

# ANALISIS VALIDITAS, RELIABILITAS, TINGKAT KESULITAN DAN DAYA BEDA BUTIR SOAL UJIAN AKHIR SEMESTER TEMA 7 KELAS III SDN KARET 1 SEPATAN

Ina Magdalena<sup>1</sup>, Septy Nurul Fauziah<sup>2</sup>, Siti Nur Faziah<sup>3</sup>, Fika Sulaehatun Nopus<sup>4</sup>  
Universitas Muhammadiyah Tangerang  
inapgsd@gmail.com , septynurulfauziah2609@gmail.com

## Abstract

*This study aims to (1) determine the quality items for the Fourth Semester Final Exam in the thematic subjects for the 7-year 2020/2021 academic year in terms of validity, reliability, level of difficulty, and distinguishing power, (2) create a program form to calculate validity, reliability, Level of difficulty and distinguishing power. The research method used is quantitative descriptive technique and documentation technique. The data consists of 20 student answer sheets for social class IV for the 2020/2021 academic year. The data were analyzed using several formulas for validity, reliability, difficulty level, power differentiation and the use of IBM SPSS Statistics25 software. The results showed that 16 questions (80%) were valid and 4 questions (20%) were invalid. The reliability of the item is quite good where the value of the reliability coefficient  $r_{11} > r_{table}$  is 0.4438 based on three formulas, namely Anova Hoyt, Cronbach alpha and IBM SPSS Statistics 25. The level of difficulty of the item is not good because the item balance includes easy, medium and difficult media, disproportionately, balance the proportion is 3-5-2 or 3-4-3. The discriminatory power of item items in the good category is 1 question or 5%, the questions in the good category are 9 questions or 45%, the questions in the sufficient category are about 4 questions or 20%, the questions in the poor category are about 5 questions or 25% and Questions that are categorized as less good have 1 question or 5%.*

**Keywords:** *Analysis, Items, Elementary School*

**Abstrak :** Penelitian ini bertujuan untuk (1) menentukan item kualitas Ujian Akhir Semester kelas IV mata pelajaran tematik tema 7 tahun ajaran 2020/2021 dalam hal validitas, reliabilitas, tingkat kesulitan, dan daya pembeda, (2) membuat bentuk program untuk menghitung validitas, reliabilitas, Tingkat kesulitan dan daya pembeda. Metode penelitian yang digunakan adalah teknik deskriptif kuantitatif dan teknik dokumentasi. Datanya terdiri dari 20 lembar jawaban siswa pada kelas sosial IV tahun ajaran 2020/2021. Data-data tersebut dianalisis dengan menggunakan beberapa rumus validitas, reliabilitas, tingkat kesulitan, pembedaan daya dan penggunaan perangkat lunak IBM SPSS Statistics25. Hasil penelitian menunjukkan bahwa 16 pertanyaan (80%) valid dan 4 pertanyaan (20%) tidak valid. Keandalan item cukup bagus dimana nilai koefisien reliabilitas  $r_{11} > r_{tabel}$  adalah 0,4438 berdasarkan tiga rumus yaitu Anova Hoyt, Cronbach alpha dan IBM SPSS Statistics 25. Tingkat kesulitan item tidak baik karena saldo item termasuk media yang mudah, sedang Dan sulit tidak proporsional, keseimbangan proporsinya 3-5-2 atau 3-4-3. Daya pembeda item soal dalam katagori baik sekali 1 soal atau 5%, soal dalam katagori baik ada 9 soal atau 45%, soal dalam katagori cukup ada sekitar 4 soal atau 20%, soal yang berkatagori kurang baik ada sekitar 5 soal atau 25% dan soal yang berkatagori kurang baik ada 1 soal atau 5%.

**Kata Kunci:** Analisis, Butir Soal, Sekolah Dasar

## PENDAHULUAN

Pendidikan merupakan sebagai usaha sadar dan terencana dalam mewujudkan proses pembelajaran pengetahuan, keterampilan, serta kebiasaan yang dilakukan suatu individu untuk mengembangkan potensinya dan mencerdaskan generasi bangsa. Dengan adanya Pendidikan dapat berpengaruh untuk pertumbuhan dan perkembangan manusia. Tanpa adanya Pendidikan manusia akan kebingungan atas perkembangan yang ada dimasyarakat.

Setiap manusia wajib untuk mendapatkan Pendidikan, hal ini jelas menegaskan bahwa Pendidikan memiliki nilai yang sangat penting dalam proses kehidupan. sebagai mana yang di jelaskan dalam UU No.20 Tahun 2003 pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa, dan negara. Berdasarkan undang-undang di atas Sisdiknas menegaskan bahwa peserta Pendidikan harus secara sadar dalam mewujudkan proses Pendidikan yang diinginkan. Peserta pendidikan melaksanakan pendidikan yang didalamnya terdapat proses pembelajaran yang bertujuan agar ia dapat mengembangkan kemampuan dalam dirinya serta kemampuan yang belum pernah ia dapatkan sebelumnya.

Pendidikan itu sendiri terdiri dari beberapa jenjang, sebagaimana yang di tuangkan dalam UU No. 20 tahun 2003 Bab VI Pasal 13 Ayat 1 jalur pendidikan terdiri atas pendidikan formal, non-formal, dan informal yang dapat saling melengkapi dan memperkaya. Pendidikan formal adalah jalur pendidikan yang terstruktur dan berjenjang yang terdiri atas pendidikan dasar, pendidikan menengah, dan pendidikan tinggi. Pendidikan non-formal adalah jalur pendidikan di luar pendidikan formal yang dapat dilaksanakan secara terstruktur dan berjenjang. Pendidikan informal adalah jalur pendidikan keluarga dan lingkungan.

Menurut (Sukardi, 2011) Pendidikan yang berkualitas berhubungan erat dengan proses kegiatan pembelajaran yang dilaksanakan di sekolah. Tiga unsur dalam proses pembelajaran adalah tujuan pengajaran (tujuan instruksional), pengalaman belajar (proses belajar mengajar), dan hasil belajar. Kegiatan

penilaian dalam dunia pendidikan dikenal dengan istilah kegiatan evaluasi. Evaluasi merupakan bagian dari proses belajar mengajar. Evaluasi hasil belajar dilakukan oleh pendidik untuk memantau proses, kemajuan, pencapaian dan perbaikan hasil belajar peserta didik secara berkesinambungan,

Evaluasi hasil belajar yang diadakan pada tiap-tiap mata pelajaran, akhir semester, menjadi sangat penting (urgent) kedudukan dan fungsinya dalam mengukur tingkat kemampuan dan pemahaman siswa. Aktivitas evaluasi sebenarnya harus selalu dilakukan pada saat akhir pelajaran, gunanya untuk menilai sampai seberapa besar tingkat penguasaan ilmu pengetahuan yang diberikan dan diserap siswa. Dalam hal ini, proses persiapan, pembuatan soal, pelaksanaan tes, observasi dan penilaian tes, hendaknya direncanakan secara teratur dan kontinyu sehingga guru dapat benar-benar mengevaluasi dan membimbing perkembangan siswa secara positif sesuai dengan penerapan Kurikulum Tingkat Satuan Pendidikan (KTSP) 2006,2010 maupun Kurikulum 2013.

Dalam kegiatan evaluasi diperlukan alat atau teknik penilaian, sehingga pelaksanaannya akan lebih terarah. Alat evaluasi dalam pendidikan yang digunakan untuk mengumpulkan data dapat berupa tes atau nontes. Berkaitan dengan pemahaman kita terhadap tes, Nurkencana dan Sunartana menyatakan tes adalah suatu cara untuk mengadakan penilaian yang berbentuk suatu tugas atau serangkaian tugas yang harus dikerjakan oleh siswa atau sekelompok siswa sehingga menghasilkan nilai tentang tingkah laku atau prestasi siswa sebagai peserta didik.

Salah satu alat yang digunakan sebagai sarana untuk penilaian hasil belajar yaitu tes. Menurut (sudijono, 2012) tes adalah cara dalam mengukur dan menilai di bidang pendidikan dalam bentuk pemberian tugas atau serangkaian tugas yang harus dikerjakan oleh siswa, sehingga dapat diketahui nilai prestasi siswa. Nilai tersebut kemudian dibandingkan dengan nilai-nilai yang dicapai oleh testee lainnya, atau dibandingkan dengan nilai standar tertentu. Oleh karena itu, untuk mengetahui kualitas butir soal yang digunakan untuk tes perlu dilakukan analisis butir soal.

Analisis butir soal merupakan suatu kegiatan yang harus dilakukan guru untuk meningkatkan kualitas soal yang telah ditulis. Menurut Aiken (1994) dalam Depdiknas (2008), tujuan analisis butir soal yaitu meningkatkan kualitas butir tes dan

mengetahui informasi diagnostik siswa. Soal yang berkualitas yaitu soal yang dapat memberikan informasi setepat tepatnya, sehingga dapat diketahui siswa yang telah menguasai materi dan yang belum.

Suatu tes evaluasi yang baik memiliki ciri dan sifat yang merupakan persyaratan-persyaratan yang harus dipenuhi, yaitu tes tersebut harus valid atau memiliki tingkat validitas yang absah/baik. Sebuah tes evaluasi dikatakan valid apabila tes tersebut secara tepat dan benar dapat mengukur apa yang hendak diukur. Validitas di sini, dapat berupa validitas isi, prediktif atau ramalan dan validitas konstruksi, kemudian tes tersebut harus reliabel, obyektif, praktis dan ekonomis.

Dalam evaluasi pendidikan baik tes maupun nontes, keduanya merupakan instrumen atau alat bantu pengumpulan dan pengolahan data tentang variabel-variabel yang diteliti. Ciri-ciri/karakteristik instrument yang baik sebagai alat evaluasi adalah memenuhi persyaratan validitas dan reliabilitas. Inilah alasan mengapa alat evaluasi yang baik dapat dilihat dari beberapa segi antara lain: (1) validitas, (2) reliabilitas, (3) objektivitas, (4) praktikabilitas, (5) daya pembeda, (6) taraf atau derajat kesukaran, (7) efektivitas option, (8) efisiensi

Soal ulangan akhir semester (UAS) menjadi penting ketika penulis menemukan sebuah pernyataan bahwa UAS termasuk dalam kategori tes sumatif. Menurut (Widoyoko, 2014) fungsi tes sumatif yaitu untuk mengetahui sejauh mana pencapaian kompetensi siswa dalam mata pelajaran tertentu. Hasil tes sumatif siswa, selanjutnya dibandingkan dengan tujuan yang seharusnya dicapai atau kriteria ketuntasan minimal (KKM). Oleh karena itu, nilai tes sumatif dijadikan sebagai catatan kemajuan belajar siswa dan penentu seorang siswa dapat atau tidaknya menerima program berikutnya.

Berdasarkan uraian di atas, maka perlu dilakukan analisis butir soal secara kuantitatif untuk mengetahui kualitas soal tersebut. Kualitas soal UAS dapat dilihat dari hasil uji validitas, reliabilitas, tingkat kesukaran dan daya beda. Peneliti tertarik untuk melakukan penelitian dengan judul “Analisis Validitas dan Reliabilitas Tes Soal Butir Soal UAS Tema 7 Kelas III SDN Karet 1 Sepatan Kabupaten Tangerang Tahun 2020/2021”.

Penelitian ini bertujuan untuk mengetahui validitas reliabilitas tingkat kesukaran dan daya beda butir soal Ujian Akhir Semester Ganjil TEMA 7 kelas III IPS di SDN Sepatan 1 tahun ajaran 2020/2021. Penelitian ini bertujuan untuk (1) mengetahui kualitas butir soal ujian akhir semester ganjil kelas III mata pelajaran TEMA 7 tahun ajaran 2020/2021 ditinjau dari validitas, reliabilitas, tingkat kesukaran, dan daya beda, (2) membuat bentuk sebuah program untuk menghitung validitas, reliabilitas, tingkat kesukaran dan daya beda.

## METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif dan Teknik dokumentasi. Data penelitian berupa data lembar soal dan lembar jawaban uas tematik tema 7 tahun ajaran 2020/2021 sebanyak 20 orang. Data di analisis berupa data validitas, reliabilitas, tingkat kesukaran dan daya beda serta menggunakan software IBM SPSS Statistik 25. Bentuk tes yang dianalisis adalah pilihan ganda. Validitas adalah indeks yang menunjukkan sejauh mana suatu alat pengukur itu mengukur apa yang ingin di ukur. Menurut Sodian Effendi (2011). Analisis validitas dilakukan dengan menggunakan teknik korelasi point biserial yaitu:

$$r_{pbi} = \frac{m_{\bar{p}} - m_{\bar{q}}}{SD_{\bar{p}}} \sqrt{\frac{p}{q}}$$

rumus lainnya dengan menggunakan korelasi product moment

$$r_{xy} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{[n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2] [n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2]}}$$

nilai  $r_{xy}$  atau  $r_{pbi}$  akan dibandingkan dengan koefisien *table* nilai “r” *product moment* pada taraf signifikan 5%. Apabila nilai  $r_{xy}$  atau  $r_{pbi}$  hasil koefisien korelasi lebih besar ( $>$ ) dari nilai  $r_{table}$ , maka hasil yang diperoleh adalah signifikan, artinya butir soal tes dinyatakan valid. Nilai  $r_{xy}$  adalah nilai koefisien korelasi dari setiap butir/ item soal sebelum dikoreksi. Kemudian dicari corrected item-total correlation coefficient dengan rumus sebagai berikut.

$$R_{x(y-1)} = \frac{r_{xy} s_y - s_x}{\sqrt{s_y^2 + s_x^2 - 2r_{xy} (s_y) (s_x)}}$$

Nilai  $r_{x(y-1)}$  akan dibandingkan dengan koefisien korelasi tabel  $r_{\text{tabel}} = r_{(a,n-2)}$ . Jika  $r_{x(y-1)} \geq r_{\text{tabel}}$ , maka instrument valid. Pada output SPSS, *corrected item-total correlation coefficient*  $t \geq r_{\text{tabel}}$ , maka instrument valid. Reliabilitas adalah indeks yang menunjukkan sejauh mana suatu alat pengukuran dapat dipercaya atau dapat diandalkan. Suatu intrumen pengukuran dikatakan reliabel jika pengukurannya konsisten, cermat dan akurat. (Sofian Effendi dan Masri Singarimbun, 2011) Rumus-rumus yang digunakan pada pengujian reliabilitas sebagai berikut :

- 1) Rumus Speaman Brown:

$$r_i = \frac{2rb}{1+rb}$$

- 2) Rumus KR 20 (Kuder Richardson)

$$r_i = \frac{k}{k-1} \left\{ \frac{st^2 - \sum p_i \cdot q_i}{st^2} \right\}$$

- 3) Rumus KR 21 (Kuder Richardson)

$$r_i = \frac{k}{(k-1)} \left\{ 1 - \frac{m(k-m)}{k \cdot st^2} \right\}$$

- 4) Rumus Rulon

$$r_{11} = 1 - \frac{s_d^2}{s_t^2}$$

- 5) Flanagan

$$r_{11} = 2 \left( 1 - \frac{s_1^2 + s_2^2}{s_t^2} \right)$$

- 6) Analisis Varians Honyt (Anova Hoyt)

$$r_i = 1 - \frac{MKe}{MKs}$$

- 7) Alfa Cronbach

$$r_i = \frac{k}{(k-1)} \left\{ 1 - \frac{\sum s_i^2}{st^2} \right\}$$

nilai koefisien ( $r_i$ ) akan dibandingkan dengan korelasi tabel  $r_{\text{tabel}} = r_{(a,n-2)}$ . Jika  $r_{11} > r_{\text{tabel}}$ , maka intrumen reliabel. Pada output SPSS, jika *Cronbach's Alpha*  $> r_{\text{tabel}}$ , maka intrumen reliabel. Analisis tingkat kesukaran soal artinya mengkaji soal-soal tes dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk mudah, sedang dan sukar. Menurut (Nana Sudjana, 2009) asumsi yang digunakan untuk memperoleh kualitas soal yang baik, disamping memenuhi validitas dan reliabilitas adanya keseimbangan dari tingkat kesulitan soal tersebut. Keseimbangan adalah

ananya soal – soal yang termasuk mudah, sedang dan sukar secara proporsional. Perbandingan antara soal mudah-sedang-sukar bisa dibuat 3-4-3, artinya 30% soal katagori mudah, 40% soal katagori sedang dan 30% soal katagori sukar. Perbandingan lain yang termasuk sejenis dengan proposi di atas adalah 3-5-2. Artinya 30% soal katagori mudah, 50% soal katagori sedang dan 20% soal katagori sukar. (Novalia dan Muhammad Syazali, 2014) Rumus yang digunakan untuk mengetahui tingkat/indek kesukaran dari tiap butir soal yaitu :

$$I = \frac{B}{J}$$

Penafsiran tingkat kesukaran butir tes digunakan kriteria menurut *Witherington* dalam Anas sudijono sebagai berikut.

**Tabel 1: Interpretasi tingkat kesukaran butir tes**

Interval	Interprestasi
0,00 – 0,30	Soal Sukar
0,31 – 0,70	Soal Sedang
0,71 – 1,00	Soal Mudah

Cara yang lain yang dapat ditempuh untuk menafsirkan analisis tingkat kesukaran butir soal dalam anates adalah :

**Tabel 2 : Interpretasi tingkat kesukaran anates**

Interval	Interprestasi
0% - 15%	Sangat Sukar
16% - 30%	Sukar
31% - 70%	Sedang
71% - 85%	Mudah
86% - 100%	Sangat Mudah

Menurut (Novalia dan Muhammad Syazali, 2014) menganalisis daya pembeda artinya mengkaji soal-soal tes dari segi keangguapan tes tersebut dalam membedakan siswa yang termasuk kedalam katagori lemah/rendah dan katagori kuat/tinggi presentasinya. Penentuan daya beda butir soal pada Anates dapat diketahui dalam tabel daya pembeda pada kolom DP persen. Butir soal yang memiliki indeks daya beda  $\geq 0,30$  dinyatakan baik dan butir soal yang indeks daya beda  $< 0,30$  dinyatakan tidak baik. Daya pembeda butir soal memiliki manfaat yaitu untuk meningkatkan mutu setiap-butir soal melalui data empiriknya dan untuk mengetahui seberapa jauh masing-masing butir soal dapat membedakan kemampuan siswa, yaitu siswa yang telah memahami atau belum memahami materi yang diajarkan pendidik. Menurut (Kusaeri Supranto, 2012), Adapun untuk menghitung daya beda digunakan rumus sebagai berikut:

$$D = P_A - P_B$$

Sebagai catatan, dalam keadaan di mana jumlah testee adalah cukup besar (100 orang atau lebih). Daya pembeda item cukup dihitung berdasarkan 27% testee kelompok atas dan 27% dari testee kelompok bawah, sedangkan testee yang terletak diantara dua ujung ekstrem itu tidak perlu diikut sertakan dalam perhitungan analisis menurut (Sujono Anas , 2013).

**Tabel 3 : Interpretasi Indeks Daya Pembeda Butir**

<b>Daya Pembeda</b>	<b>klasifikasi</b>	<b>intreprestasi</b>
0,70 – 1,00	Excellent	Baik Sekali
0,40 – 0,69	Good (baik)	Baik
0,20 – 0,39	Satisfactory (memuaskan)	cukup
0,00 – 0,19	Poor (lemah)	Kurang Baik
Bertanda Negatif	-	Jelek Sekali



## HASIL PENELITIAN DAN PEMBAHASAN

Hasil data yang di peroleh peneliti berupa soal dan jawaban UAS Tematik Tema 7 siswa mata pelajaran Matematika Kelas III SDN Karet 1 Sepatan Kabupaten Tangerang Tahun 2020/2021.

**Tabel 4 : Data nilai peserta didik kelas III**

Jawaban Benar	Nilai	Banyak Siswa
5	25	4
6	30	1
7	35	3
8	40	1
9	45	2
10	50	1
12	60	4
14	70	2
16	80	1
17	85	1

Berdasarkan tabel tersebut, nilai maksimum atau nilai tertinggi didapatkan oleh 2 orang siswa yang mendapatkan nilai 90. Nilai minimum atau nilai terendah didapatkan oleh 3 orang siswa yang mendapat nilai 40. *Mean* atau rata-rata skor yang diperoleh untuk mata-mata pelajaran matematika adalah  $X = \frac{\sum_{i=1}^n x_i}{\sum f_i} = \frac{960}{20} = 48$ . Jadi mean atau rata-rata skor yang diperoleh untuk mata pelajaran tema 7 adalah 48. *Median* atau nilai tengah dari hasil ujian akhir semester adalah 45. *Modus* atau nilai sering muncul didapatkan oleh 4 siswa dengan nilai 25 dan 60. Jumlah siswa yang memperoleh nilai lebih besar dari *mean* adalah 1 orang dengan nilai 85.

Validitas adalah derajat yang menunjukkan suatu tes mengukur apa yang hendak diukur. Sebuah data atau informasi dapat dikatakan valid apabila sesuai dengan keadaan sebenarnya. Analisis validitas butir soal dilakukan pada soal Tematik



Rebialitas tes adalah ketetapan alat dalam menilai apa yang dinilainya. Reliabilitas adalah indeks yang menunjukana sejauh mana suatu alat pengukur dapat dipercaya. Analisis reliabilitas penelitian ini menggunakan rumus *software* IBM SPSS Statistics 25. Berhasilan dari perhitungan rumus *software* IBM SPSS Statistics 25 mempunyai indek reliabilitas yaitu 0,4438.

### Case Processing Summary

		N	%
Cases	Valid	20	100.0
	Excluded <sup>a</sup>	0	.0
	Total	20	100.0

a. Listwise deletion based on all variables in the procedure.

### Reliability Statistics

Cronbach's Alpha	Part 1	Value	.761
		N of Items	6 <sup>a</sup>
	Part 2	Value	.737
		N of Items	6 <sup>b</sup>
Total N of Items			12
Correlation Between Forms			.767
Spearman-Brown Coefficient	Equal Length		.868
	Unequal Length		.868
Guttman Split-Half Coefficient			.868

a. The items are: soal1, soal2, soal3, soal4, soal5, soal6.

b. The items are: soal7, soal8, soal10, soal16, soal18, soal19.

Mengutip dari jurnal (Heri Susanto, Achi Rinaldi dan Novalia, 2015) Formula *spearman-Brown* dilakukan untuk menentukan raliabilitas tes hasil belajar bentuk objektif dengan cara membelas tes menjadi 2 bagian, dengan menggunakan dua model yaitu Model Gasal Genap dan Model Kiri-Kanan, sedangkan formula Flanagan dan Rulon sama-sama menggunakan Teknik belah dua, namun sasaran yang dijadikan landasan

berpijak dalam penentuan reliabilitas tersebut berbeda. Dengan formula Spearman-Brown, pengujian dan penentuan reabilitas tes hasil belajar bentuk objektif didasarkan pada korelasi antara belahan pertama dan belahan kedua. Formula Flanagan menentukan reliabilitas tes hasil belajar tidak berdasarkan korelasi belah pertama dan belah kedua, melainkan pada deviasi (selisih skor dengan rata-rata totalnya). Pada formulasi Rulon, penentuan reliabilitas tes hasil belajar bentuk objektif didasarkan pada selisih skor yang dimiliki oleh belahan pertama dengan belahan kedua tes tersebut.

Berbeda dengan formula sebelumnya, formula Kuder-Richardson dan formula C. Hoyt tidak menggunakan Teknik belah dua, penelitian reliabilitas dengan formula Kuder-Richarson dilakukan menganalisis skor-skor item tes hasil belajar yang bersangkutan secara langsung. Formula C.Hoyt menentukan reliabilitas tes hendaknya kita menganggap bahwa data yang berupa skor-skor hasil tes itu kita anggap sebagai data hasil eksperimen, factor pertama adalah subjektif dan factor kedua adalah item. Selanjutnya kit acari interaksi antara testee (subjek yang menjawab item) dengan item hasil belajar itu sendiri (yang dijawab oleh subjek), Teknik ini yang lebih dikenal dengan istilah Teknik Analisis Varian (ANAVA). Namun kelemahannya yaitu perhitungan harus lebih teliti dan memerlukan waktu yang lebih lama dibandingkan dengan rumus alpha Cronbach.

Kelebihan rumus pertama ( $KR_{20}$ ) adalah perhitungannya lebih teliti dari pada rumus kedua ( $KR_{11}$ ), namun kelemahannya adalah proses perhitungannya lebih sulit. Ada pun kelebihan rumus kedua ( $KR_{11}$ ) adalah perhitungan relative lebih sederhana atau lebih mudah, sedangkan kelemahannya adalah hasil perhitungannya kurang teliti. Dalam pendapat di atas peneliti dapat menyimpulkan bahwa ketujuh rumus tersebut punya kelebihan masing-masing dalam penentuan reliabilitas butir soal. Hal ini ditunjukkan oleh proses pengerjaan dengan menggunakan rumus diatas tidak terlalu sulit, waktu pengerjaan tidak terlalu singkat, dan semua rumus mempunyai derajat reliabilitas yang berbeda-beda, selain dari rumus  $r_6$  dan  $r_7$ . Akan tetapi rumus  $r_7$  lebih baik digunakan karena proses perhitungannya lebih sederhana, lebih mudah untuk di pahami dan hasilnya yang sama dengan  $r_6$ . Selain itu kemungkinan kesalahan perhitungan sangat sedikit dari pada rumus lainnya. Rumus  $r_7$  dapat digunakan untuk

soal tes objektif maupun subjektif dan hasilnya sama dengan penggunaan *Software* IBM SPSS 25.

Analisis tingkat kesukaran artinya mengkaji soal-soal tes dari segi kesulitannya, sehingga dapat diperoleh soal-soal yang layak untuk diberikan kepada peserta didik, untuk mengetahui mana soal yang mudah, mana soal yang sedang dan mana soal yang sulit. Tingkat kesukaran soal dapat dilihat dari seberapa banyak siswa tersebut dapat menjawabnya, bukan dilihat dari sudut guru yang membuat soal. Analisis tingkat kesukaran pada penelitian ini menggunakan satu rumus. Hasil analisis Tingkat Kesukaran yaitu:

**Tabel 5 : Tingkat Kesukaran**

Soal	$p = \frac{\text{jumlah jawaban benar}}{\text{banyak siswa}}$	
	Indeks	Kriteria
Soal 1	0,70	Soal sedang
Soal 2	0,35	Soal sedang
Soal 3	0,70	Soal sedang
Soal 4	0,45	Soal sedang
Soal 5	0,50	Soal sedang
Soal 6	0,30	Soal sukar
Soal 7	0,40	Soal sedang
Soal 8	0,55	Soal sedang
Soal 9	0,70	Soal sedang
Soal 10	0,25	Soal sukar
Soal 11	0,45	Soal sedang
Soal 12	0,65	Soal sedang
Soal 13	0,10	Soal sukar
Soal 14	0,60	Soal sedang
Soal 15	0,40	Soal sedang
Soal 16	0,60	Soal sedang
Soal 17	0,50	Soal sedang
Soal 18	0,55	Soal sedang
Soal 19	0,65	Soal sedang
Soal 20	0,30	Soal sukar

Peneliti menyimpulkan dari hasil perhitungan rumus terhadap 20 butir soal item hasil belajar terdapat 16 butir soal (80%) dalam katagori sedang, dan terdapat 4 butir soal (20%) dalam katagori sukar. Butir soal yang berkatagori sedang ( tidak terlalu sukar tidak terlalu mudah) terdapat pada nomor 1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19. Sedangkan butir soal yang terdapat dalam katagori sukar terdapat pada nomer 6, 10, 13, 20. Rumus ini mempunyai pendekatan dengan cara mencari presentase siswa yang menjawab salah. Artinya semakin siswa menjawab salah maka soal itu terlihat tingkat kesukaran butir soal semakin sukar.

Dapat diartikan dari pendapat di atas bahwa butir soal UAS TEMATIK tema 7 bahwa butir soal tema 7 yang bentuknya pilihan ganda memiliki kualitas soal yang baik, disamping memenuhi validitas dan reliabilitas adlah adanya keseimbangan dari tingkat kesulitan soal tersebut. Keseimbangan yang dimaksud adalah jumlah soal-soal yang tergolong mudah, sedang dan sukar seimbang. Tingkat kesukaran soal perlu dilihat dari kemampuan siswa dalam menjawab soal yang diberikan, bukan dilihat dari sudut pandang guru yang yang membuat soal. Perbandingan antara soal yang mudah, sedang dan sukar dapat dibuat 3-4-3 atau 3-5-2, yang diartikan adalah 30% soal berkatagori mudah, 40% soal yang berkatagori sedang dan 30% soal bekatagori sukar atau 30% soal berkatagori mudah, 50% soal berkatagori sedang dan 20% soal berkatagori sukar.

Mengutip dari jurnal (Heri Susanto, Achi Rinaldi dan Novalia, 2015) Butir-butir soal yang termasuk dalam katagori sedang sebaiknya sebagai dicatat dalam buku bank soal. Selanjutnya, butir-butir soal tersebut dapat dikeluarkan lagi dalam tes hasil belajar pada waktu yang akan datang. Untuk butir soal yang sukar kemungkinan ada tiga tindak lanjut yaitu: (1) butir soal tersebut tidak lagi dipakai dan tidak dikeluarkan lagi dalam hasil tes berikutnya. (2) dikaji ulang, diteliti dan ditelusuri sehingga dapat diketahui factor penyebab anak tidak dapat menjawab soal tersebut. (3) butir soal yang sukar sebaiknya memiliki manfaat yaitu butir soal dapat dimanfaatkan pada tes seleksi yang bersifat longgar, dalam artian bahwa Sebagian besar dari testee akan dinyatakan lulus dalam tes seleksi tersebut. Dalam kondisi ini, pemberian butir soal yang mudah akan memberikan kesempatan bagi banyak testee untuk lolos dalam tes seleksi atau ujian yang diadakan.

Analisis daya beda artinya mengkaji soal-soal tes dari segi kesanggupan tes tersebut dalam membedakan siswa yang termasuk kedalam kategori rendah dan kategori tinggi. Daya pembeda butir adalah kemampuan suatu butir tes untuk dapat membedakan antara testee yang berkemampuan tinggi dan berkemampuan rendah. Analisis daya beda pada penelitian ini menggunakan rumus. Hasil analisis daya beda yaitu:

**Tabel 6 : Daya Beda**

Soal	D = PA – PB	
	Indeks	Kriteria
Soal 1	0,60	BAIK
Soal 2	0,30	CUKUP
Soal 3	0,40	BAIK
Soal 4	0,70	BAIK SEKALI
Soal 5	0,40	BAIK
Soal 6	0,60	BAIK
Soal 7	0,60	BAIK
Soal 8	0,50	BAIK
Soal 9	0,20	CUKUP
Soal 10	0,30	CUKUP
<b>D = PA – PB</b>		
SOAL	INDEKS	KRITERIA
Soal 11	0,10	KURANG BAIK
Soal 12	0,10	KURANG BAIK
Soal 13	0,00	KURANG BAIK
Soal 14	0,20	CUKUP
Soal 15	0,00	KURANG BAIK
Soal 16	0,60	BAIK
Soal 17	-0,20	JELEK SEKALI
Soal 18	0,50	BAIK
Soal 19	0,50	BAIK
Soal 20	0,00	KURANG BAIK

Berdasarkan hasil analisis dengan menggunakan rumus, maka peneliti memperoleh bahwa soal dalam katagori baik sekali 1 soal atau 5%, soal dalam katagori baik ada 9 soal atau 45%, soal dalam katagori cukup ada sekitar 4 soal atau 20%, soal yang berkatagori kurang baik ada sekitar 5 soal atau 25% dan soal yang berkatagori kurang baik ada 1 soal atau 5%. Mengutip dari jurnal (Heri Susanto, Achi Rinaldi dan Novalia, 2015) Adapun tindak lanjut atas hasil menganalisis mengenai daya pembeda item tes hasil belajar tersebut adalah, butir-butir soal tersebut dapat dikeluarkan lagi dalam tes hasil belajar pada waktu yang akan datang. Untuk butir soal yang sukar kemungkinan ada tiga tindak lanjut yaitu: (1) butir soal tersebut tidak lagi dipakai dan tidak dikeluarkan lagi dalam hasil tes berikutnya. (2) dikaji ulang, diteliti dan ditelusuri sehingga dapat diketahui factor penyebab anak tidak dapat menjawab soal tersebut. (3) butir soal yang sukar sebaiknya memiliki manfaat yaitu butir soal dapat dimanfaatkan pada tes seleksi yang bersifat longgar, dalam artian bahwa Sebagian besar dari testee akan dinyatakan lulus dalam tes seleksi tersebut. Dalam kondisi ini, pemberian butir soal yang mudah akan memberikan kesempatan bagi banyak testee untuk lolos dalam tes seleksi atau ujian yang diadakan.

## **KESIMPULAN**

Berdasarkan hasil penelitian dan pembahasan, dapat diperoleh kesimpulan yaitu terdapat 16 butir soal yang valid (80%) dan 4 butir soal (20%) tidak valid dari 20 soal tes bentuk objektif. Reliabilitas tergolong baik berdasarkan perhitungan menggunakan *software*. Tingkat kesukaran belum cukup baik karena tidak ada keseimbangan soal dari mudah, sedang dan sukar. Seharusnya persinya 3-5-2. Dalam daya beda butir soal memperoleh bahwa soal dalam katagori baik sekali 1 soal atau 5%, soal dalam katagori baik ada 9 soal atau 45%, soal dalam katagori cukup ada sekitar 4 soal atau 20%, soal yang berkatagori kurang baik ada sekitar 5 soal atau 25% dan soal yang berkatagori kurang baik ada 1 soal atau 5%.



## DAFTAR PUTAKA

- Heri Susanto, Achi Rinaldi dan Novalia. (2015). Analisis Validitas Reabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir . *Al-Jabar: Jurnal Pendidikan Matematika*, 203-217.
- Kusaeri Supranto. (2012). *Pengukuran dan Penilaian Pendidikan*. Yogyakarta : Graha Ilmu.
- Nana Sudjana. (2009). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: PT. Remaja Rosdakarya.
- Novalia dan Muhammad Syazali. (2014). *Olah Data Penelitian Pendidikan*. Bandar Lampung: AURA.
- Sofian Effendi dan Masri Singarimbun. (2011). *Metode Penelitian Survie*. Jakarta: LP3ES.
- Sudijono. (2012). *pengantar evaluasi pendidikan*. jakarta: Pt. Raja Grafindo Persada.
- Sujono Anas . (2013). *Pengantar Evaluasi Pendidikan*. Bandung : Alfa Beta.
- Sukardi. (2011). *Evaluasi Pendidikan: Prinsip dan Oprasionalnya*. Jakarta: Bumi Aksara.
- Widoyoko. (2014). *Penilaian Hasil Belajar di Sekolah*. Yogyakarta: Pustaka Pelajar.