

PENGEMBANGAN INSTRUMEN TES HASIL BELAJAR KOGNITIF FISIKA KELAS XI MIA SMA NEGERI 5 PINRANG

DEVELOPMENT OF INSTRUMENT PHYSICS LEARNING RESULT IN CLASS XI SCIENCE SENIOR HIGH SCHOOL 5 IN PINRANG

¹⁾Muh Al Ihwan, ²⁾Salamang Salmiah Sari, ³⁾Muhammad Sidin Ali
Universitas Negeri Makassar
Kampus UNM Parangtambung Jln. Daeng Tata Raya, Makassar, 90224
¹⁾e-mail : Muhammadalihwan9@gmail.com

Abstrak. Pengembangan Instrumen Tes Hasil Belajar Kognitif Fisika Kelas XI MIA SMA Negeri 5 Pinrang. Penelitian ini bertujuan untuk mengembangkan instrumen tes hasil belajar fisika untuk materi Kesetimbangan dan Dinamika Rotasi, Elastisitas dan Hukum Hooke, Fluida Statik, Fluida Dinamik, Suhu dan Kalor, Teori Kinetik Gas serta Termodinamika. Dari tujuh Kompetensi dasar pada kelas XI MIA semester ganjil, dari 105 butir soal terdapat 103 butir soal valid hasil validasi teoretik oleh 2 pakar, 103 butir soal valid dengan koefisien konsistensi internal sebesar 0,98 setelah validasi empirik dari 103 butir soal diperoleh 62 butir soal valid setelah dilakukan uji coba sebanyak 2 kali. Pada ujicoba awal diperoleh 69 butir soal valid, pada ujicoba akhir diperoleh 62 butir soal yang valid dan diperoleh nilai reabilitas tes sebesar 0,93. tingkat kesukaran soal 7 butir soal (11,29%) pada kategori sukar, 55 butir soal (88,71%) pada kategori sedang. Untuk daya pembeda 3 butir soal (4,83%) terdapat pada kategori cukup, 18 butir soal (29,03%) pada kategori baik, dan 41 butir soal (66,12%) terdapat pada kategori sangat baik. efektivitas pengecoh semua pengecoh telah berfungsi dengan baik karena tidak terdapat pengecoh yang dipilih kurang dari 5 peserta didik (5%). Sehingga ke 62 butir soal tes hasil belajar fisika yang dikembangkan valid secara teoretik dan reliabel serta layak digunakan.

Kata kunci : Hasil Belajar, Instrumen Tes, Pengembangan, Reabilitas, Valid

Abstract. Development Of Instrument Physics Learning Result In Class Xi Science Senior High School 5 In Pinrang. This research aims to develop physics learning test instruments for equilibrium and dynamics of rotation, elasticity and Hooke's law, static fluid, dynamic fluid, temperature and heat, kinetic theory of gas and thermodynamics. Of the seven basic competencies in class XI MIA odd semester, out of 105 items there are 103 valid questions the results of theoretical validation by 2 experts, 103 items are valid with internal consistency coefficients of 0.98 after empirical validation of 103 items obtained 62 items valid after testing twice. In the initial trial, 69 items were valid, in the final trial 62 items were valid and the test reliability was 0.93. The difficulty was 7 items (11.29%) in the difficult category, 55 items (88.71%) in the medium category. For distinguishing 3 items (4.83%) in the sufficient category, 18 items (29.03%) in the good category, and 41 items (66.12%) in the good category. the effectiveness of the deceitfulness of all the deceiters has functioned well because there were no deceivers chosen less than 5 students (5%), so that the 62 items of test results of physics learning that were developed were valid theoretically and reliably and were worthy of use.

Keywords : Development, Instrument Test, Learning Result, Reliably, Valid.

PENDAHULUAN

Evaluasi hasil belajar yang baik pembelajaran adalah mengadakan penilaian hasil belajar peserta didik. Oleh karena itu pendidik perlu menggunakan instrument tes yang bermutu untuk menyusun suatu perangkat penilaian yang dapat membantu pendidik meningkatkan kualitas digunakan untuk mengukur ketuntasan hasil belajar pembelajaran dan memberikan informasi dengan peserta didik yang dilihat dari pencapaian indikator tepat tentang peserta didik yang belum atau yang hasil belajar dan tujuan pembelajaran khusus yang sudah mencapai tujuan pembelajaran. Evaluasi dicapai peserta didik. Perangkat penilaian yang berkaitan dengan kegiatan mengukur dan menilai. demikian pada akhirnya dapat dijadikan sebagai salah satu tugas utama pendidik dalam proses

acuan pendidik dalam pengambilan keputusan yang tepat terhadap peserta didik.

Hasil belajar terdiri dari tiga aspek yaitu afektif, kognitif dan psikomotorik. Berdasarkan informasi dari hasil observasi dan wawancara awal yang telah dilakukan kepada para guru fisika di SMA Negeri 5 Pinrang pada hari rabu tanggal 8 agustus 2018 ternyata umumnya evaluasi yang dibuat tidak melalui proses pembakuan. Para guru menyatakan bahwa instrumen tes hasil belajar kognitif fisika yang selama ini diujikan masih memerlukan penyempurnaan serta membutuhkan instrumen tes yang valid dan reliabel. Karena pada dasarnya, menggunakan instrumen yang valid dan reliabel dalam mengukur hasil belajar peserta didik merupakan hal yang penting. Instrumen yang valid dan reliabel dapat memberikan data atau gambaran tentang peserta didik secara benar sesuai dengan kenyataan atau keadaan sesungguhnya. Dari permasalahan tersebut peneliti melakukan penelitian pengembangan instrumen pada ranah kognitif sebagai Usaha untuk membantu guru dari masalah tersebut, oleh kerena itu diperlukan adanya pengembangan instrumen penilaian hasil belajar fisika peserta didik SMA yang baku, sehingga dapat digunakan untuk mengukur hasil belajar peserta didik.

Berdasarkan uraian tersebut, maka peneliti tertarik untuk mengembangkan instrumen tes yang valid dan reliabel ditinjau dari seberapa besar koefien validitas dan reabilitasnya. Peneliti mengembangkan instrumen tes yang berpatokan pada kisi-kisi, memperhatikan tingkat kesulitan soal, melakukan telaah butir-butir soal, menguji coba butir-butir yang telah di telaah dan menganalisis butir-butir soal. Sehingga peneliti melakukan penelitian yang berjudul **“Pengembangan Instrumen Tes Hasil Belajar Fisika Kelas XI MIA SMA Negeri 5 Pinrang”**

Adapun tujuan dari penelitian ini Mengetahui hasil pengembangan instrumen tes hasil belajar kognitif fisika untuk peserta didik

Kelas XI MIA di SMA Negeri 5 Pinrang ditinjau dari segi teoretik, dan Mengetahui hasil pengembangan instrumen tes hasil belajar kognitif fisika untuk peserta didik Kelas XI MIA di SMA Negeri 5 Pinrang ditinjau dari segi empirik.

Pengertian instrumen yang dinyatakan Djaali dan Muljono (2008) menjelaskan bahwa “secara umum yang dimaksud dengan instrumen adalah suatu alat yang memenuhi kebutuhan akademis, sehingga dapat dipergunakan sebagai alat untuk mengukur suatu objek ukuran atau mengumpulkan data mengenai suatu variable”. Instrumen atau alat pengumpulan data adalah alat yang digunakan untuk mengumpulkan data dalam suatu penelitian. Data yang dikumpulkan dengan menggunakan instrumen tertentu akan dideskripsikan dan ajukan dalam suatu penelitian (Djaali & Muljono, 2008).

Permendikbud RI Nomor 23 Tahun 2016 tentang Standar Penilaian menegaskan bahwa instrumen penilaian harus memenuhi persyaratan substansi, konstruksi, dan bahasa, serta memiliki bukti validasi empirik. Substansi merepresentasikan kompetensi yang dinilai, konstruksi memenuhi persyaratan teknis sesuai dengan bentuk instrumen yang digunakan, dan penggunaan bahasa yang baik dan benar serta komunikatif sesuai dengan tingkat perkembangan peserta didik.

Secara garis besar, setelah menetapkan tujuan dan menganalisis kurikulum dan materi, Menurut Djaali & Muljono (2008) langkah-langkah penyusunan dan pengembangan. instrumen adalah sebagai berikut: (1) mengkaji teori tentang suatu konsep dari variabel yang hendak diukur kemudian merumuskan konstruk dari variabel tersebut, (2) mengembangkan dimensi dan indikator variabel yang sebenarnya telah tertuang pada rumusan konstruk variabel, (3) membuat kisi-kisi instrumen berupa tabel yang memuat dimensi, indikator, nomor butir, dan jumlah butir, (4) menetapkan besaran atau parameter yang bergerak dalam suatu rentangan kontinum dari suatu kutub ke kutub yang

lain yang berlawanan. Misalnya dari rendah ke tinggi, positif ke negatif, otoriter ke demokratik, dependen ke independen, (5) menulis butir-butir instrumen yang dapat berbentuk pertanyaan atau pernyataan. Butir instrumen diklasifikasikan kelompok butir positif dan negatif, (6) melakukan validasi teoritik yaitu validasi pakar atau panel untuk menelaah seberapa jauh dimensi merupakan jabaran yang tepat dari konstruk, seberapa jauh butir-butir instrumen yang dibuat secara tepat dapat mengukur indikator, (7) memperbaiki/revisi instrumen berdasarkan saran dari pakar hingga dikatakan valid secara teoretik, (8) menggandakan instrumen secara terbatas untuk keperluan ujicoba, (9) mengujicoba instrumen di lapangan yang merupakan proses validasi empirik, (10) menguji validitas dilakukan dengan menggunakan kriteria internal. Kriteria internal adalah instrumen itu sendiri satu kesatuan yang dijadikan kriteria, (11) memutuskan valid atau tidaknya sebuah butir atau sebuah perangkat instrumen. Jika menggunakan kriteria internal, yaitu skor total instrumen sebagai kriteria, maka proses dan keputusan pengujian mengenai valid atau tidaknya butir instrumen disebut analisis butir, (12) untuk validitas insternal, butir-butir yang tidak valid berdasarkan analisis butir dikeluarkan atau diperbaiki untuk di ujicoba ulang. Sedangkan butir yang valid dirakit kembali menjadi sebuah perangkat instrumen untuk melihat kembali validitas kontennya berdasarkan kisi-kisi. Jika secara konten butir-butir yang valid tersebut dianggap valid atau memenuhi syarat perangkat instrumen yang terakhir ini menjadi instrumen final yang akan digunakan unruk mengukur variabel penelitian, (13) menghitung koefisien reliabilitas. Makin tinggi koefisien reliabilitas maka makin tinggi pula kualitas instrumen tersebut, dan (14) merakit butir-butir instrumen yang valid untuk dijadikan instrumen final.

Instrumen penilaian yang baik harus memenuhi dua hal penting yakni ketepatannya mengukur apa yang mesti di ukur dan ketepatannya atau keejangannya. Kusaeri dan Suprananto (2012)

menyatakan bahwa “Kualitas sebuah tes tergantung pada seberapa tepat dan akurasi hasil ukurannya, seberapa handal kemampuan tes dalam mengukur dan seberapa handal kemampuan tes dalam mengukur dan seberapa praktis tes tersebut dapat digunakan”.

Instrumen atau alat pengumpul data adalah alat yang digunakan untuk mengumpulkan data dalam suatu penelitian. Untuk itu kita dapat menggunakan instrument yang telah tersedia dapat pula menggunakan instrument yang dibuat sendiri. Jika instrument baku telah tersedia untuk mengumpulkan data variabel penelitian maka kita dapat langsung menggunakan instrument tersebut, jika instrument baku belum tersedia, maka instrument untuk mengumpulkan data variabel tersebut harus dibuat sendiri. Prosedur pengembangan yang digunakan dalam penelitian ini mengadaptasi pada konsep penyusunan dan pengembangan instrument oleh (Djaali & Muljono, 2004). Pengembangan instrumen penilaian sangatlah memerlukan kecermatan karena instrumen tersebut umumnya sulit distandarisasikan dan memiliki reliabilitas rendah. Validitas isi (*content validity*) yaitu, tes yang dikatakan memiliki validitas isi apabila mengukur tujuan khusus tertentu yang sejajar dengan materi atau isi pelajaran yang diberikan, atau validitas isi juga berkaitan dengan pertanyaan “sejauh mana item tes mencakup keseluruhan materi atau bahan yang diukur”.

Tes objektif sering juga disebut tes dikotomi (*dichotomously scored item*) karena jawabannya antara benar atau salah dan skornya antara 1 atau 0. Soal tes bentuk pilihan ganda dapat digunakan untuk mengukur hasil belajar yang lebih kompleks dan berkenaan dengan aspek ingatan, pengertian, aplikasi, analisis, sintesis, dan evaluasi. Pilihan jawaban (*option*) terdiri atas jawaban yang benar atau paling benar, selanjutnya disebut kunci jawaban dan kemungkinan jawaban salah yang dinamakan pengecoh (Arifin, 2012).

Hosnan (2014) mengemukakan bahwa belajar adalah suatu proses usaha yang sengaja dilakukan peserta didik untuk memperoleh suatu perubahan tingkah laku yang baru secara keseluruhan, secara sadar, dan perubahan tersebut relatif menetap serta membawa pengaruh dan manfaat yang positif bagi siswa dalam berinteraksi dengan lingkungannya. Menurut Rusman (2013:93) pembelajaran pada hakikatnya merupakan proses interaksi antara guru dengan siswa, baik interaksi secara langsung seperti kegiatan tatap muka maupun secara tidak langsung dengan menggunakan berbagai media pembelajaran. Hasil belajar adalah sejumlah pengalaman yang diperoleh siswa yang mencakup ranah sikap, pengetahuan, dan keterampilan. Hasil belajar merupakan kemampuan yang dimiliki siswa setelah ia menerima pengalaman belajarnya (Rusman, 2013:123).

KI I sasarannya adalah pengembangan sikap spiritual, yakni sikap yang berkaitan dengan keimanan dan ketaqwaan terhadap Tuhan Yang Maha Esa, sesuai dengan agama dan kepercayaan masing – masing. KI II sasarannya adalah pengembangan sikap sosial, yakni sikap yang berkaitan dengan nilai moral, etika, dan norma hukum yang berlaku dalam kehidupan bermasyarakat, berbangsa, dan bernegara. KI III sasarannya adalah pengembangan pengetahuan, wawasan, dan kemampuan berfikir. KI IV sasarannya adalah pengembangan keterampilan. Metode penelitian dan pengembangan adalah metode penelitian yang digunakan untuk menghasilkan produk tertentu, dan menguji keektifan produk tersebut (Sugiyono, 2009).

Suwarno (2009) mengemukakan syarat instrumen penilaian yang baik memiliki ciri-ciri dan harus memenuhi beberapa kaidah berikut ini 1) validitas, 2) reliabilitas, 3) objektivitas, 4) praktikabilitas, 5) ekonomis, 6) taraf kesukaran dan 7) daya pembeda

METODE

Jenis penelitian ini adalah penelitian pengembangan instrumen penilaian tes hasil belajar fisika dalam ranah kognitif kelas XI semester ganjil berdasarkan kurikulum 2013. Penelitian ini dilaksanakan dari bulan September sampai Desember 2018. Lokasi uji coba instrumen dilaksanakan di SMA Negeri 5 Pinrang, Kabupaten Pinrang, Sulawesi Selatan. Pada penelitian ini terdapat dua data yang dikumpulkan yakni data hasil validasi teoretik oleh pakar dan data hasil uji coba lapangan. Data hasil validasi pakar dikumpulkan dengan cara membagikan lembar validasi kepada dua orang pakar. Sedangkan data hasil uji coba lapangan dikumpulkan dengan cara membagikan instrumen hasil belajar kepada subjek coba. Untuk menentukan koefisien validitas isi, hasil penilaian dari kedua pakar dimasukkan ke dalam tabulasi silang 2×2 yang terdiri dari kolom A, B, C dan D. Setelah butir soal divalidasi isi oleh 2 penilai, selanjutnya dianalisis dengan menggunakan perhitungan menurut **Gregory** (Gregory, 2015:121). Untuk menghitung tingkat kesukaran tiap butir soal digunakan persamaan:

$$p = \frac{p_h - p_l}{2}$$

Keterangan:

- p = indeks kesukaran/kemudahan
- p_h = proporsi pesertad didik pada kelompok atas yang menjawab benar butir
- p_l = proporsi pesertad didik pada kelompok atas yang menjawab salah butir

Tabel 1. Klasifikasi Tingkat Kesukaran

Indeks kesukaran	Kategori
$p < 0,3$	Sukar
$0,31 < p \leq 0,70$	Sedang
$0,71 < p$	Sangat mudah

(Ali dan Khaeruddin, 2012:90-91)

Daya pembeda butir soal dihitung dengan menggunakan persamaan berikut.

$$D = P_h - P_1$$

Tabel 2. *Kategori Daya Pembeda*

Daya pembeda	Kategori
$0,4 \leq D$	Sangat baik/ soal diterima baik
$0,30 \leq D \leq 0,39$	Baik/ soal diterima tetapi perlu diperbaiki
$0,20 < D \leq 0,29$	Cukup/ soal diperbaiki
$D \leq 0,20$	Jelek/ soal dibuang

(Ali dan Khaeruddin, 2012:93)

Untuk validasi empirik menggunakan rumus koefisien korelasi biseral.(Djaali dan Muljono 2004:71). Dan reabilitas tes hasil belajar dihitung dengan menggunakan rumus KR-20. (Djaali dan Muljono 2004:78).

HASIL DAN PEMBAHASAN

Setelah membuat instrumen dan di konsultasikan di pembimbing terdapat 105 instrumen. hasil kesepakatan antar dua pakar dengan menggunakan analisis Gregory.

Tabel 3. *Hasil Analisis Gregory Instrumen Tes Hasil Belajar*

		Ahli 1	
Tabel Gregory	Skor 1-2 (Tidak relevan)	Skor 3-4 (Relevan)	
	Skor 1-2 (Tidak relevan)	A (0)	B (2 item)
Ahli 2	Skor 3-4 (Relevan)	C (0)	D (103 item)

Berdasarkan hasil perhitungan koefisien konsistensi antarpakar diperoleh nilai sebesar 0.98.

Pada uji coba pertama dilakukan analisis butir secara kuantitatif soal yang drop atau tidak

valid sebanyak 25 butir. Untuk tingkat kesukaran sebagai berikut

Tabel 4. *Hasil Analisis Taraf Kesukaran Butir*

No	Kategori	Jumlah	Persentase
11	Sukar	40	38,83%
22	Sedang	62	60,19%
33	Sangat Mudah	1	0,97%
Total Butir		103	100%

Untuk daya pembeda sebagai berikut

Tabel 5. *Hasil Analisis Daya Pembeda*

No.	Kategori	Jumlah	Persentase
11	Jelek	25	24,27%
22	Cukup	8	7,76%
23	Baik	17	16,50%
44	Sangat baik	53	51,46%
Total Butir		103	100%

kemudian dilakukan analisis validitas dan instrumen yang drop atau tidak valid sebanyak 12 butir selanjutnya dilakukan perhitungan koefisien reliabilitas yaitu 0,96.

Pada uji coba kedua dilakukan analisis butir secara kuantitatif soal yang drop atau tidak valid sebanyak 4 butir. Untuk tingkat kesukaran sebagai berikut.

Tabel 6. *Hasil Analisis Taraf Kesukaran Butir*

Untuk daya pembeda sebagai berikut.

No	Kategori	Jumlah	Persentase
11	Sukar	6	9,09%
22	Sedang	59	89,39%
33	Sangat Mudah	1	1,52%
Total Butir		66	100%

Tabel 7. *Hasil Analisis Daya Pembeda*

No.	Kategori	Jumlah	Persentase
11	Jelek	4	6,06%
22	Cukup	7	10,60%
23	Baik	16	24,24%
44	Sangat baik	39	59,09%
Total Butir		66	100%

Kemudian dilakukan analisis validitas dan instrumen yang drop atau tidak valid sebanyak 3 butir selanjutnya dilakukan perhitungan koefisien reliabilitas yaitu 0,93.

Instrumen tes hasil belajar yang dikembangkan pada penelitian ini sebanyak 105 butir instrumen pilihan ganda yang diuji validasi teoretik dan validasi empirik. Pada validasi teoretik instrumen tersebut diberikan kepada 2 orang pakar dan validasi empirik dilakukan uji coba lapangan sebanyak dua kali

Hasil dari penilaian pakar terdapat 2 soal mendapat penilaian B yang berarti butir instrumen disetujui oleh pakar 1 namun tidak disetujui oleh pakar 2. Pakar 2 memberikan nilai dua untuk kedua soal tersebut karena ranah kognitif soal yang dibuat tersebut tidak sesuai dengan indikator, sehingga 2 soal tersebut dibuang. Selanjutnya 103 soal yang mendapat penilaian D yang berarti kedua pakar menyetujui. Nilai koefisien antar kedua pakar berdasarkan hasil analisis Gregory adalah 0,98. Nilai tersebut menandakan bahwa instrumen yang dibuat telah valid secara teoretik sehingga layak untuk diuji cobakan ke tahap selanjutnya. Sebelum diuji cobakan peneliti merevisi beberapa butir instrumen berdasarkan saran dari pakar dan pembimbing. Analisis hasil validasi isi oleh pakar.

Selanjutnya akan dibahas mengenai **analisis validitas empirik** yang telah diuji cobakan sebanyak 103 butir instrumen.

Instrumen kemudian diuji cobakan pada 30 peserta didik di kelas XII.IPA 4 di SMAN 5 Pinrang sebanyak 103 butir instrumen.

Berdasarkan tabel 4 terlihat bahwa 0,97% soal pada kategori sangat mudah, 60,19% soal pada kategori sedang, dan 38,83% soal kategori sukar. Dan daya beda pada tabel 4.3 yaitu 6,06% soal memiliki daya beda jelek, 10,60% soal dengan daya beda cukup, 24,24% soal memiliki daya beda baik, dan 59,09% soal dengan daya beda sangat baik. Sehingga butir yang termasuk dalam kategori jelek dibuang sebanyak 25 butir instrumen. Butir

instrumen yang tidak termasuk pada kategori jelek sebanyak 78 butir yang diujicoba sebanyak 30 responden dengan signifikansi 5%, dari sini di dapat nilai $df = n-2$, $df = 30 - 2 = 28$. Sehingga r product moment dengan signifikansi 5% di dapat nilai r-tabel 0,361. Dengan Hasil uji coba pertama terdapat 37 butir instrumen yang tidak valid 25 termasuk kategori jelek Berdasarkan analisis dari hasil uji coba pertama terdapat 25 butir instrumen pada kategori jelek yaitu butir instrumen nomor 6, 9, 15, 17, 19, 20, 24, 25, 35, 41, 44, 45, 48, 55, 58, 61, 65, 67, 69, 72, 75, 78, 82, 86, 92, 103. dan 12 drop. Berdasarkan analisis nilai korelasi item, terdapat 12 (dua belas) butir instrumen dengan besar indeks korelasi di bawah standar minimal (0,361). Butir yang tidak valid yakni butir instrumen nomor 3, 11, 14, 24, 27, 32, 38, 40, 64, 81, 94, 101. Butir instrumen tersebut masing-masing memiliki koefisien korelasi 0,155; 0,110; 0,202; 0,147; 0,230; 0,319; 0,195; 0,360; 0,355; 0,084; 0,248; 0,334 yang di bawah standar dari nilai valid. Karena nilai tersebut dibawah standar maka butir instrumen tersebut di buang atau tidak digunakan.

Dari 37 butir instrumen tidak valid tersebut secara umum dikarenakan soal terlalu mudah ataupun terlalu sukar. Berdasarkan analisis, peneliti menduga soal yang tidak valid dikarenakan beberapa butir terlalu sukar maupun terlalu mudah sehingga peserta hanya menebak jawaban dan terdapat soal yang terlalu mudah untuk dijawab karena pengecoh pilihannya tidak berfungsi dengan baik.

Setelah analisis validitas butir kemudian dilakukan **analisis reliabilitas** instrumen pada uji coba pertama. dilakukan perhitungan analisis reliabilitas untuk 66 butir instrumen. Koefisien reliabilitas instrumen tes sebesar 0,96.

Kemudian instrumen selanjutnya kembali di uji cobakan di kelas yang berbeda yakni pada kelas XII.IPA 3 di SMAN 5 Pinrang dengan jumlah butir instrumen yang diuji cobakan sebanyak 66 butir. Setelah di ujicobakan kembali ternyata berdasarkan analisis dari hasil uji coba kedua Pada tabel 4.4

terlihat bahwa 1,52% soal pada kategori sangat mudah, 89,39% soal pada kategori sedang, dan 9,09% soal pada kategori sukar. Dan daya pembeda pada tabel 4.5 yaitu 6,06% soal memiliki daya beda jelek, 10,60% soal dengan daya beda cukup, 24,24% soal memiliki daya beda baik, dan 59,09% soal dengan daya beda sangat baik. Pada uji coba kedua terdapat 59 instrumen yang valid dan 7 butir instrumen yang tidak valid Terdapat 4 butir instrumen yang beada pada kategori jelek yaitu butir instrumen nomor 1, 14, 33, 64. dan Butir instrumen yang disusun sebanyak 62 butir yang diujicoba sebanyak 30 responden dengan signifikansi 5%, dari sini di dapat nilai $df = n-2$, $df = 30 - 2 = 28$. Sehingga r product moment dengan signifikansi 5% di dapat nilai r -tabel 0,361. nilai korelasi item, masih terdapat 3 (tiga) butir instrumen dengan besar indeks korelasi di bawah standar minimal (0,361). Butir yang tidak valid yakni butir instrumen nomor 17, 30, 39. beberapa butir instrumen yang di bawah standar tersebut masing-masing memiliki koefisien korelasi 0,060; 0,244; 0,337 yang dibawah standar dari nilai valid. Karena nilai tersebut di bawah standar maka butir instrumen dibuang atau tidak digunakan, sehingga menghasilkan instrumen yang valid dan siap digunakan sebanyak 59 butir instrumen. Pada hasil analisis uji coba kedua terdapat soal yang valid pada uji coba pertama namun pada uji coba kedua soal tersebut menjadi tidak valid sehingga soal tersebut dibuang. Karena tidak dapat memberikan gambaran ketetapan jawaban peserta didik.

Kemudian dilakukan perhitungan analisis reliabilitas untuk 59 butir instrumen. Koefisien reliabilitas instrumen tes sebesar 0,93 Nilai tersebut menunjukkan bahwa instrumen tes yang dikembangkan merupakan instrumen reliabel yang berarti bahwa jika instrumen diuji cobakan lebih dari satu kali pada subyek yang berbeda akan menghasilkan nilai yang relatif sama. Hasil analisis akhir diperoleh 59 butir instrumen yang valid.

Setelah dilakukan analisis kuantitatif instrumen yang dikembangkan oleh peneliti rata-rata merupakan soal yang memiliki tingkat kesukaran yang sedang, Rata-rata butir instrumen yang dikembangkan memiliki daya pembeda yang sangat baik. Pada efektivitas pengecoh semua pengecoh telah berfungsi dengan baik karena tidak terdapat pengecoh yang dipilih kurang dari 2 peserta didik (5%).

Berdasarkan hasil pengembangan instrumen tes hasil belajar fisika dalam ranah kognitif seperti uraian di atas, maka dapat disajikan sebagai berikut.

- a) Hasil uji coba pertama diperoleh :
- 1) Analisis butir secara kuantitatif diperoleh 78 butir dari 103 butir soal yang memenuhi kriteria sebagai butir soal yang “baik”.
 - 2) Analisis validitas item diperoleh 66 dari 78 butir soal yang valid dengan koefisien reliabilitas 0,96.
- b) Hasil uji coba kedua diperoleh:
- 1) Analisis butir secara kuantitatif diperoleh 62 dari 66 butir soal yang memenuhi kriteria sebagai butir soal yang “baik”.
 - 2) Analisis validasi item diperoleh 59 dari 62 butir soal yang valid dengan koefisien reliabilitas 0,93.

SIMPULAN

Hasil pengembangan instrumen tes hasil belajar fisika diberikan kesimpulan sebagai berikut.

1. Instrumen tes hasil belajar kognitif fisika yang dikembangkan untuk peserta didik Kelas XI MIA semester ganjil di SMA Negeri 5 Pinrang, secara teoretik melalui justifikasi pakar telah valid 103 butir instrumen dengan koefisien konsistensi antar pakar 0,98.
2. Hasil pengembangan instrumen tes hasil belajar fisika dalam ranah kognitif ditinjau dari segi empirik diperoleh 59 dari 103 butir soal yang valid dengan reliabilitas 0,93.

DAFTAR RUJUKAN

- Ali, M. S., & Khaeruddin. 2012. *Evaluasi Pembelajaran*. Makassar: Badan Penerbit Universitas Negeri Makassar.
- Arifin, Z. 2012. *Evaluasi Pembelajaran*. Bandung: PT. Remaja Rosdakarya.
- Djali. Pudji Muljono. 2004. *Pegukuran dalam Bidang Pendidikan*. Jakarta: Program Pascasarjana Universitas Negeri Jakarta.
- Djaali, dan Pudji Muljono. 2008. *Pengukuran dalam Bidang Pendidikan*. Jakarta: PT Gramedia Widiasarana Indonesia.
- Gregory, R. J. 2015. *Psychological Testing (History, Principles and Applications)* (7th ed.). United States of America: Pearson.
- Hosnan. 2014. *Pendekatan Saintifik dan Kontekstual dalam Pembelajaran Abad 21*. Bogor: Ghalia Indonesia.
- Kusaeri dan Suprananto. 2012. *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu
- Permendikbud. 2016. *Salinan Standar Penilaian Pendidikan*. Jakarta.
- Rahayu, D. dan Azizah, U. 2012. Pengembangan Instrumen Penilaian Kognitif Berbasis Komputer Dengan Kombinasi Permainan “Who Wants To Be A Chemist” Pada Materi Pokok Struktur Atom Untuk Kelas X SMA RSBI. *Jurnal Prosiding Seminar Nasional Kimia Unesa*.
- Rusman. 2013. *Metode-Metode Pembelajaran: Mengembangkan Profesionalisme Guru*. Jakarta: RajaGrafindo Persada.
- Sugiyono. 2009. *Metode penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.
- Suwarno, W. 2009. *Dasar-dasar Ilmu Pendidikan*. Jokjakarta: Ar-Ruzz Media