

## PENGEMBANGAN INSTRUMEN KOGNITIF UNTUK MENGUKUR PENALARAN SISWA SMP DI KOTA BANJARMASIN PADA MATERI CAHAYA

Heny Amelia, Mustika Wati, Sri Hartini

Program Studi Pendidikan Fisika, FKIP, Universitas Lambung Mangkurat  
henyamelia.w@gmail.com

**ABSTRAK:** Instrumen kognitif yang dibuat oleh guru untuk mengukur kemampuan siswanya selama ini belum melalui proses validasi. Berdasarkan hal tersebut perlu dilakukan penelitian untuk mendeskripsikan kualitas instrumen kognitif pada materi cahaya yang dapat mengukur kemampuan penalaran siswa. Tujuan khusus penelitian ini adalah mendeskripsikan: (1) validitas instrumen kognitif dilihat dari *Outfit* MNSQ, ZSTD, dan *Pt Measure Corr*; (2) reliabilitas instrumen kognitif yang dilihat dari nilai separasi individu, separasi butir, dan *Alpha Cronbach*; (3) tingkat kesukaran yang dilihat dari nilai *logit* pada kolom pengukuran; (4) daya beda soal dilihat dari nilai *logit* butir yang dibandingkan dengan nilai *logit* individu; (5) kemampuan penalaran siswa yang dilihat dari jumlah nilai yang diperoleh dari setiap butir soal. Penelitian pengembangan ini mengacu pada model pengembangan Borg and Gall dengan subjek penelitian 201 siswa kelas VIII yang berasal dari SMPN 22, 23, dan 28 Banjarmasin. Instrumen pengumpul data berupa instrumen validasi dan kognitif. Teknik analisis data berupa deskriptif dan kuantitatif. Hasil penelitian menunjukkan: (1) lima butir soal tergolong valid dari delapan soal yang diujicobakan; (2) reliabilitas instrumen kognitif dilihat dari nilai separasi individu tergolong cukup, nilai separasi butir tergolong istimewa, serta nilai *Alpha Cronbach* yang masuk kategori bagus sekali; (3) tingkat kesukaran soal yang dilihat dari nilai *logit* terkategori sangat sulit (25%), sulit (25%), mudah (37,5%), dan sangat mudah (12,5%); (4) daya beda soal menunjukkan lima butir soal masuk kategori bagus dan tiga soal terkategori jelek; (5) kemampuan penalaran siswa yang diperoleh untuk SMPN 23 terkategori kurang, SMPN 22 dan 28 terkategori sangat kurang.

Kata Kunci: Instrumen kognitif, cahaya, penalaran

**ABSTRACT:** *The cognitive instruments made by teachers to measure students abilities have not been through the validation process. Based on this, a reseach need to do with aim to describe the quality of cognitive instruments in light material that can measure students reasoning abilities. The specific purpose of this reseach is to describe: (1) the validity of cognitive instruments is seen from the Outfit MNSQ, ZSTD, and PtMeasure Corr scores; (2) the reliability of cognitive instruments is seen from Person reliability, Item Reliability, and Alpha Cronbach scores; (3) the level of difficulty of items seen from logit value in the measurement column; (4) the index of discrimination can view from the logit value of the item compared with the logit value of person; (5) the student reasoning abilities seen from scores in eachquestion. This development reseach refers to the development model of Bord and Gall with the subjects of study are 201 students of class VIII from SMPN 22, 23, and 28 Banjarmasin. Data collection instruments are the validation instruments and the developed cognitive instruments. Data analysis techniques are descriptive and quantitative. The result showed: (1) the five items are classified as valid of the eight items tested; (2) reliability of cognitive instruments for the person reliability is quite, the item reliability is excellent, and the Alpha Cronbach is very good; (3) the difficulty of items in category very difficult is 25%, difficult is 25%, easy is 37,5%, and very easy is 12,5%; (4) the index of discrimination view from the item and person logit value show five items have good category and three items have bad category; (5)*

*Students reasoning abilities gained for SMPN 23 Banjarmasin were categorized as less, SMPN 22 and 28 Banjarmasin were categorized very less.*

**Keywords:** *Cognitive instruments, light materials, reasoning ability*

## **PENDAHULUAN**

Berdasarkan Peraturan Menteri Pendidikan Nasional Nomor 16 Tahun 2007 tentang Kualifikasi Akademik dan Standar Kompetensi Guru menyatakan bahwa menyelenggarakan penilaian hasil belajar merupakan salah satu kompetensi inti guru yang harus dipenuhi. Seorang guru harus dapat melaksanakan evaluasi untuk mengetahui apakah materi yang diberikan dapat dipahami peserta didik atau belum (Prastya dalam Mustari, 2016). Penilaian yang sah dan objektif dapat diperoleh dengan menggunakan tes yang memuat soal atau pertanyaan-pertanyaan yang dapat mencerminkan kemampuan yang diukur (Mustari, 2016). Suatu tes yang digunakan dalam penilaian harus berkualitas baik dan mampu mengukur kemampuan siswa yang sebenarnya (Hayati dan Mardapi, 2014).

Berdasarkan hasil wawancara dan contoh soal yang diberikan oleh guru Ilmu Pengetahuan Alam (IPA) di SMPN 22, SMPN 23, dan SMPN 28 Banjarmasin, soal-soal yang digunakan untuk Ulangan Harian adalah soal yang dibuat sendiri oleh guru sesuai dengan materi yang sudah diajarkan. Setiap guru

memiliki kriteria tersendiri dalam membuat soal ulangan harian, sesuai dengan indikator dan tujuan pembelajaran yang telah ditetapkan di dalam Rencana Pelaksanaan Pembelajaran (RPP). Setelah melakukan tes kemampuan penalaran di SMPN 22 Banjarmasin khususnya kelas VIIC yang hasil jawabannya diukur dengan menggunakan Penilaian Acuan Kriteria (PAK) dapat diketahui sekitar 72,7% siswa termasuk ke dalam kategori kemampuan penalaran yang sangat kurang, hanya 27,3% siswa yang masuk ke dalam kategori kemampuan penalaran kurang. Hal tersebut disebabkan karena siswa tidak terbiasa menyelesaikan soal-soal yang dirancang untuk menggunakan kemampuan penalaran dalam menjawabnya. Kemampuan penalaran secara signifikan berpengaruh positif terhadap belajar fisika, semakin tinggi kemampuan penalarannya semakin baik pula siswa tersebut dalam menyelesaikan soal fisika (Tawil, 2008).

Berdasarkan latar belakang di atas, maka peneliti memutuskan untuk melakukan penelitian yang berjudul "Pengembangan Instrumen Kognitif Pada Materi Cahaya untuk Mengukur Kemampuan Penalaran Siswa SMP

Kelas VIII di Kota Banjarmasin". Tujuan penelitian ini adalah mendeskripsikan kualitas instrumen kognitif yang terdiri atas aspek validitas, daya beda, reliabilitas, dan taraf kesukaran, serta mendeskripsikan tingkat penalaran siswa SMP di Kota Banjarmasin yang diukur dengan soal yang telah dikembangkan.

### **KAJIAN PUSTAKA**

Pada Kurikulum Tingkat Satuan Pendidikan (KTSP) pokok bahasan Cahaya merupakan materi IPA untuk SMP Kelas VIII Semester 2 dengan standar kompetensi memahami konsep dan penerapan getaran, gelombang, dan optika dalam produk teknologi sehari-hari, sedangkan kompetensi dasar untuk materi ajar ini adalah menyelidiki sifat-sifat cahaya dan hubungannya dengan berbagai bentuk cermin dan lensa.

Kegiatan menganalisis butir soal merupakan suatu kegiatan yang harus dilakukan oleh guru untuk meningkatkan mutu soal yang telah ditulis (Adiputra, 2012). Menurut Koyan (2012) kualitas suatu tes dapat dilihat secara kualitatif dan kuantitatif. Analisis tes secara kualitatif dilaksanakan berdasarkan kaidah penulisan soal. Penelaahan tes dilakukan sebelum alat ukur tersebut digunakan. Faktor yang perlu dicermati dalam penelaahan soal secara kualitatif adalah bahasa,

konstruksi, materi, aspek isi, pedoman pemberian skor, dan kunci jawaban. Kualitas alat ukur secara kuantitatif dapat dilihat melalui validitas, reliabilitas, taraf kesukaran, daya pembeda, dan analisis pengecoh dari tes tersebut.

Instrumen merupakan alat pengumpul data yang berupa prosedur sistematis dengan aturan yang telah ditentukan (Yance, 2013). Kognitif adalah konsep, fakta, dan generalisasi yang ada dalam pikiran siswa (Widyatingtyas, 2002). Menurut Selvizia, Zainuddin, Salam (2016) tes uraian merupakan tes yang cocok untuk soal-soal dengan indikator kecerdasan logis seperti penalaran. Instrumen penilaian tes tertulis selain dapat digunakan untuk mengetahui tingkat kemampuan siswa, juga dapat digunakan sebagai sarana untuk melatih kemampuan siswa berpikir pada tingkat yang lebih tinggi (Rofiah, 2013).

Penalaran oleh Shurter dan Pierce (Sumarmo dalam Permana, 2007) didefinisikan sebagai proses pencapaian kesimpulan logis berdasarkan fakta dan sumber yang relevan. Hasil penelitian Nurhayati (2016) menunjukkan bahwa penalaran ilmiah memiliki peranan penting dalam kemampuan penyelesaian masalah fisika. Menurut Sulistiawati (2016) indikator penalaran meliputi: (1) memperkirakan jawaban dan proses

solusi; (2) menganalisis pernyataan-pernyataan dan memberikan contoh yang dapat mendukung atau bertolak belakang; (3) mempertimbangkan validitas dari argumen yang menggunakan berpikir deduktif atau induktif; (4) menggunakan data yang mendukung untuk menjelaskan mengapa cara yang digunakan serta jawaban adalah benar; dan (5) memberikan penjelasan dengan menggunakan model, fakta, sifat-sifat, dan hubungan. Adapun Sinaga (2015), indikator penalaran terdiri atas, mengidentifikasi masalah, membuat pola hubungan dengan konsep pengetahuan yang sesuai, membuat, menyangkal dan mendukung argumen serta menjelaskan fenomena secara ilmiah. Adapun indikator penalaran yang akan diukur meliputi: (1) mengumpulkan fakta; (2) mengidentifikasi masalah berdasarkan fenomena yang terjadi di dalam soal; (3) memperkirakan jawaban dan proses solusi; (4) membuat pola hubungan dengan konsep atau fakta yang sesuai; (5) menarik kesimpulan dari penjelasan atau penyelesaian yang telah diperoleh.

Teori Respon Butir dinamai juga sebagai Teori Ciri Laten atau Lengkungan Karakteristik Butir (Sudaryono, 2011). Teori respons butir mempunyai kelebihan dibandingkan teori tes klasik, yaitu statistik butir tidak tergantung pada kelompok, skor tes

yang diperoleh dapat menggambarkan kemampuan individu, tidak memerlukan tes yang paralel untuk menghitung koefisien reliabilitas, dan dapat menyediakan ukuran yang tepat untuk setiap skor kemampuan (Huriaty, 2015). Model Rasch mencoba menghubungkan perilaku atau karakteristik sebuah butir soal ketika diujikan kepada sejumlah peserta ujian dengan berbagai tingkat kemampuan (Kumaidi, 1999).

Keunggulan pemodelan Rasch dibanding metode lainnya, khususnya teori tes klasik, yaitu kemampuan melakukan prediksi terhadap data hilang berdasarkan pola respon individu. Keunggulan ini menjadikan hasil analisis statistik model Rasch lebih akurat dalam penelitian yang dilakukan, dan yang lebih penting lagi, pemodelan Rasch mampu menghasilkan nilai pengukuran eror standar untuk instrumen yang digunakan yang dapat meningkatkan ketepatan perhitungan (Ardiyanti, 2016).

## **METODE PENELITIAN**

Jenis penelitian yang digunakan adalah metode penelitian dan pengembangan. Desain penelitian pengembangan ini menggunakan model prosedur dari Borg and Gall yang diadopsi oleh Sugiyono (Mustari, 2016). Namun, pada penelitian ini hanya sampai pada tahap kesembilan. Tahap-tahap tersebut dimulai dari tahap potensi

dan masalah, pengumpulan data, desain produk, validasi desain, ujicoba produk, revisi produk, dan ujicoba pemakaian, serta revisi produk.

Pada tahap potensi dan masalah, peneliti melakukan analisis kebutuhan untuk mengetahui instrumen penilaian yang digunakan oleh guru SMP Negeri di Kota Banjarmasin dalam pembelajaran IPA. Pengumpulan data ini dilakukan dengan wawancara terhadap guru IPA Kelas VIII di SMPN 22, 23, dan 28 Banjarmasin dan tes kemampuan penalaran yang dilakukan di kelas VIII C SMPN 22 Banjarmasin. Pada tahap desain produk, peneliti mendesain produk yang akan dikembangkan yaitu pembuatan instrumen kognitif berupa soal uraian pada materi Cahaya untuk mengukur kemampuan penalaran siswa SMP kelas VIII di Kota Banjarmasin.

Kemudian tahap validasi desain, tahap ini merupakan proses penilaian rancangan produk yang dilakukan dengan memberi penilaian berdasarkan pemikiran rasional kepada dua validator yang sudah ditentukan yaitu Syubhan An'nur, M.Pd selaku akademisi dan Penguji 1 dan Saiyidah Mahtari, M.Pd selaku akademisi dan Penguji 2. Kemudian, hasil validasi pakar dianalisis untuk mengetahui validitas dan reliabilitasnya. Tahap revisi desain dilakukan untuk memperbaiki produk

yang telah dibuat sebelum produk tersebut diujicobakan.

Pada tahap ujicoba produk, produk yang telah dibuat diujicobakan pada siswa kelas IX di SMPN 28 Banjarmasin. Hasil ujicoba produk dianalisis untuk mengetahui validitas, reliabilitas, daya beda, dan tingkat kesukarannya. Analisis validitas dilakukan dengan menggunakan rumus Pearson, reliabilitas menggunakan persamaan *Alpha Cronbach*, daya beda dengan mengurangi nilai rata-rata kelompok atas dengan bawah yang kemudian dibagi dengan skor maksimum soal, serta tingkat kesukaran yang dianalisis dengan membagi rata-rata skor yang diperoleh siswa dengan skor maksimum soal. Selanjutnya produk tersebut direvisi kembali untuk memperbaiki kelemahan yang masih ada.

Pada tahap ujicoba pemakaian, produk yang dihasilkan diujicobakan ke siswa kelas VIII di SMPN 22, 23, dan 28 Banjarmasin yang seluruhnya berjumlah 201 siswa yang kemudian dianalisis dengan menggunakan program Rasch. Setelah dilakukan ujicoba pemakaian, produk tersebut direvisi kembali untuk memperbaiki kesalahan-kesalahan yang terjadi saat ujicoba pemakaian.

Subjek dari penelitian ini adalah instrumen tes kognitif yang dapat

mengukur kemampuan penalaran siswa SMP. Tes dilaksanakan selama satu kali pertemuan (2 x 40 menit). Adapun waktu penelitian dari Februari sampai Juni 2017.

Instrumen yang digunakan untuk mengumpulkan data penilaian adalah instrumen validasi soal dan instrumen kognitif yang dikembangkan. Analisis data yang digunakan dalam penelitian ini adalah Pemodelan Rasch dengan teori respon butir.

Pada program Rasch valid tidaknya suatu butir soal tergantung dari nilai MNSQ, nilai ZTSD, dan nilai korelasi pengukuran yang dihasilkan. Adapun kriteria dari ketiga nilai tersebut yaitu

1. nilai MNSQ yang baik untuk pengukuran:  $0,5 < \text{MNSQ} < 1,5$
2. nilai ZSTD yang baik untuk pengukuran:  $-2,0 < \text{ZSTD} < +2,0$
3. nilai korelasi pengukuran yang diharapkan:  $0,4 < \text{Pt Measure Corr} < 0,85$

Reliabilitas soal dengan pemodelan Rasch dianalisis dengan menggunakan nilai separasi individu dan separasi butir serta nilai *Alpha Cronbach* yang ditampilkan dalam keluaran program Rasch. Adapun kriteria yang diharapkan untuk nilai separasi individu dan butir minimal 0,67 dan Alpha Cronbach minimal 0,61 atau masuk kategori cukup.

Daya beda soal dengan pemodelan Rasch dapat dianalisis dengan menggunakan peta Wright dengan membandingkan sebaran kemampuan siswa dan kesulitan soal serta nilai logit butir dan logit individu yang terdapat dalam Tabel statistik butir dan Tabel statistik individu. Adapun kriteria yang diharapkan adalah nilai logit butir lebih besar dari nilai logit individu terendah dan lebih kecil dari nilai logit individu tertinggi atau masuk kategori daya beda bagus.

Tingkat kesukaran soal dengan pemodelan Rasch dianalisis dengan menggunakan bilangan *logit* yang terdapat dalam kolom pengukuran soal. Semakin tinggi nilai *logit*-nya maka semakin tinggi tingkat kesukaran soal tersebut dengan kategori tingkat kesukaran yaitu soal sangat sulit, sulit, mudah, dan sangat mudah. (Widhiarso dan Sumintono, 2015)

Kemampuan penalaran dianalisis dengan menjumlahkan nilai yang diperoleh dari setiap butir soal. Semakin tinggi skor keseluruhan yang diperoleh maka semakin tinggi pula kemampuan penalaran siswa tersebut. Adapun rubrik penilaian penalaran yang digunakan seperti pada Tabel 1.

Tabel 1. Kriteria Penilaian Penalaran

Skor	Kriteria
75 – 100	Sangat Baik
61 – 74	Baik
51 – 60	Sedang
35 – 50	Kurang
25 – 34	Sangat Kurang

(Bao Lei dalam Hindriana, 2014)

### HASIL DAN PEMBAHASAN

Hasil penelitian yang dijelaskan meliputi hasil analisis dari validasi desain dengan bantuan pakar, uji coba produk, dan uji coba pemakaian. Pada validasi desain diperoleh validitas dari pakar sebesar 3,7 atau masuk kategori valid dengan revisi kecil. Adapun reliabilitas yang diperoleh sebesar 0,42 atau masuk kategori bagus. Jadi, berdasarkan penilaian pakar desain instrumen kognitif yang dikembangkan valid dan reliabel sehingga dapat digunakan untuk uji coba produk dengan perbaikan terlebih dahulu sesuai dengan saran yang diberikan.

Pada uji coba produk dengan menguji 17 butir soal yang dikembangkan terdapat 13 soal valid dan 4 soal tidak valid. Analisis reliabilitas menunjukkan nilai yang diperoleh sebesar 0,89 atau masuk kategori reliabilitas tinggi. Menurut Matondang (2009), suatu instrumen tes dapat dikatakan baik jika instrumen tersebut valid dan reliabel. Jadi, terdapat 13 butir soal yang memiliki kualitas baik

dan dapat digunakan untuk uji coba selanjutnya. Adapun hasil analisis daya beda dapat dilihat pada Tabel 2

Tabel 2. Daya Beda Soal

Kategori	Nomor Soal
Bagus	16
Sedang	1, 2, 6, 10, 12, 14, 17
Jelek	3, 4, 5, 7, 8, 9, 11, 14, 15

Tabel di atas menunjukkan ada 1 butir soal yang memiliki daya beda bagus dan 7 soal memiliki daya beda sedang, serta 9 soal yang daya bedanya jelek. Sebuah soal dapat dinyatakan memiliki daya pembeda yang baik jika siswa yang kurang pandai tidak dapat mengerjakan soal dengan benar dan siswa yang pandai dapat mengerjakan soal tersebut dengan benar (Tandilling, 2012). Daya beda yang jelek disebabkan karena mayoritas soal penalaran yang dikembangkan memiliki tingkat kesukaran yang sulit sehingga banyak siswa baik dari golongan atas maupun bawah yang tidak dapat menjawabnya dengan benar. Menurut Sumintono dan Widhiarso (2015), soal yang memiliki daya beda jelek lebih baik tidak digunakan karena akan berpengaruh negatif terhadap hasil pengukuran. Hasil analisis taraf kesukaran dapat dilihat pada Tabel 3.

Tabel 3. Taraf kesukaran Soal

Kategori	Nomor Soal
Sukar	10, 13, 14, 15
Sedang	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 16, 17
Mudah	-

Tabel 3 menunjukkan ada 4 soal terkategori sukar dan 12 soal terkategori sedang, tidak ada soal yang masuk kategori mudah. Jadi, berdasarkan analisis di atas diambil 8 butir soal yang memiliki kualitas bagus untuk digunakan pada ujicoba pemakaian.

Pada tahap ujicoba pemakaian secara luas hasil analisis tentang kesesuaian butir menunjukkan keseluruhan butir soal yang diujikan valid. Hal ini dapat dilihat pada Tabel 4. berikut:

Tabel 4. Kriteria kesesuaian butir soal

Nomor Soal	Outfit		Pt Measure	Kategori
	MNS Q	ZSTD	Corr	
8, 4, 1, 2, 3,	Baik	Logis	Baik	Valid
7	Baik	Tidak dapat diprediksi Terlalu mudah diprediksi	Baik	Valid
5, 6	Baik		Baik	Valid

Berdasarkan hasil tersebut, dapat dilihat bahwa butir soal nomor 7 mempunyai kecenderungan yang kurang fit. Jika dilihat dari tiga kriteria, butir soal nomor 7 hanya tidak memenuhi syarat pada nilai *outfit* ZSTD atau data tidak dapat diprediksi, namun untuk nilai *outfit* MNSQ masih memenuhi kriteria yang baik untuk pengukuran. Adapun nilai *Pt Measure Corr* juga memenuhi kriteria. Karena itu, butir soal nomor 7 dapat dipertahankan, dan tidak

perlu diperbaiki. Hal tersebut sama dengan butir soal nomor 5 dan 6 yang memiliki nilai ZSTD yang terlalu kecil, sehingga data terlalu mudah diprediksi namun masih bisa dipertahankan untuk melakukan pengukuran kemampuan penalaran siswa. Adapun untuk butir-butir soal yang lainnya memiliki nilai MNSQ, ZSTD, dan *Pt Measure Corr* yang sesuai dengan kriteria sehingga tidak perlu diperbaiki.

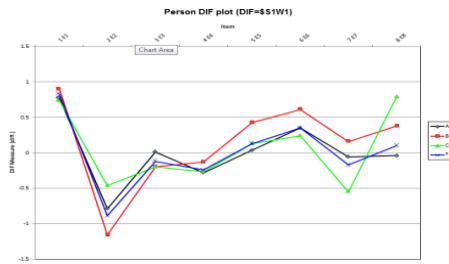
Hasil validasi butir soal juga dilihat dari bias tidaknya soal tersebut, yang dapat dilihat pada Tabel 5.

Tabel 5. Soal Bias

Nomor Soal	Kategori	Keterangan
1, 3, 4, 5, 6	Tidak bias	Tidak perlu diperbaiki
2, 7, 8	Bias	Perlu diperbaiki

Berdasarkan Tabel 5 terdapat tiga butir soal yang bias, yaitu soal nomor 2, 7, dan 8. Butir soal nomor 2 memiliki nilai probabilitas 0.0001, soal 7 memiliki nilai probabilitas 0.0031, dan soal nomor 8 memiliki nilai probabilitas 0.0111 yang menunjukkan nilainya kurang dari 0.05. Hal ini menunjukkan ketiga butir soal ini perlu diperbaiki agar tidak merugikan sekolah tertentu. Selain dengan angka, informasi adanya butir soal yang bias juga dapat diketahui melalui Gambar 1.





Gambar 1 Mendeteksi Soal Bias

Pada Gambar 1 di atas terlihat bahwa soal nomor 2 mudah dikerjakan oleh siswa yang berasal dari SMPN 22 Adapun untuk soal nomor 7 lebih mudah dikerjakan oleh siswa dari SMPN 28, dan soal nomor 8 lebih mudah dikerjakan oleh siswa dari SMPN 23 dibandingkan dengan siswa yang berasal dari SMPN 22 dan 28. Untuk butir soal yang lain, perbedaan kemampuan mengerjakan butir soal dengan benar tidak berbeda jauh. Jadi, berdasarkan bias tidaknya suatu butir soal terdapat tiga soal yang tergolong bias sehingga perlu diperbaiki agar dapat mengukur kemampuan penalaran dengan baik.

Hasil penelitian Ardiyanti (2016), menyatakan bahwa hasil perhitungan validitas menggunakan model Rasch dapat dikatakan lebih akurat daripada perhitungan klasik karena data responden yang digunakan sudah bebas dari *outliner* atau responden yang tidak sesuai dengan pengukuran yang dilakukan.

Reliabilitas soal dapat diketahui melalui nilai reliabilitas individu dan reliabilitas butir, serta nilai Alpha

Cronbach yang hasilnya dapat dilihat pada Tabel 6.

Tabel 6. Reliabilitas Soal

Kriteria	Skor	Kategori
Reliabilitas individu	0,79	Cukup memuaskan
Reliabilitas butir	0,99	Istimewa
Alpha Cronbach	0,85	Bagus sekali

Pada Tabel 6 ditunjukkan reliabilitas siswa yaitu sebesar 0.79 yang menunjukkan bahwa konsistensi jawaban dari siswa cukup bagus. Artinya siswa memang dapat menjawab soal mudah dan kesulitan dalam menjawab soal yang sukar. Nilai reliabilitas (*Cronbach Alpha*) yang diperoleh sebesar 0.85 dan masuk dalam kategori bagus sekali, sehingga hasil pengukuran yang dilakukan dapat dipercaya. Nilai reliabilitas tersebut menunjukkan interaksi antara siswa dan butir-butir soal secara keseluruhan sangat bagus. Adapun nilai reliabilitas butir soal yang diperoleh adalah 0.99 yang masuk dalam kategori istimewa. Hal itu menunjukkan bahwa kualitas butir-butir soal yang digunakan dalam pengukuran sangat bagus.

Jadi, secara keseluruhan instrumen kognitif yang dikembangkan memiliki tingkat reliabilitas yang tinggi, baik dari segi reliabilitas individu, butir soal, maupun nilai *Alpha Cronbach*. Hasil penelitian Manfaat (2013) menyatakan bahwa jika nilai reliabilitas tes yang didapatkan tinggi maka tes yang

dikembangkan dapat memberikan hasil yang sama bila diberikan pada kelompok yang sama meskipun oleh orang yang berbeda, waktu atau kesempatan yang berbeda, dan tempat yang berbeda pula sehingga konsistensi tes dianggap tinggi dan dapat dipercaya. Kereh (2015) menyatakan bahwa setidaknya ada dua sifat dasar yang dapat dijadikan acuan kualitas dari suatu instrumen pengukuran, yaitu reliabilitas dan validitasnya. Jika reliabilitas dan validitas instrumen kognitif yang dikembangkan itu valid, maka instrumen tersebut dapat digunakan seterusnya untuk melakukan pengukuran karena kualifikasinya yang baik.

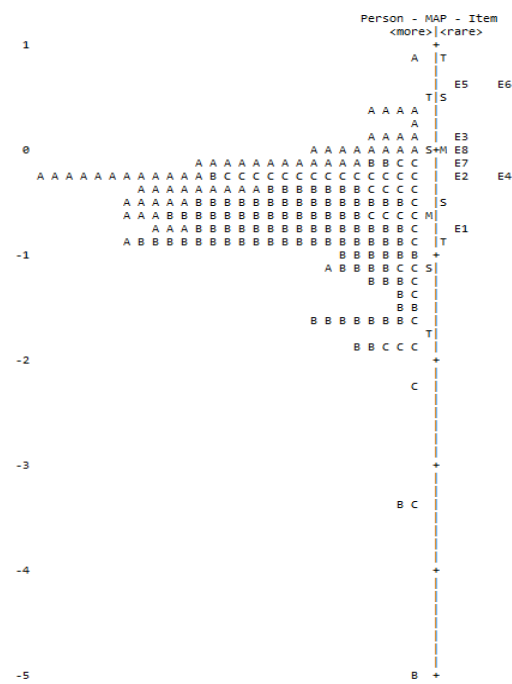
Daya beda soal dikelompokkan berdasarkan peta Wright dan perbandingan nilai logit butir dan nilai logit individu yang diperoleh dari pemrograman Rasch. Hasil analisis untuk daya beda soal dapat dilihat pada Tabel 7.

Tabel 7. Daya Beda Soal

Jenis Soal	Daya Beda Soal	Nomor Soal	Jumlah Butir
Sangat Sulit	Jelek	5, 6	2
Sedang	Bagus	2, 3, 4, 7, 8	5
Sangat Mudah	Jelek	1	1

Pada Tabel 7 soal nomor 5 dan 6 memiliki daya beda jelek karena soal memiliki tingkat kesulitan sangat sulit yang hanya dapat dijawab oleh satu siswa. Artinya, butir soal 5 dan 6 merupakan soal dengan tingkat

kesukaran tertinggi dan kemungkinan semua siswa dapat mengerjakan soal ini dengan benar sangat kecil. Adapun soal nomor 1 juga memiliki daya beda jelek, hal ini karena nilai logit butir yang paling rendah sehingga soal tergolong sangat mudah, dan dapat dijawab oleh 142 siswa. Hal ini menunjukkan hampir semua siswa dapat mengerjakan soal tersebut dengan benar. Soal nomor 2, 3, 4, 7, 8 tergolong soal dengan daya pembeda bagus hal ini karena nilai logit butir yang lebih kecil dari nilai logit individu tertinggi dan lebih besar dari nilai logit individu terendah artinya soal memiliki tingkat kesulitan sedang dan tidak dapat dijawab oleh siswa kelompok bawah namun dapat dijawab benar oleh siswa kelompok atas.



Gambar 2. Penyebaran Daya Beda Soal

Agar lebih jelas dalam memahami pengelompokkan daya pembeda soal yang dikembangkan, dapat dilihat pada Gambar 2.

Gambar 2 menunjukkan butir soal nomor 5 dan 6 merupakan butir soal yang paling sulit, dan hanya ada satu siswa dari SMPN 23 yang dapat mengerjakannya. Soal nomor 3 dan 8 termasuk dalam kategori soal dengan daya beda bagus, namun kedua soal tersebut hanya dapat diselesaikan oleh siswa dari SMPN 23 saja. Soal nomor 7 juga masuk kategori daya beda bagus dan dapat dijawab oleh siswa dari SMPN 23 dan beberapa siswa dari SMPN 28. Soal nomor 2 dan 4 juga memiliki kategori daya pembeda bagus dan tingkat kesukaran yang sama, hal itu dilihat dari letak soal nomor 2 dan 4 yang berada pada baris yang. Kedua soal tersebut dapat dijawab oleh siswa dari SMPN 22, 23, dan 28 dengan mayoritas siswa dari SMPN 23 dan 28 lebih banyak menjawab benar daripada siswa dari SMPN 22. Adapun soal nomor 1 merupakan soal paling mudah dengan daya beda jelek, hal ini ditunjukkan dari posisi butir soal yang berada di sisi kanan bawah pada gambar. Pada soal ini banyak siswa yang dapat menjawabnya dengan benar.

Hasil penelitian Rahayu, Purnomo, dan Sukidin (2014) menyatakan bahwa soal yang memiliki daya beda bagus

dinyatakan sebagai soal yang dapat digunakan untuk menilai hasil belajar siswa. Menurut hasil penelitian Fitriatun dan Sukanti (2016) soal yang memiliki daya beda jelek terjadi karena soal tersebut merupakan soal yang sukar sehingga siswa tidak dapat menyelesaikan soal tersebut dengan baik. Hasil penelitian Samsiah, Ruslan, dan Sappaile (2015) menyatakan bahwa ada beberapa faktor yang menyebabkan daya beda butir soal jelek yaitu kunci jawaban butir soal tidak tepat, butir soal memiliki dua atau lebih kunci jawaban yang benar, kompetensi yang diukur tidak jelas, serta indeks butir soal yang terlalu tinggi dan terlalu rendah. Butir soal yang terlalu sulit atau mudah tidak dapat membedakan peserta didik pandai dan peserta didik kurang pandai sehingga tidak mempunyai daya pembeda yang baik. Jadi, terdapat lima butir soal dari dengan kategori daya beda bagus yang dapat digunakan untuk mengukur tingkat penalaran siswa SMP pada materi cahaya.

Hasil pengukuran untuk tingkat kesulitan soal dapat dilihat pada Tabel 8.

Tabel 8. Tingkat Kesukaran Soal

Tingkat Kesulitan	Nomor Soal	Jumlah Butir Soal	Persentase Soal
Sangat sulit	5, 6	2	25%
Sulit	3, 8	2	25%
Mudah	7, 4, 2	3	37,5%
Sangat mudah	1	1	12,5%

Pada Tabel 8 dapat dilihat bahwa soal nomor 5 merupakan soal tersulit dan soal nomor 1 merupakan soal termudah, hal itu sesuai dengan ranah kognitif dari soal penalaran yang dikembangkan. Untuk soal nomor 5, 6, dan 3 memiliki ranah kognitif C5, soal 8, 7, dan 4 memiliki ranah kognitif C4, dan soal nomor 1 memiliki ranah kognitif C3. Adapun untuk soal nomor 2 memiliki ranah kognitif C5, namun hal tersebut tidak sesuai dengan analisis penelitian yang menunjukkan soal tersebut terkategori sangat mudah. Hal itu juga ditunjukkan dari banyaknya jumlah siswa yang dapat menjawab benar soal tersebut. Berdasarkan tabel tersebut terdapat empat kelompok soal yaitu 2 buah (25%) soal sangat sulit yaitu soal nomor 5 dan 6, 2 buah (25%) soal sulit yaitu soal nomor 3 dan 8, 3 buah (37,5%) soal mudah yaitu soal nomor 7, 4, dan 2, serta 1 buah (12,5%) soal sangat mudah yaitu soal nomor 1.

Jadi, berdasarkan hasil tersebut ada empat soal yang tergolong sulit, tiga butir soal mudah, dan satu butir soal sangat mudah. Kereh (2015) menyatakan bahwa jika soal yang tergolong sulit bagi siswa memiliki jumlah lebih banyak daripada soal mudah, maka hal itu menegaskan kondisi siswa yang menjadi subyek penelitian ini termasuk dalam kelompok siswa dengan kemampuan penalaran

yang kurang. Menurut Manfaat (2013), ada beberapa faktor yang menyebabkan butir soal berkategori sukar yaitu: (1) butir soal itu “mungkin” salah kunci jawaban; (2) butir soal itu mempunyai 2 atau lebih jawaban yang benar; (3) Materi yang ditanyakan belum diajarkan atau belum tuntas pembelajarannya, sehingga kompetensi minimum yang harus dikuasai peserta didik belum tercapai; dan (4) pernyataan atau kalimat soal terlalu kompleks dan panjang. Namun, dari hasil penelitian soal yang dikembangkan memiliki kualitas yang baik karena memiliki tingkat kesukaran yang beragam, yaitu sangat sulit, sulit, mudah, dan sangat mudah.

Hasil kemampuan penalaran siswa dapat dilihat pada Tabel 9.

Tabel 9. Tingkat Penalaran Siswa

Sekolah	Rata-rata	Kategori
SMPN 23	40,99	Kurang
SMPN 22	18,28	Sangat Kurang
SMPN 28	24,99	Sangat Kurang

Berdasarkan tabel tersebut dapat dilihat bahwa SMPN 23 memiliki tingkat penalaran yang lebih baik dibandingkan dengan SMPN 22 dan 23. Namun, yang perlu diperhatikan adalah kemampuan penalaran SMPN 22 yang jumlah siswa dengan tingkat kemampuan penalaran sangat kurang lebih banyak dari SMPN 28. Menurut Sugiyono (2014) hasil penelitian sangat tergantung dari obyek yang diteliti, sehingga peneliti harus mampu

mengendalikan obyek yang diteliti tersebut.

Hasil penelitian Utami (2014) menyatakan bahwa kemampuan penalaran yang kurang disebabkan oleh materi yang sulit, seperti materi yang membutuhkan tingkat ketelitian yang tinggi dan cakupan yang luas, sehingga siswa tidak efektif untuk menyelesaikannya. Menurut Amir (2014), kemampuan penalaran sangat diperlukan dalam memahami suatu konsep materi pokok. Tanpa adanya kemampuan penalaran, maka peserta didik akan mengalami kesulitan dalam menyelesaikan suatu permasalahan. Tawil (2008) menyatakan bahwa kemampuan penalaran berpengaruh positif yang signifikan terhadap hasil belajar fisika. Hal ini berarti apabila siswa senantiasa dibiasakan untuk mengembangkan kemampuan penalarannya, seperti mengerjakan soal-soal fisika yang membutuhkan penalaran untuk diselesaikan, maka hasil belajar fisika dapat lebih baik lagi.

## SIMPULAN

Berdasarkan data yang dihasilkan dari penelitian dengan mengacu pada tujuan yang telah dibuat sebelumnya, secara umum dapat disimpulkan bahwa terdapat lima butir soal yang memiliki kualitas bagus dan dapat digunakan untuk mengukur kemampuan penalaran pada siswa SMP. Instrumen kognitif

yang dikembangkan secara keseluruhan memiliki validitas yang valid. Reliabilitas untuk keseluruhan soal yang ditinjau dari nilai reliabilitas individu masuk kategori cukup, dan nilai reliabilitas soal menunjukkan istimewa, dan nilai *Alpha Cronbach* masuk kategori bagus sekali. Jumlah soal yang memiliki taraf kesukaran sangat sulit sebanyak 25%, soal sulit 25%, mudah 37,5%, dan sangat mudah 12,5%. Terdapat lima butir soal dengan daya beda bagus dan tiga butir soal dengan daya beda jelek. Adapun untuk kemampuan penalaran, SMPN 23 termasuk kategori penalaran kurang, SMPN 22 termasuk kategori penalaran sangat kurang dan SMPN 28 juga termasuk dalam kategori penalaran sangat kurang.

## DAFTAR PUSTAKA

- Adiputra, I. B. R. (2012). Analisis Butir Soal Tes Ulangan Akhir Semester IPS Terpadu Buatan MGMP IPS Kabupaten Gianyar Kelas VII Semester 1 Tahun Pelajaran 2011-2012. *Jurnal Penelitian dan Evaluasi pendidikan Indonesia*, 2(1).
- Amir, A. (2015). Kemampuan Penalaran dan Komunikasi dalam Pembelajaran Matematika. *Logaritma*, 2(01).
- Ardiyanti, D. (2016). Aplikasi Model Rasch pada Pengembangan Skala Efikasi Diri dalam Pengambilan Keputusan Karir Siswa. *Jurnal Psikologi*, 43(3), 248-263.

- Fitriatun, A., & Sukanti, S. (2016). Analisis Validitas, Reliabilitas, Dan Butir Soal Latihan Ujian Nasional Ekonomi Akuntansi Di MAN Maguwoharjo. *Kajian Pendidikan Akuntansi Indonesia*, 5(8).
- Hayati, N., & Mardapi, D. (2014). Pengembangan butir soal matematika SD di Kabupaten Lombok Timur sebagai upaya dalam pengadaan bank soal. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran*, 44(1).
- Hindriana, Anna Fitri. (2014). *Pembelajaran Fisiologi Tumbuhan Terintegrasi Struktur Tumbuhan Berbasis Kerangka Instruksional Marzano untuk Menurunkan Beban Kognitif Mahasiswa*. Disertasi Doktor. Universitas Pendidikan Indonesia
- Huriaty, D. (2015). Metode Kalibrasi dan Desain Tes Berdasarkan Teori Respons Butir (IRT). *Math Didactic: Jurnal Pendidikan Matematika*, 1(3).
- Kereh, Cicylia T. (2015). Validitas dan Reliabilitas Instrumen Tes Matematika Dasar yang Berkaitan dengan Pendahuluan Fisika Inti. *Jurnal Pendidikan*. Vol 02, No 01
- Koyan, Wayan. (2012). *Konstruksi Tes*. Singaraja: Universitas Pendidikan Ganesha Press
- Kumaidi, K. (2009). Aplikasi Teori Rasch untuk Penyekalaan Vertikal Tes Catur Wulan. *Jurnal Ilmu Pendidikan Universitas Negeri Malang*, 6(2).
- Manfaat, Budi dan Siti Nurhairiyah. (2013). Pengembangan Instrumen Tes untuk Mengukur Kemampuan Penalaran Statistik Mahasiswa Tadris Matematika. *Jurnal Jurusan Tadris Matematika*. Vol 02, No 02
- Matondang, Z. (2009). Validitas dan reliabilitas suatu instrumen penelitian. *Jurnal Tabularasa*, 6(1), 87-97.
- Mustari, M. (2016). Pengembangan Instrumen Ranah Kognitif pada Pokok Bahasan Fluida Statis SMA/MA. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, 5(1), 121-130.
- Nurhayati, N., Yuliati, L., & Mufti, N. (2016). Pola Penalaran Ilmiah dan Kemampuan Penyelesaian Masalah Sintesis Fisika. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan*, 1(8), 1594-1597.
- Permana, Y., & Sumarmo, U. (2007). Mengembangkan kemampuan penalaran dan koneksi matematik siswa SMA melalui pembelajaran berbasis masalah. *Jurnal Educationist*, 1(2), 116-123.
- Rahayu, Tika Dwi, Bambang Hari Purnomo, dan Sukidin. (2014). Analisis Tingkat Kesukaran dan Daya Beda Soal Ujian Tengah Semester Ganjil Bentuk Pilihan Ganda Mata Pelajaran Ekonomi Kelas X di SMA Negeri 5 Jember Tahun Ajaran 2012-2013. *Jurnal Edukasi UNEJ*. Vol 01, No 01
- Ranjabar, Jacobus. (2015). *Dasar-dasar Logika*. Bandung: Penerbit Alfabeta
- Rofiah, E., Aminah, N. S., & Ekawati, E. Y. (2013). Penyusunan Instrumen tes kemampuan berpikir tingkat tinggi fisika pada siswa SMP. *Jurnal Pendidikan Fisika*, 1(2).
- Samisah, H. (2016). Analisis Kualitas Tes Try Outujian Sekolah 2014/2015 Mata Pelajaran Matematika Sekolah Dasar Pada Gugus 30 Wilayah IV Kecamatan Donri-Donri Kabupaten Soppeng. *Riset Assesmen*, 1(1).

- Selvizia, N., Zainuddin, Z., & Salam, A. (2016). Pengembangan Perangkat Pembelajaran Berorientasi Kecerdasan Logis-Matematis Pada Pokok Bahasan Impuls Dan Momentum Dengan Menggunakan Model Direct Instruction Di Sma Muhammadiyah 1 Banjarmasin. *Berkala Ilmiah Pendidikan Fisika*, 4(2), 104-111.
- Sinaga, T. (2016). Pengembangan Soal Model Pisa Mata Pelajaran Ilmu Pengetahuan Alam Terpadu Konten Fisika Untuk Mengetahui Penalaran Siswa Kelas IX. *Jurnal Inovasi Dan Pembelajaran Fisika*, 2(2), 194-196.
- Sudaryono, S. (2011). Implementasi Teori Responsi Butir (Item Response Theory) Pada Penilaian Hasil Belajar Akhir di Sekolah. *Jurnal Pendidikan dan Kebudayaan*, 17(6).
- Sugiyono. (2014). *Metode Penelitian Pendidikan*. Bandung: Penerbit Alfabeta
- Sulistiawati, Didi Suryadi, dan Siti Fatimah. (2016). Peningkatan Kemampuan Penalaran Matematis Menggunakan Desain Didaktis Berdasarkan kesulitan Belajar pada materi Luas dan Volume Limas. *Jurnal Penelitian Pendidikan Matematika*. Vol 09, No 01
- Sumintono, Bambang dan Wahyu Widhiarso. (2015). *Aplikasi Pemodelan RASCH pada Assessment Pendidikan*. Jakarta: Trim Komunikata
- Tandilling, E. (2012). Pengembangan instrumen untuk mengukur kemampuan komunikasi matematik, pemahaman matematik, dan selfregulated learning siswa dalam pembelajaran matematika di sekolah menengah atas. *Jurnal Penelitian Pendidikan*, 13(1), 24-31.
- Tawil, M. (2008). Kemampuan Penalaran Formal Dan Lingkungan Pendidikan Keluarga Dikaitkan Dengan Hasil Belajar Fisika Siswa Kelas X SMA Negeri 1 Sungguminasa Kabupaten Gowa. *Jurnal Pendidikan dan Kebudayaan*, 14(75), 1047-1068.
- Utami, N. P. (2014). Kemampuan penalaran matematis siswa Kelas XI IPA SMAN 2 Painan melalui penerapan pembelajaran think pair square. *Jurnal Pendidikan Matematika*, 3(1).
- Widyatiningtyas, Reviandari. (2002). Pembentukan Pengetahuan Sains, Teknologi, dan Masyarakat dalam Pandangan Pendidikan IPA. *Jurnal Pendidikan*. Vol 01, No 01
- Yance, R. D. (2013). Pengaruh penerapan model project based learning (PBL) terhadap hasil belajar fisika siswa kelas XI IPA SMA Negeri 1 Batipuh Kabupaten Tanah Datar. *Pillar of Physics Education*, 1(1).