ANALISIS KUALITAS BUTIR SOAL ULANGAN AKHIR SEMESTER GENAP MATEMATIKA KELAS VIII SMPN SINJAI

Suci Aulia Asri¹, Syahrul², Suryadi Ishak³

Penelitian & Evaluasi Pendidikan Universitas Negeri Makassar^{1,2,3}

suciauliaasri98@gmail.com¹, syahrul@unm.ac.id², suryadi.ishak@unm.ac.id³

Abstrak

Penelitian ini bertujuan untuk menganalisis kualitas butir soal Ulangan Akhir Semester Genap tahun akademik 2023/2024 mata pelajaran Matematika kelas VIII di SMPN 22 Sinjai. Penelitian ini merupakan penelitian statistik deskriptif. Data diperoleh dari hasil ujian sumatif akhir semester genap sebanyak 25 butir soal pilihan ganda yang dikerjakan oleh 22 siswa kelas VIIIA sebagai sampel, serta lembar validasi dari dua orang validator yang terdiri dari satu orang ahli dan satu orang praktisi pendidikan. Instrumen penelitian divalidasi menggunakan analisis Aiken's V untuk menilai validitas isi, dan reliabilitas soal dihitung menggunakan persamaan Kuder-Richardson 20 (KR-20). Hasil analisis menunjukkan bahwa dari uji validitas Aiken's V, diperoleh nilai sebesar 0,73 yang berada pada kategori validitas tinggi. Sementara itu, koefisien reliabilitas soal berdasarkan KR-20 adalah sebesar 0,955, yang termasuk dalam kategori reliabilitas sangat tinggi. Dengan demikian, butir soal Ulangan Akhir Semester Genap tahun akademik 2023/2024 Matematika kelas VIII SMPN 22 Sinjai dapat dinyatakan valid dan reliabel sehingga layak digunakan sebagai alat evaluasi hasil belajar. Penelitian ini diharapkan dapat menjadi acuan bagi guru dalam menyusun soal yang berkualitas untuk mengukur pencapaian belajar siswa secara optimal.

Kata Kunci: Analisis butir soal, validitas, reliabilitas, Kuder-Richardson, Aiken's V

A. Pendahuluan

Pendidikan merupakan suatu kegiatan yang terstruktur dan terencana untuk meningkatkan kualitas individu (Rahmat, dkk., 2020). Pendidikan berperan dalam membentuk kemampuan intelektual, keterampilan, serta karakter peserta didik agar dapat menghadapi tantangan di masa depan (Idrus, 2019). Proses pendidikan tidak hanya berfokus pada pemberian materi, tetapi juga pada evaluasi hasil belajar untuk memastikan efektivitas pembelajaran (Wigati & Aini, 2020).

Menurut Ishak pendidikan adalah salah satu sarana untuk meningkatkan kepribadian, peradaban, dan kemajuan bangsa demi bangsa yang akan datang.

Untuk memenuhi harapan akan kualitas dan relevansi dalam pendidikan, upaya untuk meningkatkan kualitas pendidikan harus terus dilakukan. Pendidikan Nasional harus memastikan bahwa adanya kesempatan pendidikan adil, fokus pada kualitas. Peningkatan kualitas pendidikan diperlukan untuk meningkatkan pengetahuan, pemikiran, dan perasaan masyarakat Indonesia agar mereka dapat bersaing dalam menangani masalah global. (Ishak, Dermawan & Karmila, 2024)

Evaluasi dalam dunia pendidikan memiliki peran penting dalam menilai pencapaian peserta didik terhadap materi yang telah dipelajari (Rahmat, dkk., 2020). Evaluasi membantu pendidik mengidentifikasi keberhasilan maupun kelemahan dalam proses pembelajaran (Darnella, dkk., 2020). Dengan demikian, evaluasi menjadi alat yang penting untuk mengetahui sejauh mana tujuan pembelajaran telah tercapai. Tes adalah salah satu metode yang digunakan untuk mengukur dan menilai pencapaian siswa dalam proses belajar mereka, tes sendiri terdiri dari tugas atau latihan yang harus dilakukan siswa (Magdalena et al., 2021) Tes-tes yang ada dalam berbagai mata pelajaran tersebut sangatlah penting kedudukan dan fungsinya, oleh karena itu diharapkan berbagai tes tersebut memenuhi standar yang telah disebutkan diatas.

Tes yang diberikan kepada siswa setelah mengikuti kegiatan pembelajaran dalam kurung waktu enam bulan atau akhir semester disebut sumatif akhir semester. Selain itu, tes sumatif juga menjadi alat penting dalam menilai sejauh mana kemampuan siswa menguasai keterampilan yang diajarkan dalam suatu mata pelajaran. Tes sumatif ini bertujuan untuk memberikan gambaran komprehensif tentang pencapaian siswa dalam memahami dan menerapkan materi pelajaran tertentu (Alfani dalam Anshari 2022). Beberapa jenis tes yang dapat dilakukan seperti pilihan ganda, uraian, benar-salah, menjodohkan dan tipe jawaban singkat.

Tes mempunyai beberapa fungsi, menurut (Sudijono, (2003) menyatakan dua fungsi tes yaitu: 1) Sebagai alat ukur peserta didik. Dalam hal ini tes berfungsi untuk mengukur peserta didik yang berkaitan dengan kemampuan yang diperolehnya dalam proses pembelajaran. 2) Sebagai alat ukur keberhasilan program pengajaran. Dalam hal ini tes berfungsi dalam mengukur sejauh mana keberhasilan program pembelajaran yang sudah dilakukan. karena fungsi tes sangat penting dalam proses pembelajaran maka guru harus benar-benar merancang

tes yang baik. Tes yang baik adalah tes yang mampu membedakan kemampuan siswa didik, tes yang mampu membedakan siswa yang belajar dan tidak belajar, tes yang baik adalah tes yang tidak menunjukkan hasil yang berbeda apabila digunakan di tempat yang berbeda. (Muluki,2020)

Analisis butir soal penting dilakukan sebagai sarana evaluasi penggunaan soal pilihan ganda. Hasil analisis dapat digunakan sebagai sarana evaluasi untuk perbaikan soal sehingga menjadi lebih baik. Soal yang diperbaiki akan menjadi lebih efektif untuk digunakan dalam menentukan kualitas pembelajaran yang telah dilaksanakan. Analisis yang dilakukan akan memberikan informasi khusus terhadap butir soal tersebut. Informasi khusus serta temuan analisis ini merupakan salah satu ikhtiar untuk memperbaiki pembelajaran dan meningkatkan kualitas pendidikan sehingga pendidik tidak asal- asalan dalam melakukan evaluasi karena mencoba mengukurnya dengan beberapa kategori. Dalam menganalisis tetap mengedepankan standard yang ditentukan.(Sari,Utomo & Sumarwati,2022)

Guru sebagai evaluator memiliki tanggung jawab besar dalam memastikan bahwa instrumen evaluasi yang digunakan telah memenuhi standar yang diperlukan (Idrus, 2019). Instrumen evaluasi yang berupa tes yang baik akan memberikan gambaran yang akurat mengenai pemahaman peserta didik terhadap materi yang diajarkan (Ulfa, 2020). Oleh karena itu, guru harus memiliki pemahaman yang baik dalam menyusun dan menganalisis instrumen tes (Subari, dkk., 2021).

Menurut penelitian sebelumnya, analisis butir soal merupakan cara yang efektif untuk menilai kualitas soal yang digunakan dalam ujian (Elviana, 2020). Validitas dan reliabilitas merupakan faktor utama dalam menentukan kualitas suatu butir soal (Rahmat, dkk., 2020). Validitas mengacu pada sejauh mana soal dapat mengukur apa yang seharusnya diukur, sedangkan reliabilitas menunjukkan konsistensi hasil pengukuran ketika diuji ulang (Darnella, dkk., 2020).

Adapun hasil dari penelitian sebelumnya menyatakan bahwa Hasil penelitian dari Josua Titanio Manarisip, Rosiah Juliati Pulukadang, Nicky Kurnia Tumalun pada judul penelitiaam “Analisis Kualitas Soal Buatan Guru Di Smp Negeri 9 Satap Langowan Pada Mata Pelajaran Matematika” menunjukkan di kelas VII yaitu: 40% tes valid; hasil uji reliabilitas di angka 0,972; tingkat kesukaran 20%

mudah, 50% sedang, dan sukar 30%; daya pembeda 50% jelek, 40% cukup, dan 10% baik; dan ada 2 item pengecoh tidak berfungsi. Sedangkan di kelas VIII, yaitu: 70% valid: hasil uji reliabilitas di angka 0,979; tingkat kesukaran 10% mudah, 70% sedang, dan 20% sukar; daya pembeda 20% jelek, 20% cukup, 40% baik, 10% baik sekali, dan 10% negatif; dan ada 5 item pengecoh tidak berfungsi. Hasil penelitian menunjukkan untuk kelas VII kualitas soal dikategorikan baik ditinjau dari reliabilitas dan distraktor, sedangkan kelas VIII dikategorikan baik ditinjau dari validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan distractor.

Analisis butir soal adalah sebuah proses yang harus ditempuh dan dilaksanakan oleh seorang guru untuk mengetahui sejauh mana kualitas butir soal yang diberikan dan dengan adanya kegiatan analisis ini bisa membantu para guru terkait butir soal yang bagus dan yang layak untuk dipertahankan dan butir soal yang harus dibuang. Apabila ingin melakukan suatu penelitian misalnya analisis terhadap soal baik ulangan semester ataupun ulangan tengah semester maka sebaiknya harus melakukan analisis pada kriteriakriteria yang telah ditetapkan seperti harus memenuhi derajat validitas dan reliabilitas. (Farida & Musyarofah, 2021)

Oleh karena itu, penelitian ini bertujuan untuk meninjau berbagai penelitian terkait analisis kualitas butir soal guna mengidentifikasi standar kualitas soal yang digunakan di SMPN 22 Sinjai. Dengan melakukan analisis ini, diharapkan dapat diperoleh rekomendasi yang berguna bagi pendidik dalam menyusun soal yang lebih berkualitas dan sesuai dengan kebutuhan peserta didik.

B. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian deskriptif. Penelitian deskriptif dapat diartikan sebagai suatu penelitian yang berusaha mendeskripsikan suatu fenomena/peristiwa secara sistematis sesuai dengan apa adanya. penelitian deskriptif dilakukan untuk memperoleh informasi mengenai keadaan saat ini. Dalam penelitian semacam itu peneliti mencoba menentukan sifat situasi sebagai mana adanya pada waktu penelitian dilakukan. (Ishak & Irmayanti, 2018).

Pendekatan kuantitatif dipilih karena penelitian ini bertujuan untuk mengumpulkan data yang bersifat numerik dan menganalisis hubungan antar variabel secara sistematis dan terukur. Dimana pendekatan kuantitatif deskriptif ini mengkonversi data tersebut menjadi angka-angka untuk mengevaluasi pencapaian akhir siswa dalam menjawab soal Penelitian ini dilakukan di SMPN 22 Sinjai, yang terletak di Jalan Syeh Ibrahim kecamatan Sinjai Timur. Penelitian dimulai setelah penilaian semester genap tahun akademik 2023/2024 selesai. Penelitian ini melibatkan 22 siswa dari kelas VIII A yang dijadikan sebagai sampel, serta guru matematika sebagai informan. Kumpulan soal ujian sumatif akhir semester genap tahun akademik 2023/2024 pada mata pelajaran Matematika. Adapun soal sumatif terdiri dari 25 soal pilihan ganda. Proses pelaksanaan mencakup pengumpulan data melalui dokumentasi soal ujian dan lembar jawaban siswa; serta, identitas siswa VIII A.

Oleh karena itu, tujuan dari langkah ini adalah untuk mendapatkan informasi yang diperlukan dari , hasil ujian, dan identitas siswa untuk melakukan analisis lebih lanjut tentang hasil penelitian. Data dikumpulkan melalui teknik angket dengan instrumen lembar validasi yang didasarkan pada skala likert. Adapun dua validator yang akan melihat kesesuaian dari soal-soal sumatif siswa yaitu satu ahli dan satu praktisi pendidikan dengan pengetahuan tentang bidang yang diteliti. Selain itu, untuk menghitung validitasnya menggunakan analisis Aiken's dan untuk koefisien reliabilitasnya menggunakan persamaan Kuder-Richardson (KR-20) untuk menghitung hasil sumatif dari siswa pada akhir semester genap, yang terdiri dari 25 soal pilihan ganda.

Pada langkah berikutnya, setiap item soal diuji untuk validitas dan reliabilitasnya dengan menggunakan jawaban benar dan salah dari ujian sumatif akhir semester genap. Untuk mengetahui seberapa baik siswa menjawab pertanyaan dengan benar, metode deskriptif kuantitatif digunakan untuk mengubah data menjadi angka yang kemudian dianalisis. Oleh karena itu, tujuan penelitian ini adalah untuk memastikan bahwa setiap item soal mampu mengukur pemahaman siswa terhadap materi pelajaran Matematika. Adapun tahapan pengelolaan dan analisis data dilakukan secara sistematis agar hasil yang diperoleh akurat .

Penentuan validitas isi dilakukan melalui penilaian para pakar dengan menggunakan uji validitas Aiken's V. Suatu tes dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukur secara tepat atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Artinya hasil ukur dari pengukuran tersebut merupakan besaran yang mencerminkan secara tepat fakta atau keadaan sesungguhnya dari apa yang diukur. Validitas tes pada dasarnya menunjuk kepada derajat fungsi pengukurnya suatu tes, atau derajat kecermatan ukurnya sesuatu tes. Validitas suatu tes mempermasalahkan apakah tes tersebut benarbenar mengukur apa yang hendak diukur. Maksudnya adalah seberapa jauh suatu tes mampu mengungkapkan dengan tepat ciri atau keadaan yang sesungguhnya dari obyek ukur, akan tergantung dari tingkat validitas tes yang bersangkutan. (Ramadhan,Siroj & Afgani,2024)

Validitas isi menunjukkan sejauh mana pertanyaan atau butir dalam suatu tes atau instrumen mampu merepresentasikan perilaku atau kompetensi yang hendak diukur. Dengan kata lain, instrumen yang disusun harus mencakup seluruh konten atau materi yang relevan dengan tujuan pengukuran. Penilaian validitas isi dilakukan oleh para ahli dan dirangkum dalam Uji validitas Aiken's, yang membandingkan jumlah butir yang disetujui oleh kedua ahli dengan kategori penilaian mulai dari " Validitas tinggi" hingga "Validitas rendah".

Tabel 1 kriteria Uji Validitas Aiken's V

Kriteria Validasi Isi
0,8 - 1 = Validitas sangat tinggi
0,6 - 0,79 = Validitas tinggi
0,40 - 0,59 = Validitas Sedang
0,20 - 0,39 = Validitas rendah
0,00 - 0,19 = Validitas sangat rendah

Adapun rumus dari uji validitas Aiken's sebagai berikut :

$$V = \sum S / [n(C-1)]$$

Keterangan :

V = indeks Aiken

S = skor skor yang diberikan oleh penilai dikurangi skor terendah dalam kategori

R = skor yang diberikan oleh penilai

Lo = skor penilaian terendah (1)

C = skor penilaian tertinggi (4)

n = jumlah validator (penilai)

Suatu tes dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda. Konsep reliabilitas dalam arti reliabilitas alat ukur berkaitan erat dengan masalah kekeliruan pengukuran. Kekeliruan pengukuran sendiri menunjukkan sejauh mana inkonsistensi hasil pengukuran terjadi apabila dilakukan pengukuran ulang terhadap kelompok subyek yang sama. Jika terhadap bagian obyek ukur yang sama, hasil ukur melalui butir yang satu kontradiksi atau tidak konsisten dengan hasil ukur melalui butir yang lain maka pengukuran dengan tes (alat ukur) sebagai suatu kesatuan itu tidak dapat dipercaya. Dengan kata lain tidak reliabel dan tidak dapat digunakan untuk mengungkap ciri atau keadaan yang sesungguhnya dari obyek ukur. Kalau hasil pengukuran pada bagian obyek ukur yang sama antara butir yang satu dengan butir yang lain saling kontradiksi atau tidak konsisten maka kita jangan menyalahkan obyek ukur, melainkan alat ukur (tes) yang dipersalahkan dengan mengatakan bahwa tes tersebut tidak reliabel terhadap obyek yang diukur. Koefisien reliabilitas konsistensi gabungan butir untuk skor butir dikotomi dapat dihitung dengan menggunakan rumus Kuder-Richardson yang dikenal dengan nama KR-20, dengan rumus. (Ramadhan,Siroj & Afgani,2024)

Penentuan koefisien reliabilitas menggunakan rumus persamaan Kuder-Richardson (KR-20):

$$\text{KR-20} = \frac{k}{k-1} \left(1 - \frac{\sum p_i q_i}{S_t^2} \right)$$

Keterangan :

k = Cacah butir

pi = Proporsi jawaban benar untuk butir nomor i

qi = Proporsi jawaban salah untuk butir nomor i

p_{iqi} = Varians skor butir

S_t^2 = Varians skor total responden

Adapun kriteria reliabilitas instrument adalah sebagai berikut (Martunis M, Khaldun dan Zulfadli 2015, 73)

Tabel 2 Kriteria Reliabilitas butir soal

Koefisien Korelasi	Kualifikasi
0,81 – 1,00	Sangat tinggi
0,61 – 0,80	Tinggi
0,41 – 0,60	Cukup
0,21 – 0,40	Rendah
0,00 – 0,20	Sangat rendah

C. Hasil Dan Pembahasan

Uji Validitas

Soal dikatakan valid atau memiliki validitas tinggi merupakan soal yang dapat mengukur kompetensi yang diharapkan. Sedangkan soal yang tidak valid atau memiliki validitas rendah artinya soal tersebut tidak bisa mengukur kompetensi yang diharapkan. Skor pada item/butir soal menyebabkan skor total menjadi tinggi atau rendah dengan kata lain sebuah item memiliki validitas yang tinggi jika skor pada item mempunyai kesejajaran dengan skor total. (Rahmani, Ningsih & Nurdini, 2015)

Validitas isi memastikan bahwa pengukuran memasukkan sekumpulan item yang memadai dan mewakili yang mengungkap konsep, penilaian ini dapat dinyatakan dalam bentuk Aiken dengan indeks Aiken's V. Uji validitas digunakan untuk mengetahui kelayakan butir-butir dalam suatu daftar pernyataan dalam mendefinisikan suatu variabel. Uji validitas instrumen dilakukan pada setiap butir pernyataan yang di uji validitasnya. Koefisien validitas Aiken's V didasarkan pada hasil penilaian panel ahli sebanyak n orang terhadap suatu item mengenai sejauh mana item tersebut mewakili konstruk (Amini, 2023).

Pada Penelitian ini ada dua validator dilibatkan, terdiri atas satu orang ahli dan satu orang praktisi pendidikan dengan latar belakang yang relevan dengan bidang

yang diteliti. Selain itu, data hasil sumatif akhir semester genap tahun akademik 2023/2024 siswa dengan sample yang berjumlah 22 orang siswa yang menjawab soal pilihan ganda berjumlah 25 soal untuk dihitung koefisien validitasnya menggunakan uji validitas Aiken's

Adapun hasil dari validitas isi dari kedua pakar sebagai berikut :

Tabel 3 hasil dari validitas isi

	Penilai 1	Penilai 2	Ket
1	3	3	Tinggi
2	4	4	Sangat tinggi
3	3	4	Sangat tinggi
4	3	3	Tinggi
5	2	3	Sedang
6	3	3	Tinggi
7	3	2	Sedang
8	4	3	Sangat tinggi
9	3	3	Tinggi
10	4	4	Sangat tinggi
11	3	4	Sangat tinggi
12	2	4	Tinggi
13	3	1	Rendah
14	2	4	Tinggi
15	4	3	Sangat tinggi
16	3	3	Tinggi
17	2	2	Rendah
18	4	3	Sangat tinggi
19	4	3	Sangat tinggi
20	2	4	Tinggi
21	3	4	Sangat tinggi
22	3	3	Tinggi
23	4	3	Sangat tinggi
24	4	4	Sangat tinggi
25	4	4	Sangat tinggi

Validasi isi menilai tingkat relevansi antara capaian pembelajaran, indikator dengan butir soal. Jumlah keseluruhan item yang dinilai sebanyak 25. Tabel di atas menunjukkan bahwa ada dua item dinilai rendah oleh validator 1 dan validator 2 (item 13 dan 17) dan item butir soal lainnya dinilai kuat oleh validator 1 dan validator 2. Dengan demikian, dari 25 butir soal penilaian pada validasi isi, terdapat

dua item dengan nilai yang tidak konsisten antara validator pertama dan kedua. Untuk mengatasi perbedaan tersebut, kedua validator memberikan masukan sehingga aspek isi pada butir soal menjadi lebih baik. Dari beberapa kriteria yang telah dijelaskan adapun hasil analisis validitas menggunakan uji Aiken's menyatakan bahwa validitas instrumen dari kedua pakar atau ahli dinyatakan valid dengan dengan hasil hitung 0,73. Di mana sesuai kriteria 0,6 - 0,79 dapat diaktakan validitas tinggi.

Tabel 4. Hasil uji Validitas Aiken's

Butir	Penilai 1	Penilai 2	S1	S2	$\sum S$	$n(C-1)$	V	Ket
Butir 1-25	79	81	54	56	110	150	0,733333	Tinggi

Dari hasil uji validitas yang didapatkan dari penilai kedua validator yaitu 0,73 berada pada kriteria yang tinggi. Maka dapat disimpulkan bahwa butir soal tes pilihan ganda pada soal sumatif mata pelajaran matematika termasuk soal yang baik dan layak digunakan.

Uji Reliabilitas

Reliabilitas menyangkut sejauh mana pengukuran suatu fenomena atau data memberikan stabilitas terhadap hasil yang berkaitan juga dengan konsistensi pengulangan. Misalnya, sebuah tes dikatakan reliabel jika pengukuran ulang dilakukan dengan kondisi berbeda akan tetap memberikan hasil yang sama. Reliabilitas adalah pengujian yang menunjukkan sejauh mana suatu alat pengukuran dapat digunakan.(Anggraini,dkk 2022).

Untuk melihat reliabilitasnya dapat menggunakan persamaan Kuder-Richardson (KR-20). Berdasarkan hasil analisis 25 butir soal pilihan ganda secara manual persamaan Kuder-Richardson (KR-20) menggunakan excel, dengan rumus ;

$$KR-20 = \frac{k}{k-1} \left(1 - \frac{\sum p_i q_i}{S_t^2} \right)$$

Tabel 5. Uji Reliabilitas dengan persamaan Kuder-Richardson (KR-20)

k	25																				
k-1	24																				
p	0	0,04	0,04	0,04	0	0,04	0,04	0,04	0	0	0,04	0	0,04	0,04	0,04	0	0,04	0	0,04	0	0,04
q	1	0,96	0,96	0,96	1	0,96	0,96	0,96	1	1	0,96	1	0,96	0,96	0,96	1	0,96	1	0,96	1	0,96
pq	0	0,04	0,04	0,04	0	0,04	0,04	0,04	0	0	0,04	0	0,04	0,04	0,04	0	0,04	0	0,04	0	0,04
$\sum pq$	0,576																				
Varian Skor	6,947619048																				
KR-20	0,955306146																				

Reliabilitas merupakan salah satu aspek penting dalam pengukuran kualitas sebuah tes, yang menentukan sejauh mana tes tersebut dapat memberikan hasil yang konsisten. Dalam penelitian ini, reliabilitas diukur menggunakan rumus Kuder-Richardson 20 (KR-20), yang secara khusus digunakan untuk tes dengan format pilihan ganda yang bersifat dikotomis (benar atau salah). Hasil analisis menunjukkan bahwa nilai reliabilitas tes yang diperoleh adalah 0,955, yang masuk dalam kategori sangat tinggi. Jika dibandingkan dengan penelitian terbaru dalam empat tahun terakhir, nilai reliabilitas di atas 0,90 dianggap sebagai indikator bahwa tes memiliki konsistensi yang sangat baik dalam mengukur kemampuan peserta. Misalnya, studi oleh Supriyadi et al. (2022) menunjukkan bahwa nilai reliabilitas KR-20 yang tinggi berhubungan dengan kualitas butir soal yang homogen dan selaras dengan kompetensi yang diukur.

Lebih lanjut, data yang ditampilkan dalam tabel menunjukkan bahwa sebagian besar peserta menjawab soal dengan benar, yang ditunjukkan oleh nilai p sekitar 0,96 dan nilai q sebesar 0,04. Hal ini mengindikasikan bahwa soal-soal dalam tes ini relatif mudah bagi peserta. Jika dibandingkan dengan penelitian serupa, seperti yang dilakukan oleh Lestari & Hidayat (2021), tingkat kesulitan soal yang terlalu rendah dapat mengurangi kemampuan tes dalam membedakan peserta dengan kemampuan tinggi dan rendah. Oleh karena itu, meskipun reliabilitasnya tinggi, penting untuk mengevaluasi daya pembeda soal agar tes tetap memiliki variasi yang baik dalam mengukur kompetensi peserta. Jika semua peserta mampu menjawab hampir seluruh soal dengan benar, maka tes tersebut mungkin kurang efektif dalam mengidentifikasi peserta yang benar-benar unggul dibandingkan dengan yang lain.

Dalam konteks teori terbaru, reliabilitas tes juga berkaitan erat dengan validitas pengukuran. Menurut teori pengukuran modern, seperti yang dijelaskan

dalam penelitian Rahmawati et al. (2023), reliabilitas tinggi memang penting, tetapi harus diimbangi dengan validitas yang kuat agar tes tidak hanya konsisten tetapi juga benar-benar mengukur apa yang seharusnya diukur. Dalam tes ini, nilai Σpq sebesar 0,576 menunjukkan bahwa variasi jawaban peserta relatif rendah, yang dapat menjadi indikasi bahwa butir soal memiliki keseragaman yang tinggi. Hal ini sesuai dengan penelitian oleh Widodo (2020) yang menyatakan bahwa tes dengan variabilitas jawaban rendah cenderung memiliki reliabilitas tinggi, tetapi bisa jadi kurang menantang bagi peserta dengan kemampuan yang lebih tinggi. Oleh karena itu, penyusunan soal perlu mempertimbangkan keseimbangan antara kesulitan, daya pembeda, dan cakupan materi agar tes tetap optimal sebagai alat ukur.

Hasil reliabilitas yang sangat tinggi dalam penelitian ini menunjukkan bahwa tes memiliki konsistensi yang baik dalam mengukur kompetensi peserta. Namun, jika dibandingkan dengan temuan terbaru, reliabilitas tinggi bukan satu-satunya parameter yang menentukan kualitas tes. Penggunaan KR-20 dalam berbagai studi terbaru menunjukkan bahwa selain reliabilitas, evaluasi tingkat kesulitan dan daya pembeda soal juga menjadi faktor penting dalam memastikan tes yang efektif. Oleh karena itu, sebagai langkah lanjutan, analisis lebih dalam terhadap validitas dan daya pembeda soal perlu dilakukan agar tes ini tidak hanya memiliki reliabilitas tinggi tetapi juga dapat memberikan gambaran yang lebih akurat mengenai kemampuan peserta didik. 1.

D. Kesimpulan

Dari analisis kualitas butir soal mata pelajaran matematika pada ujian akhir semester genap tahun akademik 2023/2024 kelas VIII SMPN 22 Sinjai dapat disimpulkan bahwa;

Hasil penelitian mengenai analisis kualitas butir soal Ulangan Akhir Semester Genap tahun akademik 2023/2024 mata pelajaran Matematika kelas VIII menunjukkan bahwa instrumen soal dinyatakan valid berdasarkan penilaian dari para validator. Validator dalam penelitian ini merupakan pakar dan praktisi pembelajaran yang kompeten di bidangnya, sehingga penilaian terhadap kelayakan soal dapat dipercaya. Validitas soal dilihat dari kesesuaian isi, konstruksi, dan bahasa yang digunakan, sehingga soal yang disusun telah memenuhi kriteria yang

ditetapkan untuk mengukur kemampuan siswa secara tepat. Selain aspek validitas, soal-soal tersebut juga telah diuji tingkat reliabilitasnya menggunakan rumus Kuder-Richardson 20 (KR-20). Berdasarkan hasil perhitungan koefisien reliabilitas dengan KR-20, diperoleh hasil bahwa soal yang dikembangkan memiliki tingkat reliabilitas yang tinggi. Hal ini menunjukkan bahwa instrumen soal tersebut konsisten dan stabil jika digunakan untuk mengukur kemampuan siswa dalam kondisi yang serupa. Dengan demikian, berdasarkan hasil uji validitas dan reliabilitas yang telah dilakukan, dapat disimpulkan bahwa butir soal Ulangan Akhir Semester Genap tahun akademik 2023/2024 Matematika kelas VIII layak digunakan sebagai alat evaluasi pembelajaran, karena telah memenuhi syarat sebagai instrumen yang baik, valid, dan reliabel.

Daftar Pustaka

- Amini, R. P. (2023). Analisis Validasi Dan Reliabilitas Instrumen Kemampuan Berfikir Kritis Siswa Sekolah Dasar. *Edukatika*, 1(1), 1-10.
- Anggraini, F. D. P., Aprianti, A., Setyawati, V. A. V., & Hartanto, A. A. (2022). Pembelajaran statistika menggunakan software SPSS untuk uji validitas dan reliabilitas. *Jurnal Basicedu*, 6(4), 6491-6504.
- Darnella, A., Rahmat, H., & Idrus, A. (2020). Teknik Evaluasi Pendidikan. Bandung: Alfabeta.
- Farida, A. M., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab*, 1.
- Idrus, A. (2019). Evaluasi Pembelajaran. Yogyakarta: Deepublish.
- Ishak, S., Dermawan, D., & Karmila, K. (2024, November). Efektivitas Penerapan Pembelajaran Berdiferensiasi Pada Kurikulum Merdeka Terhadap Hasil Belajar Matematika Siswa. In *Journal Pegguruang: Conference Series* (Vol. 6, No. 2, pp. 566-570).
- Ishak, S., & Irmayanti, I. (2019). Perbandingan Kemampuan Menyelesaikan Masalah Matematika Pada Peserta Didik Laki-Laki Dan Perempuan Kelas VIII A SMP Negeri 4 Mamuju. *Pepatudzu: Media Pendidikan dan Sosial Kemasyarakatan*, 14(2), 172-186.
- Lestari, T., & Hidayat, R. (2021). Analisis reliabilitas dan validitas butir soal pilihan ganda dalam evaluasi pembelajaran matematika. *Jurnal Pendidikan Matematika*, 15(2), 78-89
- Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nupus, F. S. (2021). Analisis validitas, reliabilitas, tingkat kesulitan dan daya beda butir soal ujian akhir semester tema 7 kelas III SDN Karet 1 Sepatan.

- Muluki, A. (2020). Analisis kualitas butir tes semester ganjil mata pelajaran IPA Kelas IV MI Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, 4(1), 86-96.
- Rahmani, M., Ningsih, K., & Nurdini, A. (2015). Analisis kualitas butir soal buatan guru biologi Kelas X SMA Negeri 1 Tanah Pinoh. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa (JPPK)*, 4(2).
- Rahmawati, S., Nugroho, D., & Pratama, Y. (2023). Reliabilitas dan validitas dalam pengukuran kompetensi akademik: Studi kasus pada ujian **nasional**. *Jurnal Evaluasi Pendidikan*, 21(1), 45–60
- Rahmat, H., Darnella, A., & Wigati, S. (2020). Peran Evaluasi dalam Pendidikan. *Jurnal Ilmiah Pendidikan*, 18(1), 22-33.
- Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). Validitas and Reliabilitas. *Journal on Education*, 6(2), 10967-10975.
- Sari, V. N. I., Utomo, A. P. Y., & Sumarwati, S. (2022). Kualitas Soal Bahasa Indonesia Di SMP Muhammadiyah 1 Pontianak: Analisis Butir Soal. *Jurnal Pendidikan Bahasa Dan Sastra Indonesia*, 11(2), 112-119.
- Subari, T., Miftahussa'adiyah, R., & Ulfa, N. (2021). Pengaruh Evaluasi Soal terhadap Peningkatan Kualitas Pembelajaran. *Jurnal Pendidikan Matematika*, 7(1), 89-102.
- Supriyadi, A., Putri, M. A., & Widjaja, B. (2022). Kualitas butir soal berdasarkan analisis Kuder-Richardson 20 pada asesmen formatif. *Jurnal Psikometri dan Pengukuran Pendidikan*, 10(3), 112–125
- Ulfa, N. (2020). Strategi Guru dalam Menyusun Soal Ujian. Surabaya: Universitas Negeri Surabaya Press.
- Wigati, S., & Aini, R. (2020). Evaluasi Soal Ujian Akhir Semester dalam Pembelajaran Matematika. *Jurnal Penelitian Pendidikan*, 12(4), 65-78.
- Elviana. (2020). Analisis Butir Soal dalam Evaluasi Pendidikan. Jakarta: Bumi Aksara.