

## Analisis Butir Soal Ujian Akhir Madrasah Berstandar Nasional di Kabupaten Tanah Laut Tahun Pelajaran 2021/2022

Rahmawati<sup>1</sup>, Endar Chrisdiyanto<sup>2</sup>, Yulis Ika Wati<sup>3</sup>, Ami Ramadayanti<sup>4</sup>

Universitas Negeri Yogyakarta<sup>1,2,3,4</sup>

[rahmawatiPII@gmail.com](mailto:rahmawatiPII@gmail.com)

**Abstrak**—Penelitian ini bertujuan untuk mengetahui kualitas soal tes pilihan ganda Ujian Akhir Madrasah matapelajaran matematika wajib di salah satu madrasah di Kabupaten Tanah Laut Tahun Ajaran 2021/2022 dengan menggunakan teori respon butir. Metode penelitian ini menggunakan deskriptif kuantitatif dan instrumen yang digunakan dalam penelitian ini adalah respon peserta didik dalam bentuk tes soal pilihan ganda. Analisis data yang digunakan adalah menggunakan program *Microsoft Office Excel*. Hasil analisis butir soal yang dilakukan pada penelitian ini didapatkan bahwa soal validitas butir soal memiliki tingkat validitas yang tinggi yang dilihat dari 90% soal dalam kategori valid dan tingkat korelasi reliabilitas soal ujian tersebut reliabel atau stabil dan dapat dipercaya. Indeks daya beda didapatkan 20% dalam kategori jelek, 32,5% dalam kategori cukup, 22,5% dengan kategori baik, dan 12,5% memiliki daya pembeda yang baik sekali. Analisis indeks tingkat kesukaran soal diperoleh 12,5% soal dalam sangat sukar, 15% soal dalam katategori sedang, dan 72,5% dalam kategori mudah, sehingga dapat disimpulkan berdasarkan tingkat kesukaran soal masih belum baik. Indeks daya pembeda 35% rendah, 32,5% sedang dan 22,5% tinggi. Daya pengecoh soal 12,5% sangat baik, 22,5% baik, 45% cukup baik, 2 dan 25% kurang baik. Berdasarkan hasil analisis butir soal ini dapat disimpulkan bahwa soal ujian akhir madrasah perlu dilakukan perbaikan agar soal ujian memiliki kualitas yang baik dan layak digunakan dalam ujian akhir madrasah.

**Kata kunci:** Analisis Butir Soal, Pilihan Ganda, Ujian

### I. PENDAHULUAN

Evaluasi dalam proses pembelajaran adalah hal yang sangat penting dalam upaya meningkatkan kualitas pembelajaran dan dapat menjadi acuan dalam mengambil kebijakan selanjutnya. Proses evaluasi digunakan untuk mengetahui tingkat ketercapaian dalam proses pembelajaran sehingga evaluasi dapat menjadi bahan untuk melakukan perbaikan dan perubahan metode maupun pendekatan yang digunakan dalam proses pembelajaran. Basuki dan Hartiyanto (2014) mengatakan evaluasi sebagai proses pengumpulan, analisis dan penafsiran yang sistematis yang bertujuan untuk mengetahui ketercapaian tujuan dari pembelajaran. Sudjana (2009) evaluasi merupakan suatu proses penentuan nilai pada objek tertentu dengan menggunakan kriteria tertentu. Dalam evaluasi pendidikan, proses penentuan nilai ini diberlakukan pada peserta didik. Fungsi evaluasi pendidikan secara khusus dibedakan menjadi 3 yaitu secara psikologis, didaktik, dan administratif (Sudjono, 2011). Bagi pendidik secara psikologis, untuk mengetahui tingkat keberhasilannya dalam proses pembelajaran yang diberikan kepada peserta didik, baik dalam penggunaan metode maupun strategi pembelajaran. Secara didaktik, sebagai bahan untuk mendiganosa peserta didik terkait kendala dan kesulitan yang dihadapi pada saat pembelajaran dan sebagai laporan dan memberikan data keterangan sekaligus memberikan gambaran proses pembelajaran (Sudijono, 2011)

Evaluasi dipandang bukan sebagai hasil dari pembelajaran, namun sebagai suatu proses pembelajaran. Kualitas hasil yang diperoleh dari kegiatan evaluasi akan memiliki nilai tersendiri dalam proses pembelajaran, Sebagaimana menurut Arifin (2014) bahwa pelaksanaan evaluasi dilakukan secara berkelanjutan dan sistematis, dengan kata lain prosesnya terencana sesuai dengan prinsip dan prosedur yang dilakukan tahap demi tahap dan berdampak yang positif terhadap kualitas pembelajaran dalam jangka panjang. Evaluasi hasil belajar dapat dilakukan dengan berbagai macam cara diantaranya melalui tes (Hadiana, 2015). Tes merupakan suatu proses untuk melihat ketercapaian kompetensi yang telah dikuasai oleh peserta didik pada aspek hasil belajar tertentu. Tes ini juga dapat dikatakan sebagai evaluasi formatif, hal ini dikarenakan guru ingin mengetahui setiap penguasaan materi setiap bab yang telah diajarkan oleh guru (Suherman, 2018). Tes menurut waktunya dapat dibagi menjadi tes formatif dan sumatif. Tes formatif

merupakan tes yang dilaksanakan selama proses pembelajaran masih berlangsung atau setelah suatu topik atau pokok bahasan telah selesai dipelajari, seperti ulangan harian.

Pada salah satu tes sumatif, seperti ujian akhir semester, bentuk soal pada tes adalah pilihan ganda. Tes pilihan ganda merupakan tes yang memiliki beberapa alternatif jawaban. Peserta didik memilih salah satu jawaban yang tepat dari setiap soal dan mengisikannya pada lembar jawab. Setelah terjawab, maka penilaian dapat dilakukan dan nilai yang berwujud angka dapat mengidentifikasi hasil belajar peserta didik. Kualitas soal perlu diketahui karena jika kualitas soal belum diketahui, cenderung terdapat kesalahan pada penafsiran hasil tes. Hal ini akan berdampak pada biasanya informasi dari tes mengenai kemampuan belajar peserta didik yang sesungguhnya. Soal tes pilihan ganda dapat digunakan untuk menguji apakah alat ukur evaluasi yang digunakan tersebut memang benar-benar bisa dijadikan dasar untuk menentukan langkah yang akan dipilih selanjutnya secara tepat.

Hal pertama yang harus dilakukan sebelum tes tersebut dilaksanakan oleh peserta didik adalah menguji kualitas tes tersebut. Arikunto (2010) menyatakan bahwa cara untuk mengidentifikasi kualitas soal tersebut baik, kurang baik dan jelek adalah dengan cara menganalisis soal tersebut. Analisis butir soal pilihan ganda merupakan suatu prosedur yang sistematis, yang akan menyajikan data-data yang khusus terhadap butir soal evaluasi yang disusun (Kunandar, 2014). Secara umum, metode analisis butir soal meliputi dua hal yakni: analisis validitas dan analisis reliabilitas soal. Pada soal pilihan ganda analisis butir soal juga akan mengetahui beberapa hal diantaranya adalah validitas tes, reliabilitas butir soal, tingkat kesukaran, daya pembeda, dan daya pengecoh soal. Bermutu atau tidaknya butir soal biasanya diketahui dari tingkat kesukarannya atau taraf kesulitan dari soal tersebut yang dimiliki oleh setiap butir soal. Butir soal dapat dikatakan bagus, apabila tiap butir soal tersebut termasuk ke dalam kategori sedang yang tidak terlalu sukar dan tidak pula terlalu mudah, apabila butir soal yang digunakan itu terlalu mudah atau pun terlalu sukar maka guru tidak akan dapat membedakan peserta didik yang berkemampuan rendah dan peserta didik yang berkemampuan tinggi (Arikunto, 2013). Anwar (2014) menyatakan bahwa soal yang tergolong baik mempunyai tingkat kesukaran 0,30 hingga 0,70.

Pada soal pilihan ganda, setiap butir soal telah dilengkapi dengan beberapa kemungkinan jawab. Salah satu diantara pilihan tersebut merupakan jawaban yang benar sedangkan pilihan yang lain merupakan pengecoh (distraktor). Butir soal yang memiliki kualitas yang baik adalah jika keseluruhan distraktornya berfungsi. Jika butir soal dijawab dengan benar oleh semua peserta didik maka hal ini mengindikasikan distraktor dari butir soal tersebut tidak berfungsi karena peserta didik akan langsung merujuk kepada jawaban yang benar. Sedangkan jika butir soal dipilih secara tidak merata maka hal tersebut bertanda bahwa keseluruhan distraktornya dapat berfungsi sebagai pengecoh jawaban. Menurut Tilaar & Hasriyanti (2019) fungsi pengecoh akan berjalan sesuai dengan fungsinya jika dipilih oleh setidaknya 15% dari seluruh peserta yang ikut tes.

Pelaksanaan Ujian Akhir Berstandar Nasional di Kabupaten Tanah Laut Tahun Ajaran 2021/2022 untuk mata pelajaran matematika wajib, menggunakan soal pilihan ganda. Soal-soal pada ujian tersebut belum diujicobakan sehingga belum diketahui kualitas soal dan dikhawatirkan akan memberikan informasi yang bias pada hasil belajar peserta didik yang sebenarnya. Hasil pengukuran itu maka akan memberikan keyakinan pada guru akan tepatnya keputusan yang akan diambil terhadap peserta didik. Butir soal evaluasi pilihan ganda yang kecil akan menjadi bahan evaluasi pada pembelajaran berikutnya sedangkan butir soal evaluasi pilihan ganda yang mencukupi bahkan melampaui akan menjadi bahan pertimbangan juga pada pembelajaran berikutnya. Brennan (2006) menyatakan tes yang valid adalah tes yang dapat digunakan untuk mengukur apa yang akan diukur.

Oleh karena itu, perlu dilakukan evaluasi pada soal pilihan ganda yang digunakan pada Ujian Akhir Madrasah. Evaluasi ini nantinya digunakan untuk mengetahui kesesuaian soal untuk mengukur kemampuan peserta didik dan dapat digunakan untuk menentukan perbaikan kebijakan.

## II. METODE PENELITIAN

Metode penelitian ini menggunakan pendekatan kuantitatif dengan data berupa soal Ujian Akhir Madrasah Berstandar Nasional dan lembar jawaban peserta didik Kelas XII MA di Kabupaten Tanah Laut. Penelitian ini bertujuan untuk melihat kualitas soal ujian mata pelajaran matematika wajib. Prosedur atau tahapan penelitian pada penelitian kali ini sebagai berikut: (1) menyusun instrumen atau hal - hal yang berkaitan dengan pengukuran, (2) mengumpulkan data, (3) mendeskripsikan data yang telah dikumpulkan, (4) menganalisis data, (5) merumuskan simpulan dari hasil analisis. Analisis butir soal dilakukan oleh peneliti dengan menggunakan teori pengukuran klasik. Populasi dalam penelitian ini adalah seluruh peserta didik kelas XII yang mengikuti Ujian Akhir Madrasah Berstandar Nasional Kabupaten Tanah Laut. Sampel dalam penelitian peserta didik kelas XII yang berjumlah 93 peserta didik. Teknik analisis data kuantitatif

dilakukan untuk mengetahui validitas, reliabilitas, tingkat kesukaran, daya pembeda dan tingkat pengecoh soal menggunakan program aplikasi Microsoft Office Excel. Adapun tehnik penelitian ini yaitu:

a. Analisis Validitas Item Soal

Analisis validitas ini digunakan menentukan kevalidan dari soal yang digunakan dan digunakan dua kriteria penilaian soal yaitu valid dan tidak valid. Adapun rumus yang digunakan untuk menentukan validitas item adalah dengan menggunakan tehnik korelasi *product moment* sebagaimana berikut.

$$r_{xy} = \frac{N \sum xy - \sum x \sum y}{\sqrt{(N(\sum x^2) - (\sum x)^2)(N(\sum y^2) - (\sum y)^2)}}$$

Keterangan:

$r_{xy}$  : Koefisien Validitas

$x$  : Jawaban benar dari setiap butir soal

$y$  : Skor total

$N$  : Jumlah peserta tes

Adapun kriteria koefisien validitas butir soal yang digunakan adalah sebagai berikut:

0,91 – 1,00 : Sangat tinggi

0,71 – 0,90 : Tinggi

0,41 – 0,70 : Cukup

0,21 – 0,40 : Rendah

Negative – 0,20: Sangat Rendah

(Masidjo,2007)

b. Analisis Reliabilitas

Analisis reliabilitas digunakan untuk mengetahui bagaimana ketercapaian setelah soal digunakan untuk penelitian. Tes yang yang reliable akan memberikan hasil yang sama jika diberikan kepada kelompok yang sama meskipun dalam waktu yang berbeda. Penelitian ini menggunakan metode belah dua (*split-half method*) dengan rumus *Spearman-Broen* sebagai berikut.

$$r_{11} = \frac{2r^{1/2} 1/2}{1 + r^{1/2} 1/2}$$

Keterangan:

$r_{11}$  : Koefisien realibilitas

$r^{1/2} 1/2$  : Korelasi antara skor-skor tiap belahan tes

c. Analisis Tingkat Kesukaran

Tingkat kesukaran dihitung menggunakan indeks kesukaran anatara 0.00 sampai dengan 1.00. Kriteria indeks kesukaran soal dibagi menjadi 3 yaitu sukar, sedang, dan mudah (Sudjino, 2011). Adapun rumus yang digunakan adalah sebagai berikut.

$$P = \frac{B}{J_s}$$

Keterangan:

$P$  : Tingkat Kesukaran

$B$  : Jumlah siswa yang menjawab benar

$J_s$  : Jumlah seluruh peserta tes

Adapun kriteria indeks kesukaran soal sebagai berikut.

0,00 – 0,30 : Sukar

0,31 – 0,70 : Sedang

0,71 – 1,00 : Mudah

d. Analisis Daya Pembeda

Daya beda pembeda adalah kemampuan butir soal dapat membedakan antara peserta yang telah menguasai dan peserta didik yang belum menguasai materi yang diujikan (Kusaeri dan Suprananto,2012). Rumus yang digunakan sebagai berikut.

$$DP = \frac{BA - BB}{\frac{1}{2}N}$$

Keterangan:

$DP$  : Daya pembeda

$BA$  : Jumlah jawaban benar pada kelompok atas

$BB$  : Jumlah jawaban benar pada kelompok bawah

$N$  : Jumlah peserta tes

Kriteria daya pembeda adalah sebagai berikut.

$0,71-1,00$  : Baik sekali

$0,41 \leq 0,70$  : Baik

$0,21 < 0,40$  : Cukup

$\leq 0,20$  : Jelek

e. Analisis Pengecoh

Analisis pengecoh digunakan dengan tujuan untuk mengetahui bagaimana pengecoh pada soal tersebut dapat berfungsi dan digunakan dengan baik. Pengecoh dapat dikatakan baik jika dipilih sekurang-kurangnya 5% dari seluruh peserta didik yang mengikuti tes (Arikunto,2012). Adapun rumus yang digunakan adalah sebagai berikut.

$$D = \frac{A}{N} \times 100\%$$

Keterangan:

$D$  : tingkat distraktor (%)

$A$  : Jumlah peserta tes yang memilih opsi tersebut

$N$  : Jumlah peserta tes

### III. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mendeskripsikan kualitas butir soal pilihan ganda pada ujian akhir madrasah berstandar nasional mata pelajaran matematika wajib kelas XII di salah satu madrasah di kabupaten Tanah Laut tahun ajaran 2021/2022 yang ditinjau dari segi validitas, reliabilitas, tingkat kesukaran, daya pembeda dan daya pengecoh soal. Menelaah butir soal tersebut dibutuhkan instrumen berupa soal yang digunakan saat ujian dan kunci jawaban dari sampel penelitian. Jumlah soal yang dianalisis sebanyak 40 soal dalam bentuk pilihan ganda. Data-data hasil penelitian akan dianalisis menggunakan program microsoft office excel. Adapun hasil validitas soal yang dilakukan didapatkan hasil sebagai berikut.

TABEL 1. HASIL ANALISIS VALIDITAS BUTIR SOAL

Kategori	Jumlah	Persentase	Nomor Soal
Valid	36	90%	1,2,3,4,5,6,7,9,10,11,12,13,14,15,16,17,18,19,21,22,23,24,25,26,27,28,29,30
Tidak Valid	4	10%	8, 18,20, 30
Jumlah	40	100%	

Berdasarkan hasil analisis kuantitatif tingkat validitas butir soal ujian akhir nasional berstandar nasional mata pelajaran matematika wajib kelas XII MAN Tanah Laut sebagaimana pada tabel 1, dihasilkan bahwa 90% soal masuk dalam kategori valid dan 10% soal tidak valid. Soal yang tidak valid dilakukan perbaikan sehingga bisa valid dan layak untuk digunakan. Validitas memiliki nilai yang beragam mulai dari yang rendah hingga tinggi, semakin tinggi nilai validitas maka semakin baik tingkat kevalidannya (Susetyo, 2011). Validitas soal berkaitan dengan sejauh mana butir soal dapat mengukur hasil belajar yang hendak diukur dari peserta didik. Proporsi tingkat validitas soal secara umum adalah valid sehingga secara keseluruhan soal pilihan ganda pada ujian tersebut memiliki tingkat validitas yang tinggi dan layak untuk diimplementasikan pada ujian akhir peserta didik. Selain validitas juga dilakukan analisis reliabilitas butir soal. Adapun hasil analisis reliabilitas soal didapatkan data sebagai berikut.

Hasil analisis reliabilitas butir soal pilihan ganda pada ujian akhir madrasah berstandar nasional mata pelajaran matematika wajib Kelas XII di salah satu madrasah di kabupaten Tanah Laut di atas menunjukkan tingkat korelasi reliabilitasnya cukup tinggi dengan indeks 0,858. Indeks tersebut menunjukkan bahwa soal tersebut reliabel atau hasil penilaian satabil dan dapat dipercaya. Reliabilitas tes merupakan tingkat kekonsistenan suatu tes untuk menghasilkan skor yang konsisten sebagaimana menurut Kusaeri dan Suprananto (2012) yang mengatakan bahwa reliabilitas adalah ketepatan skor tes pada pengulangan pengukuran berulang kali. Hamimi (2020) menjelaskan bahwa koefisien reliabilitas mempunyai peranan besar pada tingkat kesukaran soal karena semakin rendah nilai reliabilitas soal maka semakin sukar soal tersebut.

Indeks Tingkat Kesukaran berfungsi untuk memberikan informasi mengenai butir soal apakah butir soal tersebut terlalu sulit atau terlalu mudah bagi peserta didik. Apabila butir soal terlalu sulit, maka peserta didik tidak dapat menjangkau jawaban dan tidak sesuai dengan kemampuan peserta didik. Apabila soal terlalu mudah, maka tidak akan ada tantangan bagi peserta didik untuk menjawabnya. Sedangkan, butir soal yang

baik adalah butir soal yang memiliki tingkat kesulitan sedang, sehingga tidak terlalu sulit dan tidak terlalu mudah bagi peserta didik. Selain itu, dengan tingkat kesukaran sedang, maka soal tersebut mengetes peserta didik sesuai dengan kemampuannya. Indeks tingkat kesukaran butir soal dibagi dalam lima kategori, yaitu  $0,00 \leq 0,30$  merupakan kategori soal sangat sukar,  $0,31 \leq 0,70$  dalam kategori sedang dan  $0,71 \leq 1,00$  masuk dalam kategori sangat mudah dan  $0,87 - 1,00$  (Arikunto, 2012). Soal yang baik mempunyai kriteria tingkat kesukaran yang tidak terlalu mudah dan tidak terlalu sukar (Arikunto, 2013). Adapun hasil analisis tingkat kesukaran pada butir soal didapatkan hasil sebagai berikut.

TABEL 2. HASIL ANALISIS TINGKAT KESUKARAN

Indeks	Kategori	Jumlah	Persentase
$0,71 \leq 1,00$	Sangat Mudah	29	72,5%
$0,31 \leq 0,70$	Sedang	6	15%
$0,00 \leq 0,30$	Sangat Sukar	5	12,5%

Berdasarkan data tabel 2 di atas, terdapat 5 butir soal atau setara dengan 12,5% masuk dalam kategori sangat sukar, 6 butir soal dalam kategori sedang (15%) dan 29 butir soal dalam ketegori sangat mudah (29%). Soal pilihan ganda pada ujian ini didominasi oleh butir sola dengan tingkat kesukaran yang sangat muda yaitu 29 butir soal (72,5%). Butir soal yang telalu mudah kurang dapat merangsang peserta didik untuk memecahkan masalah, sedangkan butir soal yang terlalu sulit dapat menyebabkan peserta didik kesulitan hingga putus asa dalam mengerjakan soal. Silverius (1991), soal tes dikatakan baik apabila memiliki tingkat soal kategori mudah mencapai 27%, kategori sedang mencapai 46% dan kategori sukar mencapai 27%. Berdasarkan teori tersebut, soal ujian akhir madarasah kelas XII MAN Tanah Laut mata pelajaran matematika wajib masih belum baik dan perlu dilakukan revisi.

Arikunto (2012) menyatakan bahwa kategori daya pembeda dibagi menjadi 4 yaitu baik sekali, baik, cukup, dan jelek. Adapun hasil analisis daya pembeda disajikan dalam tabel berikut ini.

TABEL 3. HASIL ANALISIS DAYA PEMBEDA

Indeks	Kategori	Jumlah	Persentase
0,71-1,00	Baik sekali	5	12,5%
$0,41 \leq 0,70$	Baik	9	22,5%
$0,21 < 0,40$	Cukup	18	45%
$\leq 0,20$	Jelek	8	20%

Berdasarkan hasil perhitungan daya pembeda menggunakan microsoft office excel, menyatakan bahwa 8 butir soal dalam kategori jelek (20%), 18 butir soal (32,5%) memiliki daya pembeda yang cukup dan 9 butir soal (22,5%) memiliki daya pembeda yang baik, dan 5 butir soal (12,5%) memiliki daya pembeda yang baik sekali.

Butir soal yang berkualitas adalah apabila fungsi pengecohnya dipilih rata oleh peserta didik yang menjawab salah. Sebaliknya, pengecohnya dipilih secara tidak merata itu artinya butir soal tersebut termasuk dalam kategori buruk. Pengecoh dikatakan baik bila jumlah peserta didik yang memilih pengecoh itu ideal dengan kata lain mendekati jumlah yang sama. Pengecoh dapat dikatakan baik jika dipilih sekurang-kurangnya 5% dari seluruh peserta didik yang mengikuti tes (Arikunto,2012). Adapun hasil analisis daya pengecoh sebagai berikut.

TABEL 4. HASIL ANALISIS DAYA PENGECOH

No	Efektivitas Pengecoh	Jumlah	Persentase
1.	Sangat Baik	5	12,5%
2.	Baik	9	22,5%
4.	Cukup Baik	18	45%
5.	Kurang Baik	8	25%

Berdasarkan tabel 4 diketahui butir soal yang memiliki pengecoh sangat baik sebanyak 5 soal (12,5%), butir soal yang memiliki pengecoh baik sebanyak 9 soal (22,5%), butir soal yang mempunyai pengecoh cukup baik sebanyak 18 soal (45%), butir soal yang memiliki pengecoh kurang baik sebanyak 8 soal (25%) dan butir soal yang memiliki pengecoh tidak baik sebanyak 0 soal (0%). Firmansyah (2018) menjelaskan bahwa suatu pengecoh tidak berfungsi dengan baik apabila hasil pada setiap opsi butir soal menyatakan salah satu dari kurang baik, jelek, atau sangat jelek. Oleh karena itu, 8 soal yang berada pada kategori kurang baik perlu direvisi agar valid dan layak untuk digunakan untuk ujian akhir selanjutnya.

Berdasarkan hasil analisis kevalidan soal didapatkan 90% soal dalam kategori sangat valid dan 10% soal dalam kategori tidak valid. Butir soal yang tidak valid kemudian di perbaiki agar memenuhi kategori valid dan layak untuk digunakan selanjutnya. Kemudian untuk daya pembeda diperoleh daya pembeda soal rendah sebanyak 14 butir soal. Daya pembeda soal yang rendah ini kemudian diperbaiki dan disusun kembali agar layak untuk digunakan dengan minimal dalam kategori sedang. Hal ini diperkuat oleh Firmansyah (2018) yang menyatakan bahwa butir soal dengan tingkat daya pembeda rendah masih dapat diperbaiki atau diganti dengan soal yang lain. Kusaeri dan Suprananto (2012) mengungkapkan bahwa daya pembeda soal adalah kemampuan suatu butir soal untuk dapat membedakan antara peserta didik yang telah menguasai materi dan peserta didik yang belum menguasai materi yang diujikan. Semakin tinggi indeksnya maka soal semakin mampu untuk membedakan peserta didik yang kompeten dengan peserta didik yang belum kompeten. Namun, jika ada daya pembeda yang berindeks negatif, itu artinya banyak peserta didik yang belum memahami materi dibandingkan dengan kelompok peserta didik yang memahami materi. Kemudian untuk hasil analisis daya pengecoh didapatkan bahwa soal dengan daya pengecoh cukup sebanyak 18 soal dan kurang baik sebanyak 8 soal. Soal dengan daya pengecoh ini perlu diperbaiki agar memiliki daya pengecoh yang baik ketika digunakan. Berdasarkan hasil analisis kevalidan, daya pembeda, dan daya pengecoh menunjukkan bahwa soal yang masih dalam kategori kurang valid, daya pembeda rendah, dan pengecoh kurang baik perlu diperbaiki. Analisis butir soal ini bisa digunakan sebagai evaluasi dan penentuan kebijakan dalam menyusun soal ujian akhir madrasah di tahun ajaran selanjutnya.

#### IV. SIMPULAN DAN SARAN

Berdasarkan analisis butir soal pilihan ganda pada Ujian Akhir salah satu Madrasah Tanah Laut Tahun Ajaran 2021/2022 pada mata pelajaran matematika wajib didapatkan validitas butir soal dalam kategori tingkat validitas yang tinggi, korelasi reliabilitas soal cukup tinggi dan dapat dipercaya dengan indeks sebesar 0,858. Indeks tingkat kesukaran soal diperoleh soal dengan butir soal sangat sukar (12,5%), 6 butir soal sedang (15%) dan 29 butir soal sangat mudah (72,5%), sehingga tingkat kesukaran soal masih belum baik. Indeks daya pembeda yang memiliki daya pembeda 8 butir soal dalam kategori jelek (20%), 18 butir soal (32,5%) memiliki daya pembeda yang cukup dan 9 butir soal (22,5%) memiliki daya pembeda yang baik, dan 5 butir soal (12,5%) memiliki daya pembeda yang baik sekali. Daya pengecoh yang sangat baik sebanyak 5 soal (12,5%), pengecoh sebanyak baik 9 soal (22,5%), pengecoh cukup baik sebanyak 18 soal (45%), pengecoh kurang baik sebanyak 8 soal (25%) dan pengecoh tidak baik sebanyak 0 soal (0%). Berdasarkan hasil analisis butir soal ini dapat disimpulkan bahwa soal ujian akhir madrasah perlu dilakukan perbaikan agar soal ujian memiliki kualitas yang baik dan layak digunakan dalam ujian akhir madrasah. Selain itu juga hasil analisis butir soal ini dapat digunakan sebagai pertimbangan dalam penyusunan soal ujian akhir madrasah pada tahun ajaran selanjutnya.

Saran bagi penelitian selanjutnya untuk peneliti perlunya dokumentasi mengenai pelaksanaan ujian dapat dilakukan dan untuk pembuat soal sebaiknya dilakukan analisis butir soal agar dapat meningkatkan kualitas butir soal serta layak untuk diujikan kepada peserta didik.

#### UCAPAN TERIMA KASIH

Terimakasih kepada guru dan salah satu sekolah madrasah di kabupaten Tanah Laut yang membantu dalam memfasilitasi penelitian ini.

#### DAFTAR PUSTAKA

- [1] Anwar, C. 2017, Teori – teori Pendidikan, Yogyakarta: Ircisod.
- [2] Arikunto, S. 2010. Dasar-dasar Evaluasi Pendidikan. Jakarta: Bumi Aksara.
- [3] Arikunto, S. 2012. Dasar-Dasar Evaluasi Pendidikan edisi 2. Jakarta: Bumi Aksara
- [4] Arikunto, S. 2013. Dasar-dasar Evaluasi Pendidikan. Cetakan ke-2, jakarta: Bumi Aksara. Brennan, R.L. 2006. Educational measurement. Fourth Editon. Praeger Publishers, 88 Post Road West, Westport CT. 06881.
- [5] Firmansyah. 2018. Analisis Butir Soal Tes Pilihan Ganda Mata Pelajaran Pendidikan Jasmani Olahraga Dan Kesehatan Kelas VIII SMP N 1 Wonosari Tahun Ajaran 2017/2018. Yogyakarta: Universitas Negeri Yogyakarta.
- [6] Garvin, A. D., & Ebel, R. L. (1980). Essentials of Educational Measurement. Educational Researcher, 9(9), 21. <https://doi.org/10.2307/1175572>
- [7] Hadiana, D. 2015. Penilaian Hasil Belajar Untuk Peserta didik Sekolah Dasar. Jurnal Pendidikan dan Kebudayaan, Vol. 21, Nomor 1, April 2015, Hal: 15-25.
- [8] Hamimi, L., Zamharirah, R., & Rusydy, R. (2020). Analisis butir soal ujian Matematika kelas VII semester ganjil tahun pelajaran 2017/2018. Mathema: Jurnal Pendidikan Matematika, 2(1), 57-66.
- [9] Kunandar. 2014. Penilaian Autentik (Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013). Jakarta: PT Raja Grafindo Persada.
- [10] Kurniawan, Devi, Dwi, 2019, Analisis Butir Soal Ujian Akhir Semester Matematika Dengan Teori Respon Butir, Brilliant: Jurnal Riset dan Konseptual Volume 4 Nomor 2, mei 2019, DOI: <http://dx.doi.org/10.28926/briliant.v4i2.316>

- [11] Kusaeri dan Suprananto. 2012. Pengukuran dan Penilaian Pendidikan. Yogyakarta: Graha Ilmu.
- [12] Masidjo. 2007. Penilaian Pencapaian Hasil Belajar Siswa di Sekolah. Yogyakarta: Kanisius.
- [13] Nitko, A. J., & Brookhart, S. M. (2014). Educational Assessment of Students Sixth Edition. Pearson New International Edition.
- [14] Retnawati, H. (2016). Analisis Kuantitatif Instrumen Penelitian (Panduan Peneliti, Mahapeserta didik, dan Psikometrian). Parama Publishing.
- [15] Sudaryono, dkk. 2013 Pengembangan Instrumen Penelitian Pendidikan. Yogyakarta: Graha Ilmu.
- [16] Sudjino, A. 2011. Pengantar Evaluasi Pendidikan. Jakarta: PT. Raja Grafindo Persada. Suherman. 2018. Pengelolaan Program Ulangan Harian Bersama (Uhb) Sebagai Upaya Meningkatkan Motivasi Belajar Peserta didik. Tadbir Muwahhid ISSN 2579-4876 e-ISSN. 2579-3470 Volume 2 Nomor 2, Oktober 2018
- [17] Tilaar, A.L.F. Hasriyanti. 2019. Analisis Butir Soal Semester Ganjil Mata Pelajaran Matematika pada Sekolah Menengah Pertama. JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia), 8(1), 2019, 57-68 DOI: <http://dx.doi.org/10.15408/jp3i.v8i1.13068>.