



**SOAL PILIHAN GANDA YANG VALID DAN RELIABEL UNTUK
MENGINTERPRETASIKAN KETERAMPILAN BERPIKIR SISWA SMP**

***THE VALID AND RELIABLE TEST OF MULTIPLE CHOICE TO INTERPRET THINKING
SKILLS FROM MIDDLE SCHOOLS STUDENTS***

Arifin Riadi, Noor Laila Atini

Program Studi Pendidikan Matematika STKIP PGRI Banjarmasin

arifin.riadi@stkipbjm.ac.id, noorlailaatini@gmail.com

Abstrak: penelitian ini bertujuan untuk memperoleh soal pilihan ganda yang baik, ditinjau dari valid dan reliabel. Uji validitas yang digunakan dalam penelitian ini adalah uji korelasi pearson, dan uji reliabilitasnya menggunakan alpha cronbach. Draf soal yang dibuat berjumlah 24 buah, dengan proporsi masing-masing tingkat keterampilan berpikir diwakili oleh 4 soal. Tingkat keterampilan berpikir yang dimaksud adalah berdasarkan taksonomi Bloom versi revisi yaitu Mengingat, Memahami, Mengaplikasikan, Menganalisis, Mengevaluasi, dan Mencipta. Setiap soal memiliki empat pilihan jawaban sesuai dengan ketentuan untuk siswa SMP. Dari hasil uji coba dan analisis data diperoleh bahwa soal yang valid dan reliabel berjumlah 16 buah, dengan skor reliabilitas secara keseluruhan adalah 0,602. Hasil penelitian ini diharapkan mampu digunakan untuk memperoleh gambaran mengenai tingkat keterampilan berpikir siswa SMP di sekolah maupun di daerah lain.

Kata Kunci: soal pilihan ganda, valid, reliabel, tingkat keterampilan berpikir

Abstract: This study aims to obtain good multiple choice questions, viewed from valid and reliable. Validity test used in this study is Pearson correlation test, and reliability testing using Cronbach alpha. The draft of the questions was 24, with the proportion of each level of thinking skill represented by 4 questions. The level of thinking skill in question is based on the revised version of Bloom's taxonomy namely Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating. Each question has four answer choices according to the provisions for junior high school students. From the results of the trial and data analysis, it was found that there were 16 valid and reliable questions, with an overall reliability score of 0.602. The results of this study are expected to be used to obtain an overview of the level of thinking skills of junior high school students in schools and in other regions.

Keywords: multiple choice test, valid, reliable, thinking skills level

Cara Sitasi: Riadi, A., & Atini, N.L. (2018). Soal pilihan ganda yang valid dan reliabel untuk menginterpretasikan keterampilan berpikir siswa SMP. *Math Didactic: Jurnal Pendidikan Matematika*, 4(3), 220-225.

Penilaian adalah proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar siswa (Presiden Republik Indonesia, 2013, hal. 5). Penilaian hasil belajar siswa pada jenjang pendidikan dasar dan menengah didasarkan pada prinsip-prinsip dari Permendikbud No. 66 Tahun 2013 tentang Standar Penilaian (Kementerian Pendidikan dan Kebudayaan, 2013, hal. 3), yaitu objektif, terpadu, ekonomis, transparan, akuntabel, dan edukatif. Permendikbud No. 66 Tahun 2013 (Kementerian Pendidikan dan

pengembangan soal tes baik tes bentuk pilihan ganda maupun uraian terdiri atas: (1) penentuan tujuan tes; (2) penyusunan kisi-kisi; (3) penulisan soal; (4) penelaahan soal (review dan revisi soal); (5) uji coba soal, termasuk analisis dan perbaikan; dan (6) perakitan soal menjadi perangkat tes. Lebih jauh, Kartowagiran (2012, hal. 3) menjelaskan bahwa pada Taksonomi Bloom dapat dibuat indikator berdasarkan kata kerja yang sesuai untuk setiap tingkatan HOTS seperti pada Tabel 1.

Tabel 1. Kata Kerja Taksonomi Bloom

Mengingat	Memahami	Mengaplikasikan	Menganalisis	Mengevaluasi	Mencipta
Menyebutkan	Membedakan	Menggunakan	Membandingkan konsep	Mengkritik	Mengkreasi
Menghafal	Membandingkan	Menerapkan	Memprediksikan	Memberi keputusan	Merancang
Mengidentifikasi	Menjelaskan	Membentuk	Mengorganisasikan informasi	Merekomendasikan	Menyintesis
Menuliskan	Mengilustrasikan	Mengatur	Mengupas	Menyetujui pendapat	Membuat proposal
Menunjukkan	Menduga	Menghitung	Memeriksa hubungan	Mendukung	Membangun
	Menguraikan	Menentukan	Menyimpulkan	Memilih	Mengompilasi
	Menghubungkan	Menyelesaikan		Membenarkan	Membuat estimasi
	Menerjemahkan	Melakukan eksperimen		Menyangkal	Memodifikasi
	Meringkas	Mendemonstrasikan		Memprioritaskan	Mengembangkan
	Menggolongkan			Menilai	

Kebudayaan, 2013, hal. 4) juga menjelaskan bahwa pendidik menilai kompetensi pengetahuan melalui tes tulis, tes lisan, dan penugasan.

Gronlund, Linn, & Miller (2009, hal. 139-169), menyatakan bahwa dalam perencanaan tes melibatkan hal-hal sebagai berikut.

- Menentukan tujuan pengukuran.
- Mengembangkan satu set spesifikasi/acuan.
- Memilih jenis soal tes yang tepat.
- Menyiapkan satu set item yang relevan.

Alat ukur HOTS dalam penelitian ini menggunakan tes tertulis bentuk pilihan ganda. Menurut Kartowagiran (2012, hal. 3),

ganda lebih baik daripada tes uraian dalam hal banyaknya jumlah pokok bahasan yang ditanyakan, lebih menekankan kepada keluasan materi, serta mudah dan konsisten dalam hal penskoran (Kartowagiran, 2012, hal. 13). Sedangkan menurut Reynolds, Livingston, & Willson (2009, hal. 224) tes uraian memiliki kelebihan yaitu menyediakan kebebasan kepada siswa untuk menuliskan jawaban mereka, dapat memberikan pengalaman baru serta proses kognitif peserta didik yang kreatif. Tambahkan lagi, tes uraian yang baik menantang peserta didik untuk mengorganisasi, menganalisis, mengintegrasikan dan menyintesis informasi yang ada.

Metode Penelitian

Instrumen penelitian yang digunakan adalah tes tertulis berbentuk pilihan ganda. Tes pilihan ganda ini dirancang sesuai Taksonomi Bloom versi revisi yang terdiri atas 6 tingkat (mengingat, memahami, mengaplikasikan, menganalisis, mengevaluasi, dan mencipta). Tes ini akan dibuat sebanyak 24 soal, dengan proporsi masing-masing tingkat berpikir adalah 4 soal. Banyaknya pilihan pada setiap soal adalah 4, sesuai dengan Standar Nasional untuk SMP.

Instrumen penelitian ini terlebih dahulu dirancang dengan prosedur yang akan

menjamin keabsahan data yang diambil. Prosedur yang dilakukan sebagai berikut.

- Melakukan telaah terhadap Kompetensi Dasar (KD) yang akan diukur
- Membuat Indikator yang merupakan turunan dari KD, dengan berpedoman pada kata kerja untuk Taksonomi Bloom seperti pada Tabel 1.
- Merancang draf soal
- Mengecek validitas draf soal menurut ahli materi dan ahli evaluasi
- Melakukan revisi sesuai masukan validator

Tabel 2. Contoh Lembar Validasi Butir Soal Pilihan Ganda

No.	Aspek yang divalidasi	Penilaian	
		Ya	Tidak
Aspek materi			
1	Soal sesuai dengan indikator		
2	Distraktor berfungsi		
3	Hanya ada satu jawaban yang paling tepat		
Aspek konstruksi			
4	Pokok soal dirumuskan dengan singkat, jelas, dan tegas		
5	Rumusan pokok soal dan pilihan jawaban merupakan pertanyaan yang diperlukan		
6	Pokok soal tidak memberi petunjuk ke kunci jawaban		
7	Pokok soal bebas dari pernyataan yang bersifat negatif ganda		
8	Gambar, grafik, tabel, diagram, wacana, dan sejenisnya yang terdapat pada soal jelas dan berfungsi		
9	Panjang pilihan jawaban relatif sama		
10	Pilihan jawaban tidak menggunakan pernyataan “semua jawaban di atas salah” atau “semua pilihan jawaban di atas benar” dan sejenisnya		
11	Pilihan jawaban yang berbentuk angka atau waktu harus disusun berdasarkan urutan besar kecilnya angka tersebut atau kronologisnya		
12	Butir-butir soal tidak bergantung pada jawaban soal sebelumnya		
Aspek bahasa/budaya			
13	Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia		
14	Menggunakan bahasa yang komunikatif		
15	Tidak menggunakan bahasa yang berlaku setempat (bias budaya)		
16	Pilihan jawaban tidak mengulang kata/ kelompok kata yang sama		

- f. Melakukan uji coba pada siswa di sekolah yang bukan bagian dari sampel penelitian
- g. Melakukan Analisis Butir Soal dari hasil uji coba tersebut
- h. Hasil uji coba juga digunakan untuk mengetahui apakah waktu yang disediakan sebanyak 2X40 menit cukup untuk menjawab soal
- i. Melakukan perbaikan sesuai hasil uji coba
- j. Memperoleh soal tes pilihan ganda versi final

1. Validitas Tes

Uji validitas yang digunakan adalah validitas isi dan validitas konstruk. Validitas isi dihitung melalui validasi ahli menggunakan lembar validasi yang diambil dari dua orang ahli, yaitu ahli materi dan ahli evaluasi. Berdasarkan CVR (*Content Validity Ratio*) yang dikemukakan oleh Lawshe (1975, hal. 568) bahwa jika ahli yang digunakan berjumlah 2 orang maka keduanya harus menemui kesepakatan bahwa tiap item tes valid, jika tidak valid maka harus dilakukan perbaikan. Adapun lembar validasi yang digunakan mengacu pada Tim Pusiasjian (Kartowagiran, 2012, hal. 11-12) sebagaimana terlihat pada Tabel 2.

Untuk validitas konstruk, peneliti menggunakan bantuan aplikasi SPSS versi 23. Uji validitas yang digunakan adalah validitas Product Moment Pearson Correlation dengan prinsip mengorelasikan atau menghubungkan antara masing-masing skor item dengan skor total yang diperoleh. Langkah-langkah melakukan uji validitas dengan SPSS ini adalah sebagai berikut.

- 1) Menyusun tampilan pada Data View menjadi seperti ini:

Soal_1	Soal_2	...	Soal_24	Total
(Skor siswa 1)	(Skor siswa 1)	(Skor siswa 1)	(Skor siswa 1)	(Total skor siswa 1)
(Skor siswa 2)	(Skor siswa 2)	(Skor siswa 2)	(Skor siswa 2)	(Total skor siswa 2)
(...)	(...)	(...)	(...)	(...)
(Skor siswa 384)	(Skor siswa 384)	(Skor siswa 384)	(Skor siswa 384)	(Total skor siswa 384)

- 2) Perlu diketahui bahwa untuk alasan kemudahan dalam melakukan pengelompokan gender, peneliti menyusun skor siswa berdasarkan gender sehingga semua siswa laki-laki diisi terlebih dahulu pada tabel Data View baru kemudian data semua skor siswa perempuan
- 3) Memilih menu Analyze, kemudian sub menu Correlate, lalu Bivariate
- 4) Pada kotak dialog Bivariate Correlations, masukkan semua variabel ke kotak Variables. Pada bagian Correlation Coefficients, centang Pearson. Pada bagian Test of Significance, pilih Two-tailed. Centang Flag Significant Correlations lalu klik OK untuk menampilkan Output hasilnya.
- 5) Pada Output tabel Correlations yang muncul, yang perlu diperhatikan adalah kolom Total untuk baris Sig. (*significance*) pada masing-masing item soal. Jika nilai signifikansi yang muncul kurang dari derajat signifikansi yang ditentukan yaitu 0,05 maka soal tersebut valid, sebaliknya jika nilai signifikansinya

0,05 atau lebih dari itu maka soal tersebut tidak valid.

2. Reliabilitas Tes

Untuk uji estimasi reliabilitas, peneliti menggunakan formula Alpha Cronbach melalui aplikasi SPSS versi 23. Adapun langkah-langkah yang harus dilakukan adalah sebagai berikut.

- 1) Mengubah tampilan Data View dari uji validitas konstruk dengan menghapus kolom Total sehingga menjadi seperti ini:

Soal_1	Soal_2	...	Soal_24
(Skor siswa 1)	(Skor siswa 1)	(Skor siswa 1)	(Skor siswa 1)
(Skor siswa 2)	(Skor siswa 2)	(Skor siswa 2)	(Skor siswa 2)
(...)	(...)	(...)	(...)
(Skor siswa 384)	(Skor siswa 384)	(Skor siswa 384)	(Skor siswa 384)

- 2) Memilih menu Analyze, lalu sub menu Scale, kemudian pilih Reliability Analysis
- 3) Pada kotak dialog Reliability Analysis, masukkan semua variabel ke kotak Items kemudian pada bagian Model pilih Alpha
- 4) Klik kotak Statistics, centang kotak Scale if item deleted, lalu klik Continue
- 5) Klik OK untuk mengakhiri perintah dan menampilkan Output hasil
- 6) Output yang harus diperhatikan adalah tabel Reliability Statistics, yaitu pada isian Cronbach's Alpha. Jika nilainya lebih dari r_t untuk banyak soal $N = 24$ dan derajat signifikansi 0,05 serta Two-tailed yaitu 0,404 maka soal secara keseluruhan reliabel. Sebaliknya jika nilainya sama atau kurang dari 0,404 maka soal tersebut tidak reliabel.

Hasil Penelitian dan Pembahasan

Hasil validitas konstruk dengan menggunakan uji korelasi pearson dapat dilihat pada Tabel 3.

Tabel 3. Skor Korelasi Pearson

No. Soal	Skor Korelasi Pearson	Kesimpulan
1	0,431	Valid
2	0,193	Valid
3	0,329	Valid
4	0,109	Tidak Valid
5	0,608	Valid
6	0,038	Tidak Valid
7	0,077	Tidak Valid
8	0,051	Tidak Valid
9	0,159	Tidak Valid
10	0,220	Valid
11	0,378	Valid
12	0,376	Valid
13	0,329	Valid
14	0,276	Valid
15	-0,046	Tidak Valid
16	0,241	Valid
17	0,186	Valid
18	-0,211	Valid
19	0,269	Valid
20	-	Tidak Valid
21	0,338	Valid
22	0,414	Valid
23	0,179	Tidak Valid
24	0,231	Valid

Dari Tabel 3 diperoleh 16 soal yang valid. Kemudian dari 16 soal tersebut dilakukan uji reliabilitas menggunakan alpha cronbach dengan hasil seperti pada Tabel 4.

Tabel 4. Skor Cronbach Alpha

Cronbach's Alpha	N of Item
0,602	16

Dari skor cronbach alpha yang diperoleh pada Tabel 4 disimpulkan bahwa seluruh soal yang valid juga reliabel secara keseluruhan.

Simpulan

Dari hasil penelitian, diperoleh kesimpulan bahwa soal pilihan ganda yang dibuat sudah valid dan reliabel sehingga dapat digunakan untuk mendapatkan gambaran mengenai tingkat keterampilan berpikir siswa SMP.

Daftar Pustaka

Gronlund, Linn, & Miller. (2009). *Measurement and assessment in teaching*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Kartowagiran, B. (2012). *Penulisan butir soal*. Yogyakarta: Universitas Negeri Yogyakarta.

Kementrian Pendidikan dan Kebudayaan. (2013). *Peraturan menteri pendidikan dan kebudayaan nomor 66 tahun 2013 tentang standar penilaian*. Jakarta: Kementrian Pendidikan dan Kebudayaan.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*(28), 563-575.

Presiden Republik Indonesia. (2013). *Peraturan pemerintah nomor 32 tahun 2013 tentang perubahan atas peraturan pemerintah nomor 19 tahun 2005 tentang standar nasional*

pendidikan. Jakarta: Presiden Republik Indonesia.

Reynolds, C. R., Livingston, R. B., & Willson, V. (2009). *Measurement and assessment in education* (2nd ed.). Upper Saddle River, New Jersey: Pearson Education, Inc.