesesuaian antara soal, jawaban, dan tujuan pembelajaran yang ingin dicapai.

Validasi desain

Setelah menyelesaikan tahap desain produk, langkah selanjutnya adalah melakukan validasi terhadap instrumen yang telah dikembangkan. Validasi ini bertujuan untuk menguji kelayakan instrumen sebelum digunakan dalam tahap implementasi. Proses validasi dilakukan oleh tiga orang validator yang terdiri atas dua ahli akademis dan satu praktisi.

Dua validator ahli merupakan dosen dari Program Studi Tadris Matematika, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Keduanya bertindak sebagai ahli instrumen dan ahli materi. Sementara itu, satu validator lainnya adalah seorang guru matematika di MTs Almaarif 01 Singosari, yang berperan sebagai validator praktisi. Ketiga validator ini dipilih berdasarkan latar belakang keilmuan dan pengalaman mereka yang relevan dengan pengembangan instrumen pembelajaran matematika.

Validasi instrumen dilakukan dengan mengacu pada tiga aspek utama, yaitu aspek materi, aspek konstruksi, dan aspek bahasa. Ketiga aspek tersebut kemudian dikembangkan menjadi 13 butir pernyataan dalam lembar penilaian validasi. Masing-masing butir bertujuan untuk menilai kesesuaian isi materi, kejelasan struktur dan sistematika penyajian, serta ketepatan penggunaan bahasa dalam instrumen yang dikembangkan. Adapun hasil dari validasi kelayakan instrumen oleh ketiga validator tersebut disajikan pada Tabel 7 berikut.

Tabel 7. Hasil Validasi Kelayakan Instrumen

No	Validator	Persentase	Kategori
1.	Ahli Instrumen	85%	Sangat Layak
2.	Ahli Materi	81%	Sangat Layak
3.	Praktisi	95%	Sangat Layak

Persentase dari validator ahli instrumen yaitu 85%, validator ahli materi 81%, dan validator praktisi 95%. Persentase tersebut termasuk dalam kategori kelayakan instrumen dengan interval 81-100%, yang menunjukkan bahwa instrumen tes diagnostik *four-tier multiple choice* pada materi operasi aljabar sangat layak. Hasil penelitian tentang kelayakan tes asesmen menunjukkan validasi dari ahli materi 81% dan ahli instrumen sebesar 82%, dimana keduanya masuk dalam kategori sangat layak (Saputri, dkk., 2022). Proses validasi oleh para ahli penting untuk mengevaluasi produk yang telah dikembangkan. Saran dan masukan para validator dapat membantu mengidentifikasi kekurangan dan kelemahan produk (Sarip, dkk., 2022). Hasil akhir menunjukkan bahwa instrumen diagnostik berformat *four-tier multiple choice* pada materi operasi aljabar dinyatakan sangat layak digunakan.

Revisi desain

Hasil validasi menunjukkan bahwa instrumen tes diagnostik berbentuk *four-tier multiple choice* pada materi operasi aljabar dinyatakan valid dan layak digunakan, meskipun memerlukan beberapa

revisi minor. Validasi dilakukan oleh ahli materi dan evaluasi, yang memberikan masukan terkait kejelasan isi, struktur soal, dan kesesuaian dengan tujuan pengukuran. Revisi dilakukan berdasarkan saran validator, meliputi perbaikan redaksi pertanyaan agar lebih jelas, penyesuaian pilihan jawaban, serta perbaikan alasan memilih jawaban untuk menggambarkan pemahaman konseptual secara lebih akurat. Selain itu, pengecoh juga disempurnakan agar tetap logis dan efektif dalam mengungkap miskonsepsi peserta didik. Dengan demikian, setelah proses validasi dan revisi, instrumen dinyatakan siap untuk digunakan pada tahap selanjutnya.

Uji coba skala kecil

Uji coba ini dilaksanakan di kelas VII B MTs Almaarif 01 Singosari tahun ajaran 2023/2024 dengan melibatkan 6 peserta didik yang telah mempelajari materi operasi aljabar. Uji coba ini dilaksanakan selama 2 jam pelajaran pada hari Senin, 19 Februari 2024. Tes diagnostik berformat *four-tier multiple choice* dengan 15 soal yang dinyatakan layak oleh validator sehingga dapat digunakan dalam uji coba ini. Sebelum mengerjakan soal, peserta didik diberikan pengarahan mengenai tata cara menjawab sesuai dengan petunjuk tes diagnostik. Setelah uji coba skala kecil selesai, aplikasi *Iteman 4.3* digunakan untuk menganalisis instrumen yang bertujuan untuk mengetahui kualitas instrumen yang diujikan, meliputi:

Validitas

Analisis validitas digunakan untuk mengetahui valid atau tidaknya suatu butir soal. Hasil analisis validitas ditunjukkan dalam hasil total *point biserial* pada *Iteman 4.3*. Berdasarkan hasil uji validitas instrumen pada aplikasi *Iteman 4.3* yang menghasilkan keseluruhan butir soal memiliki nilai $r_{pb} \geq 0,2$, maka dapat disimpulkan bahwa 15 soal yang dikembangkan dinyatakan valid (Syahlani & Setyorini, 2021).

Reliabilitas

Setelah melalui proses validasi item, langkah selanjutnya merupakan uji reliabilitas. Hal ini dilakukan untuk mengetahui apakah alat ukur yang digunakan konsisten dalam memberikan hasil pengukuran yang stabil dari waktu ke waktu. Hasil analisis reliabilitas pada *Iteman 4.3* dapat dilihat berdasarkan nilai Alpha Cronbach (Guyer & Thompson, 2013). Hasil uji reliabilitas instrumen disajikan dalam Tabel 8. berikut.

Tabel 8. Hasil Uji Reliabilitas

Cronbach's Alpha	N of Items
0.917	15

Sumber: Diolah dengan Iteman 4.3

Hasil reliabilitas ditunjukkan dengan besarnya koefisien yang sangat reliabel untuk suatu tes adalah 0,80 – 1,00 (Zein & Darto, 2017). Instrumen tes yang dianalisis menggunakan aplikasi *Iteman 4.3* menghasilkan cronbach's alpha sebesar 0,917, maka tes diagnostik memiliki kriteria sangat reliabel.

Tingkat kesukaran

Analisis tingkat kesukaran digunakan untuk menilai apakah soal yang disusun sesuai dengan kemampuan peserta didik. Jika sebagian besar peserta didik menjawab dengan benar, tingkat kesukaran rendah dan sebaliknya jika sebagian besar peserta didik menjawab dengan salah maka tingkat kesukarannya tinggi (Huda & Wahyuni, 2020). Analisis ini digunakan untuk mengkategorikan soal yang mudah, sedang, atau sulit. Soal yang ideal yaitu soal yang tidak terlalu mudah maupun terlalu sulit, karena yang terlalu mudah tidak mampu merangsang kemampuan peserta didik, sementara yang terlalu sulit dapat membuat peserta didik putus asa (Zein & Darto,

2017). Rekapitulasi tingkat kesukaran pada analisis butir soal pada uji coba kelompok kecil disajikan pada Tabel 9. berikut.

Tabel 9. Persentase Tingkat Kesukaran

Nomor	Kriteria	Persentase
13,14,15	Sulit	20%
1,3,4,5,6,7,9,10,11,12	Sedang	67%
2,8	Mudah	13%

Daya Pembeda

Evaluasi terhadap kemampuan suatu soal dalam memisahkan peserta didik dengan prestasi tinggi dari yang memiliki prestasi rendah merupakan fungsi dari daya pembeda (Zein & Darto, 2017). Hasil perhitungan daya pembeda dikategorikan dengan jelek, cukup, baik, dan sangat baik. Rekapitulasi daya pembeda pada analisis butir soal dapat dilihat pada Tabel 10. berikut.

Tabel 10. Persentase Daya Pembeda

Nomor	Kriteria	Persentase
2	Jelek	7%
8	Cukup	7%
-	Baik	0%
1,3,4,5,7,9,11,12,13,14,15	Sangat Baik	86%

Keberfungsian pengecoh

Hasil analisis keberfungsian pengecoh soal tes diagnostik berformat *four-tier multiple choice* pada penelitian menunjukkan bahwa pengecoh berfungsi dengan efektif jika pengecoh (distraktor) dipilih minimal 5% dari seluruh peserta tes (Mardapi, 2012). Dari keseluruhan butir soal pengecoh yang dipilih, lebih dari 5% dari seluruh peserta tes maka tergolong kategori efektif, sehingga pengecoh dari 13 butir soal berfungsi efektif namun terdapat 2 butir soal yang tidak berfungsi efektif, maka setiap butir soal yang pengecohnya berfungsi efektif langsung dipakai pada tahap selanjutnya.

Revisi produk yang dikembangkan

Berdasarkan hasil analisis butir soal pada tahap uji coba skala kecil, dilakukan proses seleksi terhadap soal-soal yang telah diujikan. Soal-soal yang tidak memenuhi kriteria kelayakan, yaitu validitas butir, reliabilitas instrumen, daya pembeda, dan keberfungsian pengecoh, dinyatakan tidak layak untuk digunakan dan oleh karena itu digugurkan. Pengguguran ini berarti bahwa butir-butir soal tersebut tidak akan disertakan dalam tahap pengumpulan data utama karena dinilai tidak mampu secara optimal untuk merepresentasikan kemampuan konseptual peserta didik.

Analisis butir soal dilakukan dengan menggunakan perangkat lunak *Iteman 4.3*, yang memberikan keluaran analisis kuantitatif terhadap sejumlah parameter penting. Kriteria validitas digunakan untuk menilai sejauh mana butir soal mengukur kompetensi yang dimaksud. Reliabilitas menunjukkan tingkat konsistensi instrumen secara keseluruhan. Daya pembeda menunjukkan kemampuan soal dalam membedakan peserta didik yang memiliki pemahaman tinggi dan rendah terhadap konsep tertentu. Sementara itu, keberfungsian pengecoh dianalisis untuk memastikan bahwa opsi jawaban yang salah mampu menarik responden yang tidak memahami konsep secara benar.

Dari seluruh butir soal yang dianalisis, diperoleh sebanyak 13 butir soal yang memenuhi seluruh kriteria yang telah ditetapkan. Ketiga belas soal tersebut dinilai valid, reliabel, memiliki tingkat kesukaran yang proporsional, daya pembeda yang baik, serta pengecoh yang berfungsi secara efektif. Oleh karena itu, ke-13 butir soal tersebut ditetapkan sebagai produk akhir yang dapat digunakan pada tahap selanjutnya, yaitu dalam pengumpulan data utama penelitian. Revisi terhadap

produk dilakukan dengan menyesuaikan struktur instrumen agar lebih representatif, memperbaiki redaksi soal jika diperlukan, serta memastikan bahwa format *four-tier multiple choice* tetap konsisten dan mudah dipahami oleh peserta didik.

Uji coba pemakaian

Uji coba penggunaan instrumen tes diagnostik *four-tier multiple choice* pada materi operasi aljabar dilakukan terhadap 35 peserta didik kelas VII A MTs Almaarif 01 Singosari. Uji coba ini dilaksanakan selama dua jam pelajaran pada hari Rabu, 21 Februari 2024. Sebagaimana telah dijelaskan sebelumnya, uji coba skala kecil ini bertujuan untuk menganalisis kualitas butir soal, yang meliputi aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta efektivitas pengecoh. Berdasarkan hasil analisis uji coba, diperoleh sebanyak 13 butir soal yang memenuhi kriteria kelayakan dan dapat digunakan dalam pengumpulan data utama. Sebelum pelaksanaan tes, peserta didik diberikan pengarahan mengenai prosedur pengerjaan instrumen sesuai dengan petunjuk pelaksanaan tes diagnostik.

Pengarahan awal dilakukan untuk memastikan bahwa peserta didik memahami karakteristik instrumen *four-tier*, yang memiliki struktur lebih kompleks dibandingkan dengan soal pilihan ganda konvensional. Setiap soal terdiri atas empat komponen, yaitu pernyataan pertanyaan, pilihan jawaban, alasan terhadap jawaban, dan pernyataan tingkat keyakinan peserta didik terhadap pilihannya. Dengan pemahaman yang utuh terhadap format soal tersebut, diharapkan peserta didik dapat memberikan respons yang mencerminkan kondisi pemahaman konseptual mereka secara objektif.

Selama proses uji coba berlangsung, peneliti turut melakukan observasi terhadap respons peserta didik dalam menjawab instrumen. Beberapa temuan penting dicatat, antara lain kendala peserta dalam memahami format soal pada tahap awal, pengelolaan waktu dalam menyelesaikan seluruh butir soal, serta kesesuaian tingkat kognitif soal dengan kemampuan peserta didik. Temuan ini menjadi bahan evaluasi dalam penyempurnaan instrumen sebelum diterapkan pada tahap pengumpulan data sesungguhnya. Uji coba ini juga berperan sebagai upaya awal untuk meningkatkan validitas internal instrumen serta memastikan bahwa instrumen benar-benar mampu mengungkap miskonsepsi dan pemahaman konseptual peserta didik secara akurat. Persentase jawaban peserta didik dalam mengerjakan instrumen tes diagnostik disajikan pada Tabel 11. berikut.

Tabel 11. Hasil Jawaban Peserta didik

No Soal	PK		TPK		M		E	
	Σ	%	Σ	%	Σ	%	Σ	%
1	28	80%	0	0%	6	17%	1	3%
2	23	66%	8	23%	2	6%	2	6%
3	24	69%	4	11%	7	20%	0	0%
4	15	43%	12	34%	7	20%	1	3%
5	13	37%	15	43%	6	17%	1	3%
6	11	31%	12	34%	8	23%	4	11%
7	10	29%	11	31%	11	31%	3	9%
8	11	31%	12	34%	9	26%	3	9%
9	13	37%	12	34%	7	20%	3	9%
10	14	40%	11	31%	10	29%	0	0%
11	10	29%	21	60%	4	11%	0	0%
12	8	23%	15	43%	9	26%	3	9%
13	10	29%	15	43%	7	20%	3	9%
Total		543%		423%		266%		69%
Rata-rata		42%		33%		20%		5%

Keterangan:

PK : Paham Konsep

M : Miskonsepsi

TPK : Tidak Paham Konsep

E : Error

Berdasarkan Tabel 11, hasil jawaban peserta didik kelas VII A MTs Almaarif 01 Singosari dengan 35 peserta didik menunjukkan bahwa rata-rata persentase paham konsep lebih besar dibandingkan dengan kategori lainnya. Peserta didik yang memahami konsep 42%, sementara tidak paham konsep 33%, miskonsepsi 20% dan error 5%.

Hasil ini mendukung temuan dari penelitian Putri & Subekti (2021), yang menemukan bahwa rata-rata persentase peserta didik yang memahami konsep cenderung lebih tinggi dibandingkan yang mengalami miskonsepsi. Namun pada penelitian ini, persentase peserta didik yang mengalami tidak memahami konsep lebih tinggi dibandingkan peserta didik yang mengalami miskonsepsi. Hasil tersebut mengindikasikan adanya kebutuhan mendesak untuk meningkatkan pemahaman konseptual peserta didik secara mendalam.

Instrumen diagnostik *four-tier multiple choice* yang dikembangkan terbukti mampu menjadi alat yang efektif untuk mengidentifikasi tingkat pemahaman konsep peserta didik. Hal ini sejalan dengan pendapat Kaltakci-Gurel, dkk. (2017), yang menyatakan bahwa instrumen diagnostik jenis ini dapat secara akurat mengidentifikasi tingkat pemahaman peserta didik dalam suatu konsep. Selain itu, hasil penelitian ini juga menunjukkan bahwa instrumen diagnostik dapat memberikan data yang relevan untuk digunakan sebagai dasar pengambilan keputusan dalam perbaikan pembelajaran.

Namun, tingginya persentase peserta didik yang tidak memahami konsep (33%) dan mengalami miskonsepsi (20%) menunjukkan bahwa masih banyak peserta didik yang menghadapi kesulitan dalam memahami materi aljabar. Berdasarkan hasil observasi dan wawancara, hal ini terjadi karena peserta didik cenderung menghafal materi tanpa benar-benar memahami konsepnya. Ketika dihadapkan pada soal yang melibatkan pengembangan konsep aljabar, peserta didik merasa bingung dan kesulitan menyelesaikannya. Selain itu, kurangnya kebiasaan membaca buku juga menjadi faktor yang berkontribusi terhadap kesalahan konsep yang sering terjadi.

Hasil ini menegaskan pentingnya penggunaan instrumen diagnostik yang dapat mengidentifikasi dengan jelas area miskonsepsi dan ketidakpahaman peserta didik. Guru dapat menggunakan data ini untuk merancang strategi pembelajaran yang lebih efektif, seperti memberikan penekanan pada pemahaman konsep melalui pendekatan yang melibatkan eksplorasi, diskusi, dan pemecahan masalah. Selain itu, peserta didik yang teridentifikasi mengalami miskonsepsi atau tidak memahami konsep memerlukan intervensi khusus, seperti bimbingan tambahan atau latihan yang dirancang untuk memperbaiki kesalahan konsep.

Setelah mengerjakan instrumen tes, peserta didik diberikan angket untuk menilai kepraktisan instrumen. Angket ini terdiri dari 16 butir pertanyaan yang mencakup aspek perencanaan dan efisiensi. Hasil angket menunjukkan rata-rata keseluruhan 80,75% dengan kategori "Praktis". Beberapa peserta didik memberikan komentar positif, seperti menyebut instrumen ini "bagus", "kreatif", dan "inovatif".

Hasil penelitian ini menunjukkan bahwa instrumen diagnostik *four-tier multiple choice* yang dikembangkan tidak hanya layak digunakan, tetapi juga praktis dan relevan untuk diterapkan di kelas VII SMP/MTs. Hal ini sesuai dengan penelitian Syahfitri dkk. (2019) yang menunjukkan bahwa instrumen diagnostik dapat dikategorikan praktis berdasarkan rata-rata persentase skor 78%. Dengan pengembangan lebih lanjut, instrumen ini dapat digunakan pada skala yang lebih luas untuk mendukung pembelajaran matematika yang lebih efektif dan berbasis pemahaman konsep.

KESIMPULAN

Pengembangan instrumen diagnostik berformat *four-tier multiple choice* pada materi operasi aljabar dilakukan menggunakan metode *Research and Development* (R & D) model Borg and Gall yang melalui delapan tahapan. Tahapan tersebut meliputi: (1) identifikasi potensi dan masalah, (2) pengumpulan data, (3) perancangan produk, (4) validasi desain, (5) revisi desain, (6) uji coba skala kecil, (7) revisi produk, serta (8) uji pemakaian. Pengembangan ini menghasilkan 13 soal yang memenuhi kriteria kelayakan dan kepraktisan. Berdasarkan hasil validasi, tingkat kelayakan dari validasi ahli instrumen,

materi, dan validator praktisi masing-masing mencapai 85%, 81%, dan 95%, yang tergolong sangat layak. Sementara itu, hasil uji kepraktisan dari angket respons peserta didik menunjukkan rata-rata 80,75%, dengan kategori praktis. Berdasarkan hasil tes diagnostik pada uji coba pemakaian, diketahui bahwa persentase peserta didik paham konsep 42%, tidak paham konsep 33%, miskonsepsi 20%, dan error 5%. Secara umum, hasil penelitian menunjukkan bahwa instrumen diagnostik *four-tier multiple choice* materi operasi aljabar layak dan praktis digunakan juga dapat mengetahui konsepsi peserta didik. Beberapa saran untuk penelitian selanjutnya sebaiknya tidak hanya mengembangkan instrumen materi operasi aljabar saja, tetapi mencakup materi lain agar lebih beragam dan tes ini juga dapat membantu guru dalam mengidentifikasi konsepsi peserta didik. Selain itu, implementasi tes ini dalam lingkup yang lebih luas diharapkan dapat lebih mendukung guru dalam menganalisis konsepsi peserta didik secara menyeluruh.

REFERENSI

- Abidin, Z., Mania, S., & Kusumayanti, A. (2019). Analisis Miskonsepsi Siswa Kelas VII SMP dengan Menggunakan Three Tier Test Pada Materi Aljabar. *Alauddin Journal of Mathematics Education*, 1(1), 19–25. <https://doi.org/10.24252/ajme.v1i1.10930>
- Agustin, U., Susilaningih, E., Nurhayati, S., & Wijayanti, N. (2022). Pengembangan Instrumen Tes Diagnostik Four-Tier Multiple Choice untuk Identifikasi Miskonsepsi Siswa pada Materi Kesetimbangan Kimia. *Chemistry in Education*, 11(1), 1–7. <https://doi.org/10.15294/chemined.v11i1.40935>
- Amin, N., Wiendartun, & Samsudin, A. (2016). Analisis Intrumen Tes Diagnostik Dynamic-Fluid Conceptual Change Inventory (DFCCI) Bentuk Four-Tier Test pada Beberapa SMA di Bandung Raya. *Prosiding SNIPS (Simposium Nasional Inovasi dan Pembelajaran Sains)*, 570–574. Bandung: FMIPA Institut Teknologi Bandung.
- Departemen Pendidikan Nasional. (2007). *Tes Diagnostik*. Jakarta: Kementrian Pendidikan Nasional.
- Elvia, R., Rohiat, S., & Ginting, S. M. (2021). Identifikasi Miskonsepsi Mahasiswa pada Pembelajaran Daring Matematika Kimia melalui Tes Diagnostik Three Tier Multiple choice. *Hidrogen: Jurnal Pendidikan Kimia*, 8(2), 84–96. <https://doi.org/10.33394/hjkk.v9i2.4422>
- Evendi, E. (2020). *Evaluasi Pembelajaran Matematika*. Mataram: Sanabil.
- Fatchiyah, F., & Utami, M. A. P. (2022). Pengembangan Platform Digital Berbasis Linktree dengan Geogebra “Barsama” pada Materi Bangun Ruang Sis Datar Kelas VIII MTs. *Galois: Jurnal Penelitian Pendidikan Matematika*, 1(2), 33–45. <https://doi.org/10.18860/gjppm.v1i2.2245>
- Guyer, R. L., & Thompson, N. A. (2013). *User's Manual for Itean: Classical Item and Test Analysis*. Minnesota: Assessment Systems Corporation.
- Harahap, I. P. P., & Novita, D. (2020). Validitas dan Reliabilitas Instrumen Tes Diagnostik Four-Tier Multiple Choice 94TMC) pad Konsep Laju Reaksi. *Unesa Journal of Chemical Education*, 9(2), 222–227.
- Huda, N., & Wahyuni, T. S. (2020). Penggunaan Aplikasi Item and Test Analysis (Iteman) pada Soal Try Out UN IPA SMA Tahun 2019. *Jurnal Pembelajaran Sains*, 4(1), 1–8. <https://doi.org/10.17977/um033v4i1p1-8>
- Ismail, I. I., Samsudin, A., Suhendi, E., & Kaniawati, I. (2015). Diagnostik Miskonsepsi melalui Listrik Dinamis Four Tier Test. *Prosiding Simposium Nasional Inovasi dan Pembelajaran Sains 2015*, 381–384. Bandung: FMIPA Institut Teknologi Bandung. Diambil dari https://www.researchgate.net/profile/Endi-Suhendi/publication/301523361_Diagnostik_Miskonsepsi_Melalui_Listrik_Dinamis_Fo

- ur_Tier_Test/links/5717652a08aeb56278c45cf7/Diagnostik-Miskonsepsi-Melalui-Listrik-Dinamis-Four-Tier-Test.pdf
- Kaltakci-Gurel, D., Eryilmaz, A., & McDermott, L. C. (2017). Development and Application of a Four-Tier Test to Assess Pre-Service Physics Teachers' Misconceptions about Geometrical Optics. *Research in Science and Technological Education*, 35(2), 238–260. <https://doi.org/10.1080/02635143.2017.1310094>
- Mardapi, D. (2012). *Pengukuran Penilaian dan Evaluasi Pendidikan*. Yogyakarta: Nuha Medika.
- Nabilah, L. Y., Ruslan, & Rusli. (2019). Pengembangan Instrumen Diagnostik Three Tier Test pada Materi Pecahan Kelas VII. *Issues in Mathematics Education*, 3(2), 184–193. <https://doi.org/10.35580/imed12421>
- Nai, N. (2019). *Pengembangan Instrumen Tes Diagnostik Pilihan Ganda Beralasan pada Materi Biologi Kelas VII MTs Madani Alauddin Kabupaten Gowa*. Universitas Islam Negeri Alauddin Makassar, Makassar.
- Permatasari, B. A. D., Setiawan, T. B., & Kristiana, A. I. (2015). Analisis Kesulitan Siswa dalam Menyelesaikan Soal Materi Aljabar Siswa Kelas VIII SMP Negeri 2 Bangil. *Kadikma*, 6(2). <https://doi.org/10.19184/kdma.v6i2.1990>
- Putranto, A., Langitasari, I., & Nursa'adah, E. (2020). Pengembangan Instrumen Three Tier Test pada Konsep Atom, Ion, dan Molekul. *Jurnal Zarah*, 8(1), 1–6. <https://doi.org/10.31629/zarah.v8i1.1349>
- Putri, R. E., & Subekti, H. (2021). Analisis Miskonsepsi Menggunakan Metode Four-Tier Certainty of Response Index: Studi Eksplorasi di SMP Negeri 60 Surabaya. *Pensa E-Jurnal: Pendidikan Sains*, 9(2), 220–226. Diambil dari <https://ejournal.unesa.ac.id/index.php/pensa/article/view/38279>
- Rusilowati, A. (2015). Pengembangan Tes Diagnostik sebagai Alat Evaluasi Kesulitan Belajar Fisika. *Prosiding Seminar Nasional Fisika dan Pendidikan Fisika*, 1–10. Solo: Universitas Sebelas Maret.
- Saputri, R., Kusumah, R. G. T., & Kasmantoni. (2022). Pengembangan Assessment Test untuk Mengukur Berpikir Tingkat Tinggi Siswa pada Materi Gelombang dan Bunyi. *ISEJ: Indonesian Science Education Journal*, 3(1), 1–8. <https://doi.org/10.62159/isej.v3i1.364>
- Sari, H. M., & Afriansyah, E. A. (2020). Analisis Miskonsepsi Siswa SMP pada Materi Operasi Hitung Bentuk Aljabar. *Mosharafa: Jurnal Pendidikan Matematika*, 9(3), 439–450. <https://doi.org/10.31980/mosharafa.v9i3.626>
- Sarip, M., Amintarti, S., & Utami, N. H. (2022). Validitas dan Keterbacaan Media Ajar E-Booklet untuk Siswa SMA/MA Materi Keanekaragaman hayati. *JUPEIS: Jurnal Pendidikan dan Ilmu Sosial*, 1(1), 43–59. <https://doi.org/10.57218/jupeis.Vol1.Iss1.30>
- Sriyanti, A., Mania, S., & A., N. H. (2019). Pengembangan Instrumen Tes Diagnostik Berbentuk Uraian untuk Mengidentifikasi Pemahaman Konsep Matematika Wajib Siswa MAN 1 Makassar. *de Fermat: Jurnal Pendidikan Matematika*, 2(1), 57–69. <https://doi.org/10.36277/defermat.v2i1.40>
- Syahfitri, J., Firman, H., Redjeki, S., & Sriyati, S. (2019). Pengujian Validitas dan Praktikalitas Tes Disposisi Berpikir Kritis dalam Biologi (TDBKB). *Jurnal Pendidikan Sains (JPS)*, 7(1), 30–35. <https://doi.org/10.26714/jps.7.1.2019.30-35>
- Syahlani, A., & Setyorini, D. (2021). Pengembangan Instrumen Hasil Belajar Matematika Siswa (Tes Pilihan Ganda). *Akrab Juara: Jurnal Ilmu-ilmu Sosial*, 6(3), 34–46. <https://doi.org/10.58487/akrabjuara.v6i3.1523>
- Wahid, Hartoyo, A., & Mirza, A. (2015). Miskonsepsi Siswa pada Materi Operasi pada Bentuk Aljabar Kelas VII SMP Haebat Islam. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa (JPPK)*, 4(1), 1–12. <https://doi.org/10.26418/jppk.v4i1.8519>

- Yakubi, M., Zulfadli, & Latifa, H. (2017). Menganalisis Tingkat Pemahaman Siswa pada Materi Ikatan Kimia Menggunakan Instrumen Penilaian Four-Tier Multiple Choice (Studi Kasus pada Siswa Kelas X SMA Negeri 4 Banda Aceh). *JIMPK (Jurnal Ilmiah Mahasiswa Pendidikan Kimia)*, 2(1), 19–26. <https://doi.org/10.82415/jimpk.v2i1.3400>
- Yuliyani, R. (2016). Pembelajaran Matematika Realistik pada Materi Operasi Aljabar di Kelas VII MTs Daarussa'adah Ciganjur Jakarta Selatan. *Formatif: Jurnal Ilmiah Pendidikan MIPA*, 6(3), 1. <https://doi.org/10.30998/formatif.v6i3.997>
- Zein, M., & Darto. (2017). *Evaluasi Pembelajaran Matematika*. Pekanbaru: Daulat Riau.