

ANALISIS KARAKTERISTIK SOAL UJIAN SEKOLAH HASIL MUSYAWARAH GURU MATEMATIKA DI TASIKMALAYA

Rosid Bahar¹, Edi Istiyono², Widihastuti³, Sudji Munadi⁴,
Zalik Nuryana⁵, Syarief Fajaruddin^{6*}

^{1,2,3,4,6} Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

¹ STAI Al-Andina, Sukabumi, Indonesia

⁵ Nanjing Normal University, China

*Corresponding author.

E-mail: rosidbahar@gmail.com¹⁾
edi_istiyono@uny.ac.id²⁾
widihastuti@uny.ac.id³⁾
sudji.munadi@uny.ac.id⁴⁾
31193033@njnu.edu.cn⁵⁾
syarief.fa@outlook.com^{6*)}

Received 29 October 2021; Received in revised form 16 November 2021; Accepted 28 December 2021

Abstrak

Ujian sekolah merupakan tahap akhir seorang siswa dalam proses pembelajaran. Di tahap ini, seorang siswa akan ditentukan kelulusannya dengan cara melihat salah satu nilai mata pelajaran matematika. Tujuan penelitian ini adalah untuk mendeskripsikan karakteristik butir soal tes matematika dalam instrumen penilaian Ujian Sekolah yang dibuat oleh tim Musyawarah Guru Mata Pelajaran (MGMP). Analisis butir soal untuk membuktikan validitas isi, mengecek estimasi reliabilitas, tingkat kesulitan, daya pembeda soal, dan untuk mengidentifikasi kategori soal yang memiliki nilai rendah, sedang dan tinggi, serta menganalisis faktor-faktor yang menyebabkan item soal-soal tersebut sulit bagi siswa. Validitas content butir instrumen penilaian Ujian Sekolah dihitung menggunakan rumus Formula Aiken dengan melibatkan 3 orang ahli pada bidang matematika, dilanjutkan dengan validitas konstruk menggunakan analisis faktor eksploratori (*exploratory factor analysis*, EFA), sedangkan untuk reliabilitas dihitung dengan rumus KR-20. Subjek dalam penelitian ini berjumlah 125 orang yang diambil dari siswa SMA Negeri dan swasta di Kabupaten Tasikmalaya, Jawa Barat, Indonesia. Pengumpulan data diperoleh melalui hasil jawaban siswa yang dibandingkan dengan kunci jawaban Ujian Sekolah. Analisis data kuantitatif dilakukan dengan menggunakan pendekatan *Classical Test Theory* (CTT) dan *Item Response Theory* (IRT) model *Rasch*. Analisis data kualitatif dilakukan untuk mendeskripsikan tingkat kesulitan siswa dalam menyelesaikan soal-soal tes dengan kategori sulit. Hasilnya menunjukkan bahwa estimasi reliabilitas cukup reliabel. Berdasarkan hasil analisis CTT dan IRT, rata-rata tingkat kesukaran dikategorikan sedang, dan rata-rata kemampuan membedakan dikategorikan baik. Dalam analisis kualitatif diperoleh satu butir soal menjadi sulit dijawab karena pemahaman konsep yang rendah, materi pembelajaran yang belum tuntas, dan kesulitan belajar selama di kelas.

Kata Kunci: *Classical test theory*; *exploratory factor analysis* karakteristik soal; MGMP; matematika.

Abstract

School exams are the final stage of students' learning process. In this stage, the students' graduation will be determined by the grades in four compulsory subjects, one of them is Mathematics. This study aims to describe the characteristics of the mathematics test items in the School Exam instrument composed by the Mathematics Teachers Forum (MGMP). Item analysis is to prove the content validity, check reliability estimate, level of difficulty, discriminant, and identify categories of the questions with low, medium and high scores, as well as analyze the factors that make these items difficult for students. The validity of the item content of the School Exam instrument was calculated using the Aiken formula involving three mathematics experts and using exploratory factor analysis (EFA) for the construct validity. Reliability was calculated using the KR-20 formula. The subjects in this study were 125 people taken from public

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

and private high school students in Tasikmalaya, West Java, Indonesia. Data collection was obtained through the results of students' answers. Quantitative data analysis was carried out using the Classical Test Theory (CTT) and Item Response Theory (IRT) Rasch models. Qualitative data analysis was conducted to describe the level of difficulty of students in solving test questions with complex categories. The results show that the reliability estimate is quite reliable. Based on CTT and IRT analysis results, the average difficulty level is categorized as moderate, and the average ability to distinguish is categorized as good. The qualitative data analysis reveals that one item was difficult to answer because of common concept understanding, incomplete learning materials, and learning difficulties during class.

Keywords: *classical test theory; exploratory factor analysis; item characteristics; MGMP; mathematics*



This is an open access article under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

PENDAHULUAN

Pendidikan merupakan salah satu upaya dalam meningkatkan mutu sumber daya manusia di suatu negara. Pendidikan juga merupakan intervensi kehidupan dan agen perubahan untuk memperluas akses sosial dalam masyarakat secara vertikal dan horizontal (Pakpahan, 2015). Salah satu upaya dalam meningkatkan kualitas pendidikan tersebut, pemerintah Indonesia melalui Menteri Pendidikan Nasional menyelenggarakan ujian sekolah pada semua jenjang pendidikan dari mulai tingkat dasar sampai dengan menengah atas. Ujian ini secara tidak langsung merupakan salah satu cara untuk menaksir tingkat kemampuan peserta didik (Jusrianto et al., 2018; Mardapi & Kartowagiran, 2013).

Pada penelitian ini, ujian yang dimaksud adalah ujian sekolah tingkat menengah atas (SMA) dengan mata pelajaran matematika. Ujian tersebut berupa penilaian secara komprehensif yang memuat soal-soal matematika mulai dari materi kelas 1 sampai dengan kelas 3 tingkat SMA. Soal-soal tersebut biasa dibuat oleh guru melalui musyawarah guru mata pelajaran (MGMP) dengan mempertimbangkan standar kelulusan dari pemerintah.

Soal-soal yang dibuat oleh MGMP pada tiap-tiap kabupaten sudah seharusnya dianalisis. Hal itu penting

dilakukan guna mendapatkan butir-butir soal yang berkualitas untuk mengukur yang seharusnya diukur. Selain itu, penting untuk sebuah instrumen di analisis sebagai bentuk evaluasi atas kelayakan alat ukur yang akan digunakan. Selama ini, hasil analisis instrumen yang dibuat oleh MGMP kabupaten Tasikmalaya tidak pernah dipublikasikan, sehingga timbul keraguan apakah instrumen yang dibuat telah dianalisis secara ilmiah.

Sebagai contoh, penelitian yang dilakukan oleh Hasnah (2017) di SMA Negeri Watansoppeng menggunakan pendekatan teori respon butir (IRT). Hasnah (2017) melaporkan bahwa soal ujian sekolah yang dianalisis dengan model 1 PL, 77,5 % berkategori buruk. Model 2 PL 62,5 % buruk, sedangkan dengan 3 PL, 50 % buruk.

Penelitian lain yang dilakukan oleh Irena et al. (2020) menyatakan bahwa kisi-kisi soal USBN yang dibuat oleh MGMP di salah satu kabupaten di Jawa Barat tidak sesuai dengan soal USBN. Selain itu, soal juga masih dalam kategori LOTS bukan HOTS.

Dua penelitian ini menjadi gambaran untuk analisis instrumen ujian sekolah yang dibuat oleh MGMP matematika di Tasikmalaya, dengan tujuan untuk memperoleh informasi kelayakan instrumen ujian sekolah yang akan digunakan.

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

Penelitian tersebut juga tidak memberikan gambaran atau perbandingan akan dua pendekatan analisis, yaitu *Classical Theory Test* (CTT) dan *Item Response Theory* (IRT). Padahal, masing-masing pendekatan ini bisa menjadi alternatif dalam menganalisis instrumen. Harapan yang sangat besar jika salah satu dari keduanya bisa menunjukkan hasil instrumen yang baik dan menghasilkan kualitas instrumen USBN yang baik, terutama instrumen yang dibuat oleh tim MGMP.

Seperti diketahui bersama bahwa MGMP merupakan kegiatan profesi guru mata pelajaran yang sama di tingkat Kabupaten/Kota yang terdiri dari beberapa guru di sejumlah sekolah sebagai wadah peningkatan kemampuan dan kinerja guru, saling bertukar pikiran dalam memecahkan masalah belajar bersama dan melanjutkan pengembangan profesional (Daryanto et al., 2020; Yulvisriani et al., 2020).

Sebagai organisasi yang berfungsi untuk meningkatkan kemampuan dan kinerja guru, MGMP matematika senantiasa bermusyawarah dalam menyusun instrumen ujian sekolah. Dalam kapasitasnya sebagai seseorang yang memiliki kemampuan kompetensi pedagogik, pendidik harus mampu mengembangkan instrumen yang dapat memotret secara tepat kompetensi yang telah dicapai peserta didik (Istiyono, 2018). Peningkatan kemampuan kompetensi pedagogik tersebut telah direalisasikan oleh MGMP dalam bentuk penyusunan instrumen penilaian Ujian Sekolah secara bersama-sama. Kaitannya dengan pelajaran matematika, instrumen tersebut harus memenuhi berbagai prosedur spesifikasi tes, dimulai dengan menentukan tujuan tes, menyusun kisi-kisi tes, menentukan bentuk dan panjang tes (Mardapi & Kartowagiran, 2013).

Prosedur-prosedur ini sudah selayaknya ada pada instrumen penilaian matematika ujian sekolah. Instrumen ini juga harus memiliki unsur validitas dan reliabilitas guna memperoleh data yang akurat. Validitas instrumen menjadi penting untuk mengukur sejauh mana instrumen itu mampu mengukur apa yang seharusnya diukur, sedangkan reliabilitas terkait pula dengan kesalahan pengukuran, tinggi rendahnya kesalahan pengukuran tergantung dari nilai reliabilitas suatu instrumen (Retnawati, 2016). Kedua unsur ini tentu harus dipenuhi, agar instrumen ini layak digunakan untuk ujian sekolah, dan juga mampu mengukur tingkat kemampuan, pertumbuhan dan perkembangan peserta didik.

Selain dua parameter tersebut, kualitas suatu instrumen juga dapat dilihat dari karakteristiknya melalui analisis butir. Hal tersebut dapat dilakukan dengan dua pendekatan yaitu *Classical Theory Test* (CTT) dan *Item Response Theory* (IRT). Model pengukuran pada *Classical Theory Test* (CTT) dapat diilustrasikan sebagai skor observasi yang terdiri dari skor aktual dan kesalahan pengukuran (Bichi, 2016; Cappelleri et al., 2014; Ismail, 2018). IRT adalah hubungan antara keterampilan atau karakteristik yang diukur dengan instrumen dan respon item (Retnawati, 2014). IRT terkait dengan pengukuran konstruksi hipotesis laten dan hanya dapat diukur secara tidak langsung melalui pengukuran variabel manifes lainnya. Konstruksi hipotesis ini merupakan variabel laten dan seringkali merepresentasikan kompetensi dan keterampilan seseorang. Variabel laten juga disebut sebagai parameter kompetensi yang disimbolkan sebagai θ (Fatkhudin et al., 2016; Natesan, 2011). Sederhananya, Item Response Theory (IRT) adalah

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

hubungan antara kompetensi atau karakteristik yang diukur dengan instrumen dan tanggapan item (Retnawati, 2014).

Informasi lain berkaitan dengan instrumen ini, bahwa tim MGMP melengkapi instrumen dengan lima standar kemampuan matematika, yaitu: pemecahan masalah, penalaran dan pembuktian, komunikasi, koneksi, dan representasi (Ariyanti, 2016; Bernard & Senjayawati, 2019; Kusumawardani et al., 2018). Lima standar tersebut dibuat menjadi 30 butir item soal berbentuk pilihan ganda dengan memperhatikan konstruksi ranah kognitif yaitu Mengingat (C1) Memahami (C2) Mengaplikasikan (C3) Menganalisis (C4) Mengevaluasi (C5) dan Mencipta (C6) (Adams, 2015; Haolader et al., 2015; Irvine, 2017). Butir-butir tersebut diambil dari materi kelas 1 sampai dengan kelas 3 yang terdiri dari soal aljabar, kalkulus, geometri, trigonometri dan statistika. Materi-materi ini diharapkan bisa menjadi ukuran kemampuan matematika peserta didik secara khusus dalam sebuah kategori ahli.

Selain menjadi ukuran tingkat kemampuan, ujian sekolah juga diharapkan bisa menjawab kecenderungan arah orientasi studi lanjutan para peserta didik. Karena selain matematika, ada mata pelajaran lain yang juga di ujikan. Seperti bahasa indonesia, bahasa inggris, geografi, fisika dan kimia. Hasil-hasil ini lah yang nanti bisa menjadi acuan peserta dalam memilih program studi lanjutan. Peserta yang hasil ujiannya bagus di matematika, bisa melanjutkan ke jurusan teknik, dan mereka yang bagus hasil ujiannya di bahasa inggris, mungkin bisa melanjutkan ke jurusan hubungan internasional.

Manfaat lain dari penelitian ini yaitu mengevaluasi kinerja guru dan

MGMP dalam hal pembuatan konstruk instrumen ranah kognitif. Kelayakan instrumen tersebut juga bisa menjadi ukuran dalam meningkatkan kemampuan dan profesionalisme guru di masa yang akan datang. Di sisi lain, guru juga bisa menjadi lebih produkti dalam mengembangkan instrumen untuk penilaian-penilaian berikutnya. Hasil penelitian ini diharapkan dapat mendeskripsikan karakteristik butir soal tes matematika ujian sekolah yang dibuat oleh tim musyawarah guru mata pelajaran (MGMP).

METODE PENELITIAN

Penelitian ini merupakan penelitian deskriptif eksploratif dengan menerapkan pendekatan kuantitatif dan kualitatif. Pengumpulan data dilakukan melalui pendokumentasian ujian sekolah (US) tahun pelajaran 2020/2021 yang berbasis online/daring karena pandemi. Dokumentasi berupa hasil tes matematika pada ujian sekolah yang terdiri dari 30 butir soal, dan karena adanya keterbatasan dalam pendokumentasian, sehingga respon yang didapat hanya sebanyak 125 siswa pada saat ujian sekolah di Kabupaten Tasikmalaya, Jawa Barat.

Analisis data dilakukan secara kuantitatif dan kualitatif. Pertama dilakukan melalui observasi estimasi reliabilitas instrumen, kemudian dilanjutkan dengan analisis *classical theory test* (CTT) dan *item response theory* (IRT) dibantu software R Studio untuk menghitung tingkat kesulitan dan diskriminasi. Yang kedua dilakukan dengan mendeskripsikan item-item yang dikategorikan sulit. Langkah-langkah dalam analisis kualitatif meliputi mengidentifikasi soal-soal tes dengan tingkat kesukaran tinggi dan menganalisis faktor-faktor yang membuat soal-soal itu sulit bagi siswa.

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

Pembuktian validitas suatu butir dikategorikan berdasarkan indeknya. Indeks kurang dari sama dengan 0,4 dikatakan validitasnya kurang, 0,4 – 0,8 dikatakan validitas sedang, dan lebih dari 0,8 dikatakan tinggi. Umumnya, pembuktian validitas yang dapat diterima memiliki nilai indeks lebih dari 0,4 yang berkategori sedang atau tinggi (Istiyono, 2018). Sedangkan untuk reliabilitas, estimasi yang dapat diterima umumnya memiliki koefisien minimum mulai dari 0,80 hingga 0,90 atau minimal sebesar 0,85 (Crocker & Algina, 2008; Saputra et al., 2021). Dalam pendekatan CTT, perkiraan tingkat kesulitan dari 0,3 hingga 0,7 adalah yang terbaik (Dahlan et al., 2020; Jones et al., 2014; Retnawati, 2016). Diskriminasi item masih dapat digunakan jika nilainya tidak negatif (Bagiyono, 2017; Istiyono, 2018; Saepuzaman et al., 2021). Dalam pendekatan IRT, tingkat kesulitan item atau parameter b secara teoritis berkisar dari $-\infty$ hingga $+\infty$ namun secara umum instrumen yang baik berkisar antara -2 hingga +2 (DeMars, 2018; Fajaruddin et al., 2021)

HASIL DAN PEMBAHASAN

Hasil

Instrumen soal yang disusun ini berjumlah 30 item. Setelah direview, pembuktian validitas isi menggunakan rumus Aiken menunjukkan bahwa terdapat 3 item soal memiliki nilai V kurang dari 0,4 yaitu nomor 5, 15 dan 27. Itu artinya, butir – butir tersebut tidak memenuhi standar validitas isi yang telah ditetapkan (Istiyono, 2018).

Saran dari para ahli, item-item tersebut tidak harus dibuang, tetapi hanya perlu diperbaiki. Selanjutnya, pembuktian validitas konstruk menggunakan eksploratori faktor analisis (EFA) melalui uji coba secara empiris kepada 110 siswa. Hasil analisis

menunjukkan bahwa tiap butir memiliki terdapat nilai factor loading kurang dari 0,4 yaitu butir nomor 10, 23, 28 dan 30. Karena butir-butir yang tidak valid antara validitas isi dan konstruk, maka dilanjutkan dengan diskusi bersama validator.

Setelah berdiskusi dengan para validator, maka validator menyepakati bahwa soal yang dipakai untuk ujian sekolah adalah hasil validitas konstruk. Maka butir-butir yang tidak valid pada saat dianalisis menggunakan EFA tersebut dibuang. Sehingga jumlah butir yang digunakan pada instrumen ujian sekolah ini menjadi 26 butir.

Untuk analisis estimasi reliabilitas, instrumen penilaian matematika Ujian Sekolah di uji menggunakan rumus KR 20. Dari hasil analisis diperoleh koefisien reliabilitas sebesar 0,81 dengan kategori tinggi. Selanjutnya, data respon siswa dianalisis menggunakan software program R Studio untuk mengetahui tingkat kesulitan butir soal dan diskriminasi berdasarkan pendekatan CTT. Tabel 1 menunjukkan tingkat kesulitan, dan Tabel 2 menunjukkan daya beda.

Tabel 1. Hasil tingkat kesulitan (*itemMean*)

No. Soal	<i>itemMean</i>	Kategori
1	0,400	Moderat
2	0,440	Moderat
3	0,384	Moderat
4	0,352	Moderat
5	0,400	Moderat
6	0,624	Moderat
7	0,512	Moderat
8	0,648	Moderat
9	0,088	Difficult

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

No. Soal	itemMean	Kategori
11	0,440	Moderat
12	0,400	Moderat
13	0,480	Moderat
14	0,592	Moderat
15	0,496	Moderat
16	0,568	Moderat
17	0,488	Moderat
18	0,504	Moderat
19	0,616	Moderat
20	0,440	Moderat
21	0,424	Moderat
22	0,408	Moderat
24	0,336	Moderat
25	0,184	Difficult
26	0,504	Moderat
27	0,424	Moderat
29	0,288	Difficult

Hasil pada Tabel 1 menunjukkan 3 item sulit, 23 soal sedang dan tidak ada soal mudah. Dari keseluruhan, Rerata tingkat kesulitan adalah 0,448 dan dikategorikan sedang.

Tabel 2. Hasil diskriminan/daya beda (*pBis*)

No. Soal	pBis	Kategori
1	0,498	Good
2	0,313	Good
3	0,403	Good
4	0,365	Good
5	0,523	Good
6	0,364	Good
7	0,582	Good

No. Soal	pBis	Kategori
8	0,277	Good
9	0,265	Good
11	0,250	Good
12	0,441	Good
13	0,451	Good
14	0,364	Good
15	0,453	Good
16	0,278	Good
17	0,504	Good
18	0,281	Good
19	0,413	Good
20	0,398	Good
21	0,185	Good
22	0,293	Good
24	0,366	Good
25	0,209	Good
26	0,451	Good
27	0,393	Good
29	0,000	Good

Tabel 2 menunjukkan bahwa dari 26 item, keseluruhan item berkategori baik dengan rerata diskriminasi keseluruhan adalah sebesar 0,360 dan dikategorikan baik. Secara keseluruhan, dari hasil daya beda pada instrumen tersebut, menunjukkan bahwa instrumen tes memiliki daya pembeda yang baik untuk membedakan kemampuan siswa.

Selanjutnya, dilakukan analisis dengan pendekatan IRT model Rasch menggunakan software program R Studio untuk mengetahui tingkat kesulitan soal. Hasil dari analisis tersebut secara jelas dapat dilihat pada Tabel 3 yang telah disajikan.

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

panjang, dimulai dengan memfaktorkan fungsi yang pertama yaitu fungsi kuadrat $(x^2 + 2x - 3)$ yang menghasilkan dua faktor, kemudian mensubstitusikan hasil pemfaktoran tersebut ke dalam sisa bagi fungsi yang kedua yaitu fungsi linear $(3x - 4)$, dilanjutkan dengan membuat fungsi baru yang dibentuk dari pensubstitusian faktor kepada fungsi kuadrat yang kedua $(x^2 - x - 2)$ dikali fungsi linear baru $(ax + b)$ ditambah fungsi linear yang kedua $(2x + 3)$. Hasil-hasil

substitusi dari fungsi baru tersebut akan membentuk fungsi linear baru yang kemudian dengan sistem persamaan linear dua variabel akan ditemukan nilai a dan b . Langkah terakhir adalah mensubstitusikan kembali nilai a dan b tersebut ke dalam fungsi baru seperti telah disebutkan, maka akan didapat hasil jawaban pada soal tersebut. Kunci jawaban soal tersebut adalah B. Untuk memperjelas langkah-langkah ini, akan disajikan pada Gambar 2.

$$\begin{aligned}
 x^2 + 2x - 3 &= 0 \\
 (x + 3)(x - 1) &= 0 \\
 x &= -3, x = 1 \\
 f(-3) &= 3(-3) - 4 \\
 &= -9 - 4 = -13 \\
 f(1) &= 3 \cdot 1 - 4 \\
 &= 3 - 4 = -1 \\
 f(x) &= P(x) \cdot H(x) + s(x) \\
 f(-3) &= ((-3)^2 - (-3) - 2)(ax + b) + 2x + 3 \\
 -13 &= (9 + 3 - 2)(-3a + b) + 2(-3) + 3 \\
 -13 &= 10(-3a + b) - 3 \\
 -13 &= -30a + 10b - 3 \\
 -30a + 10b &= -13 - (-3) \\
 -30a + 10b &= -13 + 3 \\
 -30a + 10b &= -10 \dots(1) \\
 f(1) &= (1^2 - 1 - 2)(ax + b) + 2x + 3 \\
 -1 &= -2(a + b) + 2 \cdot 1 + 3 \\
 -1 &= -2a - 2b + 2 + 3 \\
 -1 &= -2a - 2b + 5 \\
 -2a - 2b &= -1 - 5 \\
 -2a - 2b &= -6 \dots(2) \\
 -30a + 10b &= -10 \\
 -2a - 2b &= -6 \\
 -30a - 30b &= -90 \\
 \hline
 40b &= 80 \\
 b &= 80/40 = 2 \\
 -2a - 2b &= -6 \\
 -2a - 2(2) &= -6 \\
 -2a - 4 &= -6 \\
 -2a &= -6 - (-4) \\
 -2a &= -6 + 4 \\
 -2a &= -2 \\
 a &= -2/(-2) = 1 \\
 f(x) &= (x^2 - x - 2)(x + 2) + 2x + 3 \\
 f(x) &= x^2 + 2x^2 - x^2 - 2x - 2x - 4 + 2x + 3 \\
 f(x) &= x^2 + x^2 - 2x - 1 \\
 \text{maka suku banyak tersebut adalah } & \mathbf{x^2 + x^2 - 2x - 1}
 \end{aligned}$$

Gambar 2. Algoritma jawaban nomor 9

Berdasarkan analisis CTT tingkat kesulitan soal pada Gambar 2 yaitu 0,088, sedangkan analisis IRT model Rasch sebesar 2,640, keduanya mengategorikan sulit. Hanya 18,4 %

siswa yang menjawab benar pada soal ini, dan kebanyakan siswa memilih A sebagai jawaban dengan persentase 40%. Soal ini masuk tipe *High Order Thinking Skill*, karena siswa harus

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

memahami betul materi prasyarat yang dibutuhkan untuk menjawab soal ini. Siswa harus memahami betul materi sistem persamaan linear dua variabel yang di dalamnya ada metode eliminasi dan substitusi, kemudian harus mengetahui konsep hasil bagi fungsi kuadrat terutama dalam pembentukan fungsi baru. Hal tersebut sepemahaman dengan penelitian (Hasibuan, 2015) yang menyatakan bahwa siswa harus memahami materi prasyarat sebelumnya agar mudah mempelajari materi lanjutan.

Instrumen yang dibuat oleh tim MGMP menyebutkan bahwa soal nomor 9 ini masuk kategori C6, dimana siswa harus menemukan terlebih dahulu nilai-nilai yang menjadi syarat dalam pembentukan fungsi baru. Diskusi dengan guru matematika menjelaskan bahwa sebetulnya soal ini bisa dianggap mudah jika siswa memahami betul konsep hasil bagi pada fungsi. Hanya saja, soal ini menjadi sulit karena memiliki persyaratan yang luar biasa yang harus dipahami betul oleh siswa. Sedangkan materi prasyarat tersebut merupakan materi SMP dan memang masih lemah dikuasai oleh siswa. Untuk mengetahui hal tersebut, penelusuran dilanjutkan kepada guru SMP.

Berdasarkan diskusi dengan guru SMP dan beberapa penelitian sebelumnya (Islamiyah et al., 2018; Sutriningsih et al., 2018) yang menjelaskan bahwa materi sistem persamaan linear dua variabel memang rata-rata sulit dipahami oleh siswa. Tidak seperti materi yang lain yang mungkin pernah disinggung di bangku Sekolah Dasar seperti bilangan, sistem persamaan linear dua variabel memang merupakan materi baru untuk anak SMP, karena sebelumnya siswa belum diperkenalkan materi ini di tingkat Sekolah Dasar.

Selain alasan tersebut, guru matematika SMA juga menyebutkan bahwa materi ini disajikan di kelas 2 SMA yang hampir semua materi di kelas tersebut dirasa cukup sulit. Porsi pembagian materi dari kelas 1 sampai kelas 3 sebetulnya kurang seimbang dalam hal pembobotan kompleksitas. Belum lagi ditambah dengan intake siswa yang memiliki karakteristik yang sulit memahami matematika serta daya dukung yang masih kurang di setiap sekolah. Durasi waktu yang kurang juga menyebabkan guru terkadang menyampaikan materi yang kurang lengkap yang menyebabkan siswa sulit memahami materi tersebut. Oleh karena itu, perlu pengajaran yang efektif untuk menghasilkan pembelajaran yang maksimal (Ngaeni & Saefudin, 2017).

Dalam proses kegiatan belajar mengajar, salah satu komponen vital yang harus dikuasai oleh guru adalah penilaian. Seperti halnya persiapan penyampaian materi, penilaian juga harus dirancang sejak awal untuk memastikan bahwa jenis informasi yang diinginkan dapat dilakukan saat menyampaikan hasil (Opfer et al., 2012). Dalam evaluasi, supaya memperoleh informasi yang bermanfaat, maka diperlukan penilaian yang baik sebagai bentuk tanggung jawab dan tepat pada saat pengambilan keputusan. Informasi yang dibutuhkan selama penilaian diperoleh melalui pengukuran. Pengukuran merupakan suatu proses untuk mendapatkan gambaran numerik mengenai sejauh mana karakteristik tertentu yang dimiliki (Islami & Marthanthi, 2018). Dengan kata lain, model pengukuran yaitu kerangka kerja yang digunakan untuk berkomunikasi dengan orang lain tentang bukti yang diperoleh dari observasi, dapat digunakan untuk menarik kesimpulan mengenai

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

karakteristik siswa yang terdapat dalam variabel konstruk (Opfer et al., 2012). Untuk itu diperlukan alat ukur pada saat ujian berupa sekumpulan item soal untuk memenuhi kebutuhan informasi.

Walaupun soal dibuat oleh tim MGMP, tetapi, salah satu faktor terbesar dalam mempersiapkan butir soal adalah seorang guru (Friatma & Anhar, 2019). Seorang guru harus memperhatikan karakteristik soal. Pengaruh pengabaian ciri-ciri butir dalam tes menyebabkan hasil pengukuran tidak sesuai dengan kompetensi siswa yang sebenarnya akibat butir yang terlalu sulit atau terlalu mudah. Selain itu, dua faktor lain yang berpengaruh adalah kompetensi dan kesiapan siswa. Semakin tinggi kemampuan siswa, semakin banyak pertanyaan yang mampu mereka jawab (Saputra et al., 2021). Bisa dikatakan bahwa soal yang sulit akhirnya tidak bisa menggambarkan kondisi atau kompetensi peserta tes yang sebenarnya. Ini akan berdampak pada hasil penilaian dan tindak lanjut. Hal tersebut tentunya akan berdampak pada tindak lanjut hasil asesmen yang menyebabkan guru tidak dapat memahami siswa atau kelompok mana yang memiliki kompetensi tinggi dan rendah.

Dalam pembahasan disebutkan bahwa soal-soal yang dikategorikan sulit disebabkan oleh pemahaman materi prasyarat yang rendah (Laily, 2019) dan durasi waktu yang dirasa masih kurang. Selain itu, faktor kesulitannya adalah materi yang cenderung banyak dan kompleks serta mengandalkan kemampuan hafalan rumus. Faktor lain seperti kompetensi dan profesionalisme guru, pengalaman guru, dan pemilihan model pembelajaran yang digunakan guru di kelas juga sangat berpengaruh. Untuk itu diperlukan evaluasi terhadap hal-hal tersebut.

Masukan dalam pembelajaran matematika bisa dilakukan dengan mereview atau mengubah bentuk apapun yang menjadi dasar untuk evaluasi ke arah yang lebih baik. Panel Akademi Pendidikan Nasional merekomendasikan untuk memberikan perhatian khusus pada aspek kognisi siswa seperti representasi masalah, penggunaan strategi dan keterampilan pengaturan diri, serta perumusan penjelasan dan interpretasi sebagai aspek pertimbangan yang diperlukan untuk penilaian nasional dan kemajuan pendidikan dalam menilai pencapaian yang lengkap dan akurat (Ketterlin-Geller, 2005; Opfer et al., 2012). Pemahaman konsep siswa juga sangat penting mengingat banyaknya konsep dalam materi matematika terutama di kelas 2 SMA ini. Hal tersebut dapat dilakukan melalui kegiatan pembelajaran dengan membekali siswa dengan model masalah khususnya yang berkaitan dengan konteks nyata (Retnawati et al., 2017). Kemudian, siswa yang telah menguasai pemahaman konseptual tidak hanya mampu mendeskripsikan konsep, tetapi juga dapat menggali konsep dalam situasi yang berbeda dan memanfaatkannya pada bidang yang sesuai (Karagoz & Cakir, 2011).

Selanjutnya materi-materi matematika tersebut sangat erat kaitannya dengan abstrak, oleh karena itu guru harus memahami lingkungan sekitarnya untuk menemukan sumber belajar yang dapat dihubungkan dengan kehidupan sehari-hari seperti etnomatematika. Durasi waktu yang dianggap kurang, menyebabkan materi tidak tersampaikan seluruhnya. Oleh karena itu, sangat penting dilaksanakan program pelatihan pendidikan yang mendukung profesionalisme kinerja guru terutama pelatihan metode pembelajaran dan penguasaan kelas.

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

KESIMPULAN DAN SARAN

Validitas konstruk menunjukkan bahwa 27 butir soal bisa membuktikan validitas, sedangkan tiga soal lain tidak menunjukkan bukti validitas. Estimasi reliabilitas dalam item tes ujian sekolah menunjukkan koefisien 0,80 dan berkategori tinggi. Analisis CTT menunjukkan bahwa tingkat kesulitan dikategorikan sedang dan daya beda juga menunjukkan kategori baik. Hasil analisis IRT, tingkat kesulitan dikategorikan sedang. Beberapa item pada hasil analisis tingkat kesulitan menggunakan CTT dan IRT diperiksa lebih lanjut dengan faktor-faktor yang menyebabkan kesulitan siswa. Hasil pemeriksaan lebih lanjut menunjukkan bahwa item menjadi sulit karena pemahaman konsep siswa terhadap materi pelajaran sangat rendah, ditambah juga dengan materi prasyarat yang belum dipahami dengan baik dan durasi waktu pembelajaran yang dirasa masih kurang. Hal ini diperkuat dengan hasil diskusi dengan para guru MGMP matematika Kabupaten Tasikmalaya yang menyatakan bahwa durasi waktu masih sering terganggu dengan kegiatan-kegiatan sekolah, sehingga diperlukan evaluasi untuk menyelesaikan masalah itu. Selanjutnya, siswa yang telah menguasai pemahaman konseptual yang baik, bisa ditambah dengan menggali konsep dalam situasi yang berbeda dan memanfaatkannya pada bidang yang sesuai dalam pemecahan masalah. Selain itu, guru juga harus memahami lingkungan sekitar untuk menemukan sumber belajar yang dapat dihubungkan dengan kehidupan sehari-hari. Siswa yang telah menguasai pemahaman konseptual tidak hanya mampu mendeskripsikan konsep dengan baik, tetapi juga dapat menggali konsep dalam situasi yang berbeda dan memanfaatkannya pada bidang yang

sesuai dalam pemecahan masalah. Terkait permasalahan durasi waktu yang kurang yang menyebabkan materi tidak tersampaikan seluruhnya, sebaiknya guru harus senantiasa mengikuti program pelatihan pendidikan yang mendukung profesionalisme kinerja guru terutama pada pelatihan metode pembelajaran dan penguasaan kelas.

Saran untuk penelitian selanjutnya adalah menganalisis karakteristik butir instrumen yang dibuat oleh MGMP masing-masing di setiap kabupaten. Hal ini penting dilakukan untuk melihat seberapa besar kelayakan instrumen tes Ujian Sekolah di setiap kabupaten, sebagai bahan evaluasi. Nantinya, soal yang dianalisis pada setiap kabupaten tersebut bisa dijadikan arsip bank soal yang bisa digunakan di masing-masing sekolah antar kabupaten, sehingga bisa mengukur kemampuan siswa di tingkat Nasional.

DAFTAR PUSTAKA

- Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association: JMLA*, 103(3), 152–153. <https://doi.org/10.3163/1536-5050.103.3.010>
- Ariyanti, H. P. (2016). Pengaruh pembelajaran matematika realistik terhadap kemampuan matematis siswa (ditinjau dari kemampuan representasi dan komunikasi). *Buana Matematika: Jurnal Ilmiah Matematika Dan Pendidikan Matematika*, 6(2:), 37–44. <https://doi.org/10.36456/buanamatematika.v6i2:367>
- Bagiyono, B. (2017). Analisis tingkat kesukaran dan daya pembeda butir soal ujian pelatihan radiografi tingkat 1.

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

- Widyanuklida, 16(1), 1–12.
<http://jurnal.batan.go.id/index.php/widyanuklida/article/view/4068>
- Bernard, M., & Senjayawati, E. (2019). Meningkatkan kemampuan koneksi matematik siswa SMP dengan menggunakan pendekatan metaphorical thinking berbantuan software geogebra. *Jurnal Mercumatika : Jurnal Penelitian Matematika Dan Pendidikan Matematika*, 3(2), 79–87.
<https://doi.org/10.26486/jm.v3i2.558>
- Bichi, A. A. (2016). Classical Test Theory: an introduction to linear modeling approach to test and item analysis. *International Journal for Social Studies*, 2(9), 27–33.
- Cappelleri, J. C., Jason Lundy, J., & Hays, R. D. (2014). Overview of Classical Test Theory and Item Response Theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clinical Therapeutics*, 36(5), 648–662.
<https://doi.org/10.1016/j.clinthera.2014.04.006>
- Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory* Mason. Cengage Learning.
- Dahlan, D., Permana, L., & Oktariani, M. (2020). Teacher's competence and difficulties in constructing hots instruments in economics subject . *Cakrawala Pendidikan*, 39(1), 111–119.
<https://doi.org/10.21831/cp.v39i1.28869>
- Daryanto, D., Agung, I., & Siswantari, S. (2020). MGMP teacher organization empowerment in improving students' problem solving ability. *Journal of Educational and Social Research*, 10(1), 152.
<https://doi.org/10.36941/jesr-2020-0014>
- DeMars, C. E. (2018). Classical test theory and item response theory. In *The Wiley Handbook of Psychometric Testing* (pp. 49–73). John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9781118489772.ch2>
- Fajaruddin, S., Retnawati, H., Yusron, E., & Sofyaningsih, V. (2021). Exploring the final examination test item characteristics of Pancasila and civic education in the regency of Gunungkidul. *International Journal of Education and Learning*, 3(3).
<https://doi.org/10.31763/ijele.v3i3.435>
- Fatkhudin, A., Surarso, B., & Subagio, A. (2016). Item response theory model empat parameter logistik pada computerized adaptive test. *Jurnal Sistem Informasi Bisnis*, 4(2), 121–129.
<https://doi.org/10.21456/vol4iss2p121-129>
- Friatma, A., & Anhar, A. (2019). Analysis of validity, reliability, discrimination, difficulty and distraction effectiveness in learning assessment. *Journal of Physics: Conference Series*, 1387(1), 012063.
<https://doi.org/10.1088/1742-6596/1387/1/012063>
- Haolader, F. A., Ali, M. R., & Foyso, K. M. (2015). The taxonomy for learning, teaching and assessing: Current practices at polytechnics in Bangladesh and its effects in developing students' competences. *International Journal for Research in Vocational Education and*

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

- Training*, 2(2), 99–118.
<https://doi.org/10.13152/IJRVET.2.2.9>
- Hasibuan, I. (2015). Hasil belajar siswa pada materi bentuk Aljabar di Kelas VII SMP Negeri 1 Banda Aceh tahun pelajaran 2013/2014. *Jurnal Peluang*, 4(1).
<http://www.jurnal.unsyiah.ac.id/peluang/article/view/5853>
- Hasnah, H. (2017). Analisis kualitas soal matematika ujian sekolah kelas XII IPA SMA Negeri di Watansoppeng berdasarkan teori respon butir. *PEP Educational Assessment*, 1(1).
<https://ojs.unm.ac.id/UEA/article/view/3776>
- Irvine, J. (2017). A comparison of revised Bloom and Marzano's New Taxonomy of Learning. *Research in Higher Education Journal*, 33.
- Islami, V., & Marthanthi, L. R. (2018). Analisis pengukuran kinerja direktorat kepelabuhanan pada Kementerian Perhubungan Kota Jakarta tahun 2016. *Jurnal ECODEMICA*, 2(1), 39–48.
<https://ejournal.bsi.ac.id/ejurnal/index.php/ecodemica/article/view/2642>
- Islamiyah, A. C., Prayitno, S., & Amrullah, A. (2018). Analisis kesalahan siswa SMP pada penyelesaian masalah sistem persamaan linear dua variabel. *Jurnal Didaktik Matematika*, 5(1), 66–76.
<https://doi.org/10.24815/jdm.v5i1.10035>
- Ismail, I. H. (2018). *Estimasi kesalahan pengukuran pendekatan teori tes klasik dan teori respon butir soal tryout SBMPTN TPA*. Universitas Negeri Makassar.
- Istiyono, E. (2018). *Pengembangan instrumen penilaian dan analisis hasil belajar fisika dengan teori tes klasik dan modern*. UNY Press.
- Jones, J., Gaffney-Rhys, R., & Jones, E. (2014). Handle with care! An exploration of the potential risks associated with the publication and summative usage of student evaluation of teaching (SET) results. *Journal of Further and Higher Education*, 38(1), 37–56.
<https://doi.org/10.1080/0309877X.2012.699514>
- Jusrianto, J., Zahir, A., & Megawati, M. (2018). Analisis kualitas tes ujian akhir semester mata kuliah pengetahuan komputer. *Jurnal Studi Guru Dan Pembelajaran*, 1(1), 1–9.
<https://doi.org/10.30605/jsgp.1.1.2018.7>
- Karagoz, M., & Cakir, M. (2011). Problem solving in genetics: Conceptual and procedural difficulties. *Educational Sciences: Theory and Practice*, 11(3), 1668–1674.
- Ketterlin-Geller, L. R. (2005). Knowing what all students know: procedures for developing universal design for assessment. *The Journal of Technology, Learning and Assessment*, 4(2).
<https://ejournals.bc.edu/index.php/jtla/article/view/1649>
- Kusumawardani, D. R., Wardono, W., & Kartono, K. (2018). Pentingnya penalaran matematika dalam meningkatkan kemampuan literasi matematika. *PRISMA, Prosiding Seminar Nasional Matematika*, 1, 588–595.
<https://journal.unnes.ac.id/sju/index.php/prisma/article/view/20201>
- Laily, N. I. (2019). *Profil pemahaman*

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

- konseptual dan pengetahuan prosedural siswa dalam menyelesaikan masalah bangun ruang sisi datar ditinjau dari level Van Hiele* [Fakultas Keguruan dan Ilmu Pendidikan Universitas Jember].
<http://repository.unej.ac.id/handle/123456789/102635>
- Mardapi, D., & Kartowagiran, B. (2013). Pengembangan instrumen pengukur hasil belajar nirbias dan terskala baku. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 15(2), 326–341.
<https://doi.org/10.21831/pep.v15i2.1100>
- Natesan, P. (2011). A review of bayesian item response modeling: Theory and applications. *Journal of Educational and Behavioral Statistics*, 36(4), 550–552.
<https://doi.org/10.3102/1076998611411919>
- Ngaeni, E. N., & Saefudin, A. A. (2017). Menciptakan pembelajaran matematika yang efektif dalam pemecahan masalah matematika dengan model pembelajaran problem posing. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 6(2), 264.
<https://doi.org/10.24127/ajpm.v6i2.896>
- Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, 49(6), 744–777.
<https://doi.org/10.1002/tea.21028>
- Pakpahan, R. (2015). Ujian sekolah sebagai upaya pemetaan mutu sekolah dasar. *Jurnal Pendidikan Dan Kebudayaan*, 21(2), 167–182.
- Retnawati, H. (2014). *Teori respons butir dan penerapannya: Untuk peneliti, praktisi pengukuran dan pengujian, mahasiswa pascasarjana*. Nuha Medika.
- Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian*. Parama Publishing.
- Retnawati, H., Kartowagiran, B., Arlinwibowo, J., & Sulistyaningsih, E. (2017). Why are the mathematics national examination items difficult and what is teachers' strategy to overcome it? *International Journal of Instruction*, 10(3), 257–276.
<https://doi.org/10.12973/iji.2017.10317a>
- Saepuzaman, D., Istiyono, E., Haryanto, H., Retnawati, H., & Yustiandi, Y. (2021). Analisis karakteristik butir soal fisika dengan pendekatan IRT penskoran dikotomus dan politomus. *Radiasi : Jurnal Berkala Pendidikan Fisika*, 14(2), 62–75.
<https://doi.org/10.37729/radiasi.v14i2.1200>
- Saputra, A. N. S., Retnawati, H., & Yusron, E. (2021). Analysis difficulties and characteristics of item test of on biology national standard school examination. *6th International Seminar on Science Education (ISSE 2020)*, 8–14.
<https://doi.org/10.2991/assehr.k.210326.002>
- Sari, R. M. M., Luritawaty, I. P., Dahlan, J. A., & Siregar, G. M. A. (2020). Analisis soal ujian sekolah berstandar nasional matematika sekolah menengah pertama tahun ajaran 2018/2019. *Jurnal THEOREMS (The Original*

DOI: <https://doi.org/10.24127/ajpm.v10i4.4359>

Research of Mathematics), 4(2),
195.

<https://doi.org/10.31949/th.v4i2.1862>

Sutriningsih, N., Pratiwi, R., & Utami, B. H. S. (2018). Penerapan model pembelajaran kooperatif tipe numbered head together (NHT) pada materi sistem persamaan linear dua variabel (SPLDV). *JURNAL E-DuMath*, 4(2), 11–20. <https://doi.org/10.52657/je.v4i2.717>

Yulvisriani, Y., Suryawati, E., & Suzanti, F. (2020). The role of empowerment of dumai science teachers association (MGMP IPA) in improving the professional competency of junior high school science teacher. *Journal of Educational Sciences*, 4(4), 821. <https://doi.org/10.31258/jes.4.4.p.821-833>