

Pengembangan Tes Diagnostik *Two-Tier*: Mengidentifikasi Pemahaman Konseptual dan Prosedural Materi Prasyarat Struktur Aljabar Grup

Firmansyah B.¹, Ali Umar²

¹Prodi Pendidikan Matematika, FTK, UIN Alauddin Makassar, Makassar-Indonesia

²Prodi Pendidikan Matematika, FTIK, UIN Mahmud Yunus Batusangkar, Batusangkar-Indonesia

¹firmansyahb@uin-alauddin.ac.id, ²aliumar@uinmybatusangkar.ac.id

Abstrak: Penelitian ini bertujuan mengembangkan tes diagnostik *two-tier* untuk mengidentifikasi pemahaman konseptual dan prosedural mahasiswa pada materi prasyarat Struktur Aljabar, khususnya teori grup. Penelitian ini menggunakan model pengembangan ADDIE yang meliputi Tahap Analisis (*Analyze*), tahap Perencanaan (*Design*), Tahap Pengembangan (*Develop*), Tahap Penerapan (*Implement*), dan Tahap Evaluasi (*Evaluate*). Instrumen yang dikembangkan terdiri dari lima butir soal uraian yang mencakup konsep himpunan, relasi, fungsi, dan operasi biner sebagai fondasi pembelajaran aljabar abstrak. Hasil validasi ahli diperoleh *Content Validity Coefficient* (CVC) sebesar 0,8, sedangkan analisis reliabilitas menghasilkan *Cronbach's Alpha* sebesar 0,651 yang tergolong memadai. Analisis butir soal menunjukkan distribusi tingkat kesukaran yang proporsional dan relatif berdistribusi normal dengan daya pembeda positif dan baik. Tes diagnostik ini efektif mengidentifikasi kesiapan konseptual mahasiswa, sejalan dengan teori *Zone of Proximal Development* (ZPD) Vygotsky dan teori *Meaningful Learning* Ausubel, karena membantu dosen memahami posisi awal kognitif mahasiswa untuk merancang strategi pembelajaran yang lebih adaptif, terarah, dan responsif terhadap kebutuhan belajar mahasiswa.

Kata Kunci: tes diagnostik, *two-tier*, prasyarat teori grup, validitas, reliabilitas.

Abstract: This study aims to develop a two-tier diagnostic test to identify students' conceptual and procedural understanding of prerequisite materials in Abstract Algebra, particularly group theory. The study employed the ADDIE development model, which consists of the Analysis, Design, Development, Implementation, and Evaluation stages. The instrument developed comprises five open-ended two-tier items covering the concepts of sets, relations, functions, and binary operations as foundational components of abstract algebra learning. Expert validation yielded a Content Validity Coefficient (CVC) of 0.8, while reliability analysis produced a Cronbach's Alpha value of 0.651, indicating acceptable reliability. Item analysis showed a proportional distribution of difficulty levels and a relatively normal distribution of responses, with positive and satisfactory discrimination indices. This diagnostic test effectively identifies students' conceptual readiness, aligning with Vygotsky's Zone of Proximal Development (ZPD) and Ausubel's Meaningful Learning theory, as it helps instructors understand students' initial cognitive positions to design more adaptive, targeted, and responsive instructional strategies.

Keywords: diagnostic test, two-tier, prerequisite for group theory, validity, reliability.

Pendahuluan

Pembelajaran hakikatnya merupakan kegiatan yang dirancang untuk mencapai tujuan pembelajaran. Pembelajaran matematika di perguruan tinggi menuntut mahasiswa mampu berpikir abstrak dan memahami konsep secara mendalam. Salah satu mata kuliah pada program studi pendidikan matematika yang berperan dalam membentuk kemampuan berpikir formal adalah struktur aljabar. Struktur aljabar sendiri dalam kajian matematika didefinisikan sebagai pasangan suatu himpunan tak kosong dengan satu atau lebih operasi biner yang

didefinisikan pada himpunan tersebut. Grup merupakan salah satu struktur aljabar yang fundamental dan menjadi dasar pengembangan konsep struktur aljabar seperti ring, field, dan ruang vektor. Berbagai penelitian menunjukkan bahwa tidak sedikit mahasiswa yang mengalami kesulitan dalam memahami teori grup karena lemahnya penguasaan materi prasyarat seperti himpunan, relasi, fungsi, dan operasi biner (Agustyaningrum et al., 2021; Hanifah & Abadi, 2018).

Kesulitan dalam memahami konsep-konsep dasar pada teori grup disebabkan ketidakmampuan mahasiswa menghubungkan ide-ide formal dengan representasi konseptual konkret (Dubinsky et al., 1994). Proses berpikir tersebut memerlukan kemampuan kognitif yang berkembang melalui tahapan tindakan, proses, objek, dan skema menurut teori APOS (Action, Process, Object, Schema) (Dubinsky & McDonald, 2001). Tanpa pemahaman yang baik dan kuat mahasiswa cenderung melakukan prosedur simbolik melalui hafalan tanpa pemaknaan konseptual secara mendalam (Dawkins, 2015; Tall, 2004).

Salah satu penyebab kesulitan mahasiswa adalah ketidaksiapan konseptual dan prosedural pada materi prasyarat. Hal tersebut berkaitan dengan kemampuan memahami definisi konsep dan sifat-sifatnya serta keterampilan dalam menerapkan prinsip dasar secara prosedural (Widyawati, 2025). Dengan demikian jika materi sebelumnya belum dikuasai sepenuhnya oleh mahasiswa, maka akan berpengaruh pada penguasaan konsep pada materi teori grup. Oleh karena itu diperlukan suatu instrumen asesmen diagnostik untuk mengidentifikasi tingkat pemahaman mahasiswa sebelum masuk pada materi struktur aljabar.

Tes diagnostik bisa menjadi jembatan antara kemampuan aktual mahasiswa dan tuntutan pembelajaran lanjutan di perguruan tinggi (Manaster, 1991; Sheridan, 2013). Dalam pendidikan matematika, tes diagnostik biasanya digunakan untuk mendeteksi miskonsepsi atau kesalahan pemahaman dan juga untuk memetakan kesiapan belajar peserta didik terhadap materi prasyarat (Elvia et al., 2021; Mutmainna et al., 2018). Salah satu bentuk tes diagnostik yang dinilai dapat mengungkap pemahaman yang lebih mendalam secara efektif adalah dengan format *two-tier diagnostic test*. Tes diagnostik two-tier terdiri dari dua bagian. Tier pertama berupa pertanyaan konseptual dan tier kedua berupa alasan yang mendasari jawaban. Dengan demikian memudahkan dosen dalam mengidentifikasi mana jawaban benar karena pemahaman dan jawaban yang didasari dari hafalan atau bahkan miskonsepsi (Halim & Nanda, 2018). Pada penelitian terhadap pemahaman konseptual mahasiswa pada teori grup, menunjukkan hasil bahwa informasi kesulitan belajar dan miskonsepsi mahasiswa yang diperoleh dari tes diagnostik berbasis two-tier lebih akurat dibanding bentuk tes konvensional (Veith et al., 2022). Selanjutnya pengembangan tes yang menggunakan analisis butir dinilai mampu meningkatkan validitas dan reliabilitas instrumen (Akveld & Kinnear, 2024).

Pengembangan tes diagnostik pada materi prasyarat ini merujuk pada teori pembelajaran bermakna (*meaningful learning theory*) dari David Ausubel. Teori tersebut menekankan bahwa proses belajar akan lebih efektif jika materi baru terhubung dengan pengetahuan yang sudah dimiliki peserta didik (da Silva, 2020; Sexton, 2025). Pada teori tersebut pembelajaran bermakna dapat terjadi apabila struktur kognitif awal mahasiswa dikanali dan dijadikan titik awal pembelajaran. Sejalan dengan itu, teori *Zone of Proximal Development* (ZPD) Vygotsky menjelaskan bahwa pembelajaran efektif ketika pengajaran menargetkan jarak antara apa yang telah dikuasai dengan kemampuan yang akan mereka capai melalui bantuan (Bruner, 1984; Fani & Ghaemi, 2011; Silalahi, 2019). Sehingga asesmen diagnostik tidak hanya sebagai alat evaluasi tetapi juga sebagai sarana menemukan zona perkembangan potensial mahasiswa sebelum memulai pembelajaran inti.

Berdasarkan uraian tersebut, penelitian ini bertujuan mengembangkan tes diagnostik two-tier yang mampu mengidentifikasi pemahaman konseptual dan prosedural mahasiswa pada materi prasyarat struktur aljabar terkhusus grup. Instrumen dirancang agar bersifat praktis dan dapat dilaksanakan pada awal perkuliahan dengan memberikan hasil diagnostik yang informatif dalam waktu yang singkat. Hasilnya diharapkan memberikan kontribusi teoritis terhadap kajian asesmen diagnostik, serta praktis sebagai alat evaluasi bagi dosen untuk memetakan kesiapan mahasiswa dan dalam merancang strategi pembelajaran yang efektif.

Metode Penelitian

Penelitian ini merupakan penelitian pengembangan (*Research and Development*) yang bertujuan menghasilkan tes diagnostik dua tingkat (*two-tier diagnostic test*) untuk mengukur pemahaman konseptual dan prosedural mahasiswa pada materi prasyarat sebelum mempelajari Struktur Aljabar, khususnya konsep Grup. Model pengembangan yang digunakan adalah model ADDIE, yang meliputi lima tahap, yaitu *Analyze*, *Design*, *Develop*, *Implement*, dan *Evaluate* (Branch & Varank, 2009). Model ini dipilih karena memberikan alur sistematis dan fleksibel yang sesuai untuk pengembangan instrumen asesmen pendidikan (Molenda, 2015). Tahap Analisis (*Analysis*) pada penelitian ini mengidentifikasi kebutuhan, konteks, dan karakteristik mahasiswa yang menjadi dasar pengembangan tes diagnostik *two-tier* pada materi prasyarat struktur aljabar. Analisis dilakukan melalui kajian dokumen dan observasi awal mahasiswa. Tahap Perancangan (*Design*) dalam pengembangan tes diagnostik two-tier ini dirancang instrumen yang sesuai dengan tujuan identifikasi pemahaman konseptual dan prosedural mahasiswa pada materi prasyarat struktur aljabar. Tahap Pengembangan (*Development*) berfokus pada proses validasi dan penyempurnaan instrumen sebelum dilakukan uji coba terbatas. Tahap Penerapan (*Implementation*) bertujuan untuk menguji penerapan instrumen tes diagnostik yang telah divalidasi pada kelompok mahasiswa yang menjadi sasaran penggunaannya. Tahap Evaluasi (*Evaluation*) bertujuan untuk menilai kelayakan instrumen tes diagnostik setelah melalui proses validasi dan uji coba.

Subjek penelitian adalah mahasiswa Program Studi Pendidikan Matematika semester III pada Fakultas Tarbiyah dan Keguruan UIN Alauddin Makassar, yang akan menempuh mata kuliah Struktur Aljabar I. Pemilihan subjek dilakukan dengan *purposive sampling*, mempertimbangkan bahwa mahasiswa pada tahap ini sedang berada dalam transisi dari pemahaman aljabar elementer menuju abstraksi formal. Konteks penelitian dilakukan sebelum perkuliahan dimulai untuk memetakan tingkat kesiapan mahasiswa terhadap konsep-konsep prasyarat struktur aljabar grup.

Instrumen yang digunakan dalam proses pengembangan berupa lembar validasi ahli dan instrumen utama yaitu tes diagnostik yang dikembangkan. Lembar validasi ahli untuk memperoleh data kuantitatif dan komentar kualitatif. Penilaian ahli diminta untuk mengetahui validitas konten (Beck, 2020). Validitas konten instrumen dilakukan untuk memastikan bahwa instrumen sesuai dengan tujuan dan materi yang diukur sebagai salah satu indikator kualitas instrumen (Firmansyah & Rais, 2023). Hasil tes mahasiswa untuk menganalisis performa butir soal pada instrumen tes diagnostik yang dikembangkan.

Data yang diperoleh dari validator selanjutnya dianalisis validitas isi dengan rumus *Content Validity Coefficient* untuk menghitung koefisien kesepakatan isi antar dua validator (Gregory, 2013).

$$CVC = \frac{D}{A + B + C + D}$$

A adalah jumlah butir yang disepakati tidak valid oleh kedua validator (*invalid agreement*). B dan C merupakan jumlah butir yang divalidkan oleh salah satu validator saja. D adalah jumlah butir yang disepakati valid oleh kedua validator (*valid agreement*). Dengan demikian indeks yang diperoleh berada pada interval 0 sampai 1 yang menunjukkan perbandingan jumlah item dari dua pakar sebagai validator dengan relevansi kuat terhadap keseluruhan kategori item (Retnawati, 2016). Validitas konten mengacu pada koefisien dengan nilai lebih besar atau sama dengan 0,8 (Subali, 2018).

Adapun data hasil tes mahasiswa yang diperoleh dari uji coba pada prototipe tes diagnostik yang dikembangkan selanjutnya dianalisis mengikuti pendekatan analisis butir soal (tingkat kesulitan, daya pembeda, indeks reliabilitas) (Cañeda et al., 2024).

Hasil dan Pembahasan

Hasil Penelitian

1. Analysis

Hasil analisis menunjukkan bahwa sebagian besar mahasiswa mengalami kesulitan memahami konsep grup karena lemahnya penguasaan materi prasyarat seperti himpunan, relasi, dan fungsi bijektif. Materi tersebut kemudian ditetapkan sebagai fokus dalam penyusunan lima butir soal diagnostik yang bersifat konseptual dan prosedural. Selain itu, hasil analisis kebutuhan menunjukkan bahwa instrumen perlu praktis dan dapat dikerjakan dalam waktu 50 menit (1 SKS) dari 150 menit yang tersedia pada pertemuan pertama, sehingga sesuai dengan tujuan penggunaannya sebagai asesmen awal perkuliahan Struktur Aljabar I.

2. Design

Berdasarkan hasil analisis kebutuhan, dikembangkan lima butir soal uraian yang mewakili konsep prasyarat utama, yaitu himpunan, relasi, fungsi, dan operasi biner. Adapun indikator kognitif yang dikembangkan berdasarkan konsep konsep tersebut antara lain: 1) Mahasiswa mampu membedakan kumpulan yang merupakan himpunan dan bukan himpunan; 2) Mahasiswa mampu menyatakan himpunan dengan notasi deskriptif; 3) Mahasiswa mampu menentukan himpunan bagian dan himpunan kuasa dari suatu himpunan; 4) Mahasiswa menjelaskan hubungan antara banyaknya elemen dengan banyaknya himpunan bagian; 5) Mahasiswa mampu mengidentifikasi pasangan elemen yang memenuhi relasi tertentu; 6) Mahasiswa mampu menilai apakah suatu relasi bersifat refleksif, simetris, dan transitif; 7) Mahasiswa menyimpulkan apakah relasi tersebut ekuivalen; 8) Mahasiswa membuktikan suatu fungsi bersifat satu-satu dan onto; 9) Mahasiswa menentukan fungsi invers dan memeriksa hasil komposisinya; 10) Mahasiswa memeriksa syarat dua fungsi dapat dikomposisikan berdasarkan domain dan kodomain; 11) Mahasiswa menentukan hasil komposisi fungsi jika terdefinisi; 12) Mahasiswa membandingkan dua hasil komposisi dan menentukan sifat komutatifnya.

Berdasarkan indikator kognitif tersebut dikembangkan lima indikator soal. Indikator soal yang dikembangkan sebagai berikut ini:

1. Setelah diberikan beberapa contoh kumpulan, mahasiswa mampu menentukan kumpulan mana yang merupakan himpunan dan menuliskannya dalam notasi pembentuk himpunan secara tepat.
2. Setelah diberikan suatu himpunan terhingga, mahasiswa mampu menentukan seluruh himpunan bagiannya dan menjelaskan hubungan antara jumlah elemen dan banyaknya himpunan bagian secara benar.

3. Setelah diberikan relasi tertentu pada himpunan bilangan terhingga, mahasiswa mampu mengidentifikasi pasangan yang memenuhi relasi, memeriksa sifat refleksif, simetris, dan transitif, serta menyimpulkan apakah relasi tersebut ekuivalen secara logis.
4. Setelah diberikan fungsi linear pada himpunan bilangan real, mahasiswa mampu membuktikan fungsi tersebut bersifat bijektif dan menentukan inversnya dengan perhitungan yang benar.
5. Setelah diberikan dua fungsi f dan g dengan domain dan kodomain tertentu, mahasiswa memeriksa kesesuaian menentukan apakah fungsi dapat dikomposisikan, kemudian menentukan hasil komposisi $g \circ f(x)$ dan $f \circ g(x)$, membandingkannya, serta menjelaskan maknanya terhadap sifat komutatif secara tepat.

Keseluruhan butir soal dirancang dalam bentuk *two-tier*, di mana *tier* pertama menilai pemahaman konseptual mahasiswa terhadap definisi dan sifat formal suatu konsep, sedangkan *tier* kedua menilai alasan dan juga kemampuan prosedural dalam menerapkan konsep tersebut pada situasi matematis tertentu. Misalnya, pada soal himpunan, mahasiswa tidak hanya diminta menentukan kumpulan yang merupakan himpunan, tetapi juga menuliskannya dalam notasi pembentuk himpunan yang benar. Pada soal terkait fungsi bijektif, mahasiswa diminta membuktikan sifat satu-satu dan onto suatu fungsi sekaligus menentukan fungsi inversnya jika memenuhi syarat.

Desain butir soal juga mempertimbangkan tingkat kognitif berdasarkan revisi Taksonomi Bloom, mulai dari C2 (memahami) hingga C5 (mengevaluasi). Setiap soal memiliki indikator spesifik yang mengaitkan konsep prasyarat dengan penerapannya dalam teori grup, misalnya: konsep himpunan mendukung pemahaman struktur grup sebagai himpunan tak kosong, relasi ekuivalen sebagai dasar pembentukan kelas ekuivalen dan *quotient group*, serta fungsi bijektif sebagai analogi untuk isomorfisme grup.

Selain kisi-kisi dan soal, pada tahap desain ini juga disusun rubrik penskoran dengan skala 0 sampai 4 untuk tiap soal, yang mengukur tingkat kelengkapan, ketepatan konsep, dan argumentasi matematis mahasiswa. Skor maksimal keseluruhan adalah 20.

3. Development

Kegiatan utama pada tahap ini meliputi validasi ahli terhadap tiga komponen utama instrumen, yaitu kisi-kisi soal, lembar tes diagnostik, dan rubrik penskoran. Validasi ahli dilakukan oleh dua ahli yang memiliki kompetensi di bidang pendidikan matematika dan berasal dari institusi berbeda dengan peneliti. Kedua validator dipilih secara purposif berdasarkan pengalaman dalam pengembangan instrumen asesmen maupun dalam mengajarkan aljabar abstrak di perguruan tinggi.

Instrumen validasi dianalisis menggunakan *Content Validity Coefficient* (CVC). Masing-masing validator memberikan penilaian terhadap relevansi terhadap indikator, konstruksi, dan kejelasan bahasa setiap butir soal dengan menggunakan skala empat. Berdasarkan hasil analisis kuantitatif terhadap data penilaian tersebut diperoleh nilai CVC = 0,8. Dengan demikian, butir soal, kisi-kisi, dan rubrik penilaian dinilai sesuai dengan konstruk yang diukur dengan tujuan sebagai asesmen diagnostik materi prasyarat untuk mata kuliah struktur aljabar I.

Namun validator juga memberikan masukan secara kualitatif untuk penyempurnaan instrumen. Saran validator mencakup aspek kebahasaan dan instruksi pada soal. Beberapa masukan dan revisi yang dilakukan disajikan pada Tabel 1 berikut:

Tabel 1. Saran validator dan perbaikan

Saran validator	Perbaikan yang dilakukan
-----------------	--------------------------

Pada soal nomor 1, tambah satu contoh lagi seolah-olah himpunan. “kumpulan mahasiswa dengan kategori berat badan oversize”. Kemungkinan mahasiswa berpikir himpunan karena berat badan diukur dengan angka.	Dilakukan penambahan kumpulan dari 3 kumpulan menjadi 4 kumpulan. Selanjutnya mahasiswa menentukan kumpulan mana yang merupakan himpunan dan menuliskannya dalam notasi pembentuk himpunan yang tepat.
Pada indikator kognitif pada soal nomor 2 Sebaiknya menggunakan KKO bloom secara eksplisit. “menjelaskan” tidak ada di KKO C3. Tambahkan menjadi mahasiswa “mampu” tanpa mampu menjadi kalimat aktifitas	Dilakukan penyesuaian kalimat pada indikator agar sesuai dengan level kognitif C3-menerapkan. Sehingga indikator soal nomor 2 menjadi: “Setelah diberikan suatu himpunan terhitung, mahasiswa dapat menerapkan konsep kardinalitas untuk menentukan seluruh himpunan bagian serta menunjukkan hubungan kuantitatif antara jumlah elemen dan banyaknya himpunan bagian secara benar”

Dengan selesainya tahap ini, selanjutnya instrumen digunakan pada tahap *Implementation* yang melibatkan uji coba terbatas pada kelompok mahasiswa untuk menilai reliabilitas serta analisis butir soal yang mencakup tingkat kesulitan dan daya pembeda.

4. Implementation

Ujicoba terbatas dilakukan pada 24 mahasiswa program studi Pendidikan Matematika semester III yang akan menempuh mata kuliah Struktur Aljabar I di Fakultas Tarbiyah dan Keguruan UIN Alauddin Makassar. Pemilihan subjek ini mempertimbangkan bahwa mereka mereka pada fase transisi dari aljabar elementer menuju aljabar abstrak.

Pelaksanaan ujicoba dilakukan secara tatap muka dengan pemberian tes secara langsung pada awal perkuliahan. Setiap mahasiswa mengerjakan lima soal tes diagnostik berbentuk uraian dalam durasi 50 menit. Selama pengerjaan, peneliti memastikan mahasiswa memahami petunjuk dan tidak menggunakan sumber selain apa yang mereka ketahui. Uji coba dilakukan terhadap 24 mahasiswa pendidikan matematika semester III. Setiap mahasiswa menjawab seluruh 5 butir tes, dengan skor setiap soal antara 0 sampai 4. Sehingga maksimum skor adalah 20. Adapun statistik hasil tes 24 orang mahasiswa pada tahap uji coba sebagai berikut.

Tabel 2. Statistik Deskriptif Skor Tes Diagnostik

Statistik	Nilai
Mean	9,33
Median	8
Mode	8
Standard Deviation	4,28
Sample Variance	18,32
Minimum	3
Maximum	19

Analisis reliabilitas menggunakan *Cronbach's Alpha*. Hasilnya disajikan pada tabel 3 berikut.

Tabel 3. Statistik reliabilitas dengan Cronbach's Alpha

Cronbach's Alpha	N of Items
,651	5

Diperoleh nilai $\alpha = 0,651$ yang berada pada kategori sedang. Dimana Rentang Cronbach's Alpha dianggap memiliki konsistensi internal "tingkat sedang" (0,60-0,69 dan 0,70-0,79) dan "tingkat tinggi" (0,80-0,89 dan 0,90-0,99), adapun tingkat rendah ($< 0,60$) dianggap tidak relevan (Karakaya & Alparslan, 2022). Hal tersebut senada dengan Vaske et al. (2017) yang mengkategorikan nilai Cronbach's Alpha pada interval 0,65 – 0,80 pada kategori "memadai". Nilai 0,651 yang diperoleh dimungkinkan karena jumlah item relatif sedikit hanya sebanyak lima soal. Sedangkan besarnya alfa bergantung pada jumlah item. Dimana jika korelasi rata-rata antar item tetap dan jumlah item meningkat, skala reliabilitas akan meningkat (Vaske et al., 2017). Namun demikian alasan penggunaan 5 soal sudah disebutkan pada tahap analisis.

Selanjutnya analisis butir soal mencakup analisis tingkat kesulitan dan daya pembeda. Hasil analisis tingkat kesulitan menunjukkan proporsi rata-rata skor pada setiap butir terhadap skor maksimum (Crocker & Algina, 1986). Kategori tingkat kesukaran berdasarkan kategori dari Surapranata (2019), dimana untuk koefisien $p > 0,7$ dikategorikan "Mudah", koefisien $0,3 \leq p \leq 0,7$ dikategorikan "sedang", dan koefisien $p < 0,3$ dikategorikan "sukar". Hasil analisis tingkat kesukaran butir soal disajikan pada Tabel 2 berikut:

Tabel 4. Tingkat kesukaran butir soal

Nomor Soal	Rata-rata Skor	p	Kategori
1	2,83	0,71	Mudah
2	2,67	0,67	Sedang
3	1,83	0,46	Sedang
4	1,21	0,30	Sedang
5	0,79	0,20	Sukar

Dari distribusi tingkat kesulitan pada Tabel 4 terlihat soal yang disusun mencerminkan tingkat kesulitan dengan distribusi normal. Sehingga dinilai sesuai untuk dijadikan tes diagnostik.

Adapun daya pembeda menunjukkan kemampuan butir membedakan mahasiswa berkemampuan tinggi dan rendah. Daya pembeda menggunakan korelasi Product Moment (r) antara skor butir dan skor total (tanpa butir tersebut) (Crocker & Algina, 1986). Kategori daya pembeda berdasarkan kategori dari Surapranata (2019), dimana untuk daya pembeda $> 0,3$ diputuskan "Diterima", daya pembeda 0,1 sampai 0,3 diputuskan "Direvisi", dan daya pembeda $< 0,1$ diputuskan "Ditolak". Hasil analisis daya pembeda butir soal disajikan pada Tabel 5 berikut:

Tabel 5. Daya pembeda butir soal

Nomor Soal	r	Kategori Daya Pembeda	Keputusan
1	0,314	Baik	Diterima

2	0,277	Cukup	Direvisi
3	0,395	Baik	Diterima
4	0,422	Sangat Baik	Diterima
5	0,683	Sangat Baik	Diterima

dari data pada Tabel 5 tersebut diketahui semua item soal dapat diterima, namun untuk nomor 2 diterima dengan terlebih dahulu dilakukan revisi.

5. *Evaluation*

Hasil validasi menunjukkan CVC sebesar 0,8 dengan revisi minor terkait kelengkapan soal dan kesesuaian indikator terhadap level kognitif, sehingga instrumen dinilai layak diujicobakan. Selanjutnya, evaluasi dilakukan berdasarkan hasil uji coba pada 24 mahasiswa. Hasil analisis menunjukkan nilai reliabilitas diperoleh Cronbach's Alpha sebesar 0,651. Hasil tersebut masih dianggap relevan dan memadai ($> 0,60$). Nilai tersebut mendekati batas bawah reliabel dimungkinkan karena jumlah item relatif sedikit sedangkan besarnya alfa bergantung pada jumlah item. Berdasarkan analisis butir soal yang dilakukan, diperoleh tingkat kesukaran telah memenuhi kriteria dan menunjukkan distribusi normal. Terdapat masing-masing satu soal berada pada kategori sukar dan mudah, sisanya berada pada kategori sedang. Pada analisis daya pembeda soal diperoleh bahwa seluruh item dapat diterima namun terdapat satu soal yang disarankan dilakukan revisi dan perbaikan. Berdasarkan hasil ujicoba tersebut maka secara umum soal tes diagnostik materi prasyarat yang dikembangkan dapat digunakan dengan dilakukan penyempurnaan pada item soal nomor 2. Revisi dilakukan dengan menambahkan konteks pada himpunan yang digunakan pada soal dan meminta pemaknaan pada setiap elemen anggota himpunan kuasa.

Pembahasan

Hasil pengembangan berupa tes diagnostik *two-tier* dinilai efektif untuk mengidentifikasi posisi awal mahasiswa dalam memahami konsep prasyarat sebelum mempelajari materi struktur aljabar terkait grup. Beberapa penelitian sebelumnya umumnya menggunakan format *two-tier* pilihan ganda (Mutmainna et al., 2018; Syaifuddin et al., 2022). Sedangkan instrumen yang dikembangkan dalam penelitian ini menggunakan bentuk uraian *two-tier*, sehingga tidak hanya mengungkap alasan konseptual mahasiswa, tetapi juga menilai kemampuan prosedural dan argumentasi matematis secara eksplisit. Kemampuan tes tersebut sejalan dengan teori *Zone of Proximal Development* oleh Vygotsky. Menurut teori tersebut pembelajaran akan lebih efektif apabila intervensi pengajaran diarahkan pada kemampuan aktual mahasiswa melalui bimbingan dan scaffolding yang tepat (Bruner, 1984; Fani & Ghaemi, 2011; Silalahi, 2019). Melalui tes ini, dosen dapat mengenali mahasiswa yang masih berada pada tahap kemampuan aktual dan merancang strategi pembelajaran yang sesuai. Selanjutnya hasil pengembangan yang diperoleh juga mendukung prinsip *meaningful learning* oleh Ausubel, yang menekankan bahwa pembelajaran yang efektif apabila konsep baru dapat dihubungkan dengan struktur kognitif yang sudah dimiliki mahasiswa (da Silva, 2020; Sexton, 2025). Dalam hal ini tes diagnostik yang dikembangkan berperan sebagai alat untuk mengungkap pengetahuan awal yang menjadi dasar terbentuknya pemahaman bermakna. Dengan mengetahui pengetahuan awal mahasiswa, dosen dapat mengorganisasikan pembelajaran struktur aljabar terkhusus teori grup secara hierarki dimulai dari konsep yang telah dikenal menuju abstraksi yang lebih kompleks.

Simpulan

Penelitian ini menghasilkan instrumen berupa tes diagnostik dengan format *two-tier* yang valid dan reliabel untuk mengidentifikasi pemahaman mahasiswa pada materi prasyarat struktur aljabar, khususnya sebelum mempelajari konsep grup. Hasil validasi oleh ahli menunjukkan nilai *Content Validity Coefficient* (CVC) sebesar 0,8, yang berarti instrumen memiliki validitas isi yang tinggi. Analisis reliabilitas menghasilkan nilai *Cronbach's Alpha* sebesar 0,651, menunjukkan konsistensi internal yang memadai dengan jumlah butir yang terbatas. Analisis butir soal menunjukkan distribusi tingkat kesukaran yang proporsional dan relatif berdistribusi normal dengan satu butir mudah, tiga butir sedang, dan satu butir sukar. Seluruh soal memiliki daya pembeda positif dan baik, dengan satu butir perlu revisi minor.

Secara praktis, tes diagnostik ini dapat digunakan dosen sebagai alat asesmen awal untuk mengidentifikasi kesiapan mahasiswa dan dalam merancang strategi pembelajaran yang sesuai dengan tingkat pemahaman mahasiswa. Pengembangan lanjutan dapat dilakukan dengan mengembangkan tes formatif dan sumatif pada mata kuliah Struktur Aljabar I terkait teori grup. Sehingga bisa menjadi acuan dalam memetakan kesiapan konseptual mahasiswa untuk memahami materi aljabar abstrak selanjutnya.

Referensi

- Agustyaningrum, N., Sari, R. N., Abadi, A. M., & Mahmudi, A. (2021). Dominant Factors That Cause Students' Difficulties in Learning Abstract Algebra: A Case Study at a University in Indonesia. *International Journal of Instruction*, 14(1), 847–866. <https://doi.org/10.29333/iji.2021.14151a>.
- Akveld, M., & Kinnear, G. (2024). Improving mathematics diagnostic tests using item analysis. *International Journal of Mathematical Education in Science and Technology*, 55(10), 2478–2505. <https://doi.org/10.1080/0020739X.2023.2167132>.
- Beck, K. (2020). Ensuring content validity of psychological and educational tests—the role of experts. *Frontline Learning Research*, 8(6), 1–37. <https://doi.org/10.14786/flr.v8i6.517>.
- Branch, R. M., & Varank, İ. (2009). *Instructional design: The ADDIE approach* (Vol. 722). Springer.
- Bruner, J. (1984). Vygotsky's zone of proximal development: The hidden agenda. *New Directions for Child Development*.
- Cañeda, M. E., Amar, R. P., & Lucin, E. L. (2024). Development of test questionnaire on selected topics in calculus 1 (final term). *International Journal of Research and Scientific Innovation*, 9(8), 244–255. <https://doi.org/10.51244/IJRSI.2024.1108020>.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. ERIC.
- da Silva, J. B. (2020). David Ausubel's Theory of Meaningful Learning: an analysis of the necessary conditions. *Research, Society and Development*, 9(4), 3. <http://dx.doi.org/10.33448/rsd-v9i4.2803>.
- Dawkins, P. C. (2015). In *Pursuit of Coherent and Formalizable Understanding: Reflections on David Tall's Three Worlds Framework: How Humans Learn to Think Mathematically: Exploring the Three Worlds of Mathematics*. Tall, D.(2013). New York, NY, USA: Cambridge University Press. Springer.
- Dubinsky, E., Dautermann, J., Leron, U., & Zazkis, R. (1994). On learning fundamental concepts of group theory. *Educational Studies in Mathematics*, 27(3), 267–305.
- Dubinsky, E., & McDonald, M. A. (2001). APOS: A constructivist theory of learning in undergraduate mathematics education research. In *The teaching and learning of*

- mathematics at university level: An ICMI study* (pp. 275–282). Springer.
- Elvia, R., Rohiat, S., & Ginting, S. M. (2021). Identifikasi miskonsepsi mahasiswa pada pembelajaran daring matematika kimia melalui tes diagnostik three tier multiple choice. *Hydrogen: Jurnal Kependidikan Kimia*, 9(2), 84–96.
- Fani, T., & Ghaemi, F. (2011). Implications of Vygotsky's zone of proximal development (ZPD) in teacher education: ZPTD and self-scaffolding. *Procedia-Social and Behavioral Sciences*, 29, 1549–1554.
- Firmansyah, B., & Rais, M. R. (2023). Development of mathematical literacy-based teaching material on mathematics. *MaPan: Jurnal Matematika Dan Pembelajaran*, 11(1), 21–37. <https://doi.org/10.24252/mapan.2023v11n1a2>.
- Gregory, R. J. (2013). *Psychological Testing: History, Principles, and Applications*. Pearson Education. <https://books.google.co.id/books?id=fY2DnAEACAAJ>
- Halim, A., & Nanda, N. (2018). Development of two-tier diagnostic test based on e-learning. *Journal of Physics: Conference Series*, 1120(1), 12030. doi:10.1088/1742-6596/1120/1/012030.
- Hanifah, H., & Abadi, A. P. (2018). Analisis pemahaman konsep matematika mahasiswa dalam menyelesaikan soal teori grup. *Journal of Medives: Journal of Mathematics Education IKIP Veteran Semarang*, 2(2), 235–244. <https://doi.org/10.31331/medives.v2i2.626>.
- Karakaya, S. P. Y., & Alparslan, Z. N. (2022). Sample size in reliability studies: A practical guide based on Cronbach's alpha. *Psychiatry and Behavioral Sciences*, 12(3), 150. doi: 10.5455/PBS.20220127074618.
- Manaster, A. (1991). Diagnostic testing: one link between university and high school mathematics. *Mathematicians and Education Reform, 1989-1990*, 2, 25.
- Molenda, M. (2015). In search of the elusive ADDIE model. *Performance Improvement*, 54(2).
- Mutmainna, D., Mania, S., & Sriyanti, A. (2018). Pengembangan instrumen tes diagnostik pilihan ganda dua tingkat untuk mengidentifikasi pemahaman konsep matematika. *MaPan: Jurnal Matematika Dan Pembelajaran*, 6(1), 56–69. : <https://doi.org/10.24252/mapan.2018v6n1a6>.
- Retnawati, H. (2016). Proving content validity of self-regulated learning scale (The comparison of Aiken index and expanded Gregory index). *REiD (Research and Evaluation in Education)*, 2(2), 155–164. <https://doi.org/10.21831/reid.v2i2.11029>.
- Sexton, S. S. (2025). Meaningful Learning—David P. Ausubel. In *Science education in theory and practice: An introductory guide to learning theory* (pp. 157–171). Springer.
- Sheridan, B. (2013). How Much Do Our Incoming First Year Students Know?: Diagnostic Testing in Mathematics at Third Level. *Irish Journal of Academic Practice*, 2(1), 3.
- Silalahi, R. M. (2019). Understanding Vygotsky's zone of proximal development for learning. *Polyglot: Jurnal Ilmiah*, 15(2), 169–186. DOI: [dx.doi.org/10.19166/pji.v15i2.1544](https://doi.org/10.19166/pji.v15i2.1544).
- Subali, B. (2018). Content validity analysis of first semester formative test on biology subject for senior high school. *Journal of Physics: Conference Series*, 1097(1), 12039. doi :10.1088/1742-6596/1097/1/012039.
- Surapranata, S. (2019). *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes Implementasi Kurikulum 2004*.
- Syaifuddin, M., Darmayanti, R., & Rizki, N. (2022). Development of a two-tier multiple-choice (TTMC) diagnostic test for geometry materials to identify misconceptions of

- middle school students. *Jurnal Silogisme: Kajian Ilmu Matematika Dan Pembelajarannya*, 7(2).
- Tall, D. (2004). Thinking Through Three Worlds of Mathematics. *International Group for the Psychology of Mathematics Education*.
- Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences*, 39(2), 163–173.
<http://dx.doi.org/10.1080/01490400.2015.1127189>.
- Veith, J. M., Bitzenbauer, P., & Girnat, B. (2022). Assessing learners' conceptual understanding of introductory group theory using the CI2GT: Development and analysis of a concept inventory. *Education Sciences*, 12(6), 376.
<https://doi.org/10.3390/educsci12060376>.
- Widyawati, E. (2025). RELEVANSI KEMAMPUAN PRASYARAT DALAM PENGUASAAN KONSEP MAHASISWA. *Jurnal Penalaran Dan Riset Matematika*, 4(1), 10–15.