

PENGEMBANGAN INSTRUMEN PENILAIAN HASIL BELAJAR BERBASIS TAKSONOMI BLOOM

Syaiful Rohman
(STIT Al-Ibrohimy Bangkalan)

ABSTRAK

Pendidikan di Indonesia diatur oleh Undang-Undang No. 20 Tahun 2003, yang menjadikan pendidikan sebagai faktor penting dalam pembentukan dan peningkatan kualitas kehidupan manusia. Fungsi pendidikan nasional, seperti yang diatur dalam undang-undang, mencakup pengembangan kemampuan, pembentukan karakter, dan peradaban bangsa. Standar Nasional Pendidikan, termasuk Standar Penilaian, diatur untuk memastikan perencanaan, pelaksanaan, dan pengawasan pendidikan yang bermutu.

Pemerintah merumuskan Standar Nasional Pendidikan sebagai dasar dalam perencanaan, pelaksanaan, dan pengawasan pendidikan untuk mencapai mutu pendidikan nasional. Salah satu standar adalah Standar Penilaian, yang bertujuan memastikan perencanaan, pelaksanaan, dan pelaporan hasil penilaian peserta didik yang profesional, terbuka, edukatif, efektif, efisien, dan sesuai dengan konteks sosial budaya.

Namun, di tengah kondisi riil sekolah, terdapat kendala dalam proses penilaian. Guru menganggapnya terlalu rumit dan terfokus pada kompetensi pengetahuan saja. Untuk mengatasi permasalahan ini, langkah-langkah perbaikan kualitas belajar harus mempertimbangkan semua komponen yang saling berhubungan dalam dunia pendidikan.

Selain itu, potensi untuk meningkatkan kualitas pendidikan melibatkan pengembangan instrumen penilaian berbasis taksonomi Bloom revisi. Dalam konteks ini, penelitian tentang "Pengembangan Instrumen Penilaian Hasil Belajar Berbasis Taksonomi Bloom di MI Raudlatul Mubtadiin Ketapang Sampang" menjadi relevan. Penelitian ini diharapkan dapat mengatasi kesulitan guru dalam proses penilaian, dengan fokus pada aspek kognitif, afektif, dan psikomotorik sesuai dengan ranah penilaian Bloom.

Mengembangkan instrumen penilaian yang sesuai dengan taksonomi Bloom revisi akan membantu mengukur berbagai aspek kompetensi peserta didik secara holistik. Langkah-langkah ini penting untuk mencapai tujuan pendidikan nasional, sebagaimana diamanatkan dalam undang-undang, serta untuk memastikan kualitas pembelajaran yang berkualitas.

Kata Kunci : *Instrumen penilaian, Taksonomi bloom, Hasil belajar*

A. Pengembangan Instrumen Penilaian Hasil Belajar

Pengembangan merupakan proses/metode yang digunakan untuk memvalidasi, mengembangkan produk yang sudah ada atau membuat produk yang baru. Richey dan Kelin dalam sugiyono mendeskripsikan pengembangan adalah kajian yang sistematis tentang bagaimana membuat rancangan suatu produk, mengembangkan/memproduksi rancangan tersebut, dan mengevaluasi kinerja produk tersebut, dengan tujuan dapat diperoleh data yang empiris yang dapat digunakan sebagai dasar untuk membuat produk, alat-alat dan model yang digunakan dalam pembelajaran atau nonpembelajaran¹.

Berdasarkan beberapa sumber tersebut tentang pengertian pengembangan, dapat disimpulkan bahwa pengembangan adalah sebuah metode dengan pendekatan ilmiah, gunanya untuk meneliti, merancang, memproduksi dan menguji validitas/reliabilitas produk yang telah dihasilkan.

Instrumen menurut kamus bahasa indonesia adalah alat yang digunakan untuk mengerjakan sesuatu, sarana penelitian, dan orang yang dipakai sebagai alat. Sedangkan kaitannya dengan bidang penelitian, instrumen penelitian adalah merupakan alat ukur seperti tes, kuesioner, pedoman wawancara dan pedoman observasi yang digunakan peneliti untuk mengumpulkan data dalam suatu penelitian.

Dengan demikian, instrumen penilaian hasil belajar adalah alat yang digunakan untuk mengumpulkan, menganalisis, dan menafsirkan data tentang proses atau cara dan hasil belajar peserta didik yang dilakukan secara sistematis, akurat, dan berkesinambungan dengan menggunakan kriteria tertentu sehingga menjadi informasi yang berkaitan dengan pencapaian kompetensi tertentu baik kognitif, psikomotorik, maupun afektif.

Sedangkan menurut Permendikbud No. 104 Tahun 2014, instrumen penilaian adalah alat yang digunakan untuk menilai capaian pembelajaran peserta didik, misalnya: tes dan skala sikap².

Dalam membuat instrumen penilaian perlu memerhatikan karakteristik dari suatu instrumen yang baik. Adapun karakteristik tersebut adalah valid, reliabel, relevan, representatis dan praktis.³

¹ Sugiyono. *Metode Penelitian dan Pengembangan*, Bandung (Alfabeta : 2016). Hlm. 28-29

² Salinan Lampiran Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 104 Tahun 2014 tentang Penilaian Hasil Belajar oleh Pendidik pada Pendidikan Dasar dan Pendidikan Menengah, 3.

³ Moh. Amin, *Validitas Logis dan Validitas Empiris*, diakses dari <http://makalahpendidikanislamlengkap.blogspot.co.id/2015/06/validitas-logis-dan-empiris-evaluasi.html?m=1>, pada tanggal 5 Nopember 2020

B. Instrumen Penilaian Berbasis Taksonomi Bloom

Instrumen penilaian hasil belajar berbasis taksonomi bloom merupakan penilaian yang dilakukan oleh guru untuk mengukur tingkat pencapaian atau penguasaan peserta didik dalam aspek pengetahuan, keterampilan, dan sikap berdasarkan konsep taksonomi bloom. Berikut penjelasan konsep tersebut.

1. Aspek Kognitif

Dalam ranah kompetensi kognitif atau pengetahuan terdapat dua hal yang harus diperhatikan yaitu proses dan dimensi. Untuk proses pengetahuan ada enam tingkat berpikir kognitif menurut Taksonomi Bloom yang telah direvisi Anderson dan Krathwohl, adalah sebagai berikut:⁴

		Proses Kognitif					
		C1 Mengingat	C2 Memahami	C3 Menerapkan	C4 Menganalisis	C5 Mengevaluasi	C6 Mencipta
Dimensi Kognitif	A Faktual	C1 Faktual	C2 Faktual	C3 Faktual	C4 Faktual	C5 Faktual	C6 Faktual
	B Konseptual	C1 Konseptual	C2 Konseptual	C3 Konseptual	C4 Konseptual	C5 Konseptual	C6 Konseptual
	C Prosedural	C1 Prosedural	C2 Prosedural	C3 Prosedural	C4 Prosedural	C5 Prosedural	C6 Prosedural
	D Metaognitif	C1 Metaognitif	C2 Metaognitif	C3 Metaognitif	C4 Metaognitif	C5 Metaognitif	C6 Metaognitif

Tabel 1.1 Matrik Tujuan Pembelajaran Aspek Kognitif

2. Aspek Psikomotorik

Pendidik menilai kompetensi keterampilan melalui penilaian kinerja, yaitu penilaian yang menuntut peserta didik mendemonstrasikan suatu kompetensi tertentu dengan menggunakan praktek, proyek dan penilaian portofolio. Instrumen yang digunakan berupa daftar cek atau skala penilaian (*rating scale*) yang dilengkapi rubrik⁵.

⁴ Imam Gunawan dan Anggarini Retno Palupi, "Taksonomi Bloom – Revisi Ranah Kognitif: Kerangka Landasan untuk Pembelajaran, Pengajaran, dan Penilaian". (Madiun: IKIP PGRI), 11.

⁵ Ibid. 205

3. Aspek Sikap

Penilaian kompetensi sikap melalui observasi, penilaian diri, penilaian teman sejawat oleh peserta didik dan jurnal. Instrumen yang digunakan berupa daftar cek atau skala penilaian yang disertai rubrik, sedangkan untuk jurnal berupa catatan pendidik⁶.

C. Metode Pengembangan

1. Model Pengembangan

Metode ini termasuk dalam Penelitian dan Pengembangan (R&D) dengan mengembangkan instrumen Penilaian Hasil Belajar hasil belajar pada aspek afektif, kognitif, dan psikomotorik yaitu dengan menguji validitas, reliabilitas, dan juga karakteristik instrumen penilaian yang telah dibuat⁷. Dalam penelitian ini, peneliti menggunakan langkah-langkah penelitian dan pengembangan dari Thiagarajan yang meliputi: 1) *Define*, 2) *Design*, 3) *Development*, 4) *Dissemination*, dan disingkat dengan 4D. Dan dimodifikasi hanya 3D, meliputi 1) *Define*, 2) *Design*, 3) *Development*. Metode R&D yang dipilih bersifat kuantitatif, hal ini bertujuan untuk menguji validitas dan reliabilitas instrumen penilaian hasil belajar siswa.

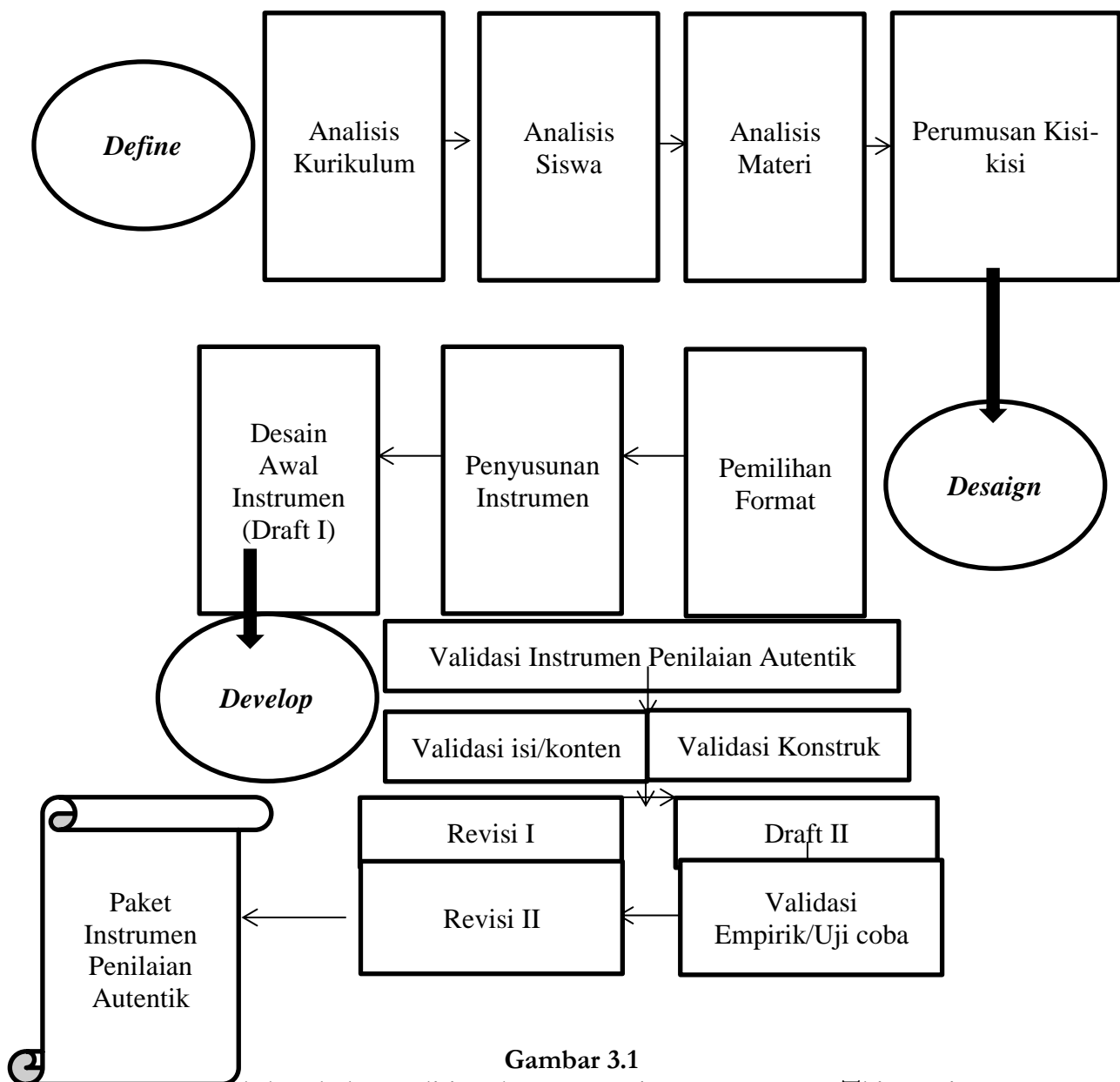
2. Langkah-langkah Penelitian dan Pengembangan

Berdasarkan keterangan tersebut dapat diberikan penjelasan sebagai berikut. *Define* (pendefinisian), analisis kurikulum, analisis siswa, analisis materi dan analisis instrumen Penilaian Hasil Belajar. *Design* (perancangan), bersisi kegiatan untuk membuat rancangan terhadap produk yang telah ditetapkan. Dalam kegiatan ini instrumen Penilaian Hasil Belajar dirancang, baik penilaian yang berupa tes maupun non-tes. *Development* (pengembangan) bersisi kegiatan membuat rancangan menjadi produk, kemudian menguji validitas dan reliabilitas produk yang berupa instrumen Penilaian Hasil Belajar sampai dihasilkan instrumen penilaian yang sesuai dengan spesifikasi yang ditetapkan⁸. Berikut skema tentang langkah-langkah dalam penelitian ini :

⁶ Ibid. 204

⁷ Nana Syaodih S, *Metode Penelitian Pendidikan*, (Bandung: Remaja Rosdakarya, 2012), 164

⁸ Sugiyono. *Metode Penelitian dan Pengembangan*, (Bandung: Alfabeta, 2016), 37-38.



Gambar 3.1
Langkah-langkah penelitian dan pengembangan menurut Thiagarajan.

D. Uji Coba Produk

1. Desain Uji Coba

Dalam proses penelitian dan pengembangan ini, uji cobanya dilakukan dengan menggunakan pendekatan satu kali pengumpulan data (*one-shout case study*)⁹:

X \xrightarrow{O}

Keterangan:

X : pengembangan instrumen Penilaian Hasil Belajar

O : Data yang diperoleh setelah pengembangan instrumen Penilaian Hasil Belajar

⁹ Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Praktek*, (Jakarta: Adi Mahasatya, 2006), 86.

2. Subjek dan Objek Uji Coba

Subjek penelitian ini adalah siswa kelas lima MI Raudlatul Mubtadiin Ketapang.

3. Jenis Data

Di dalam proses penelitian ini sumber datanya berupa instrumen Penilaian Hasil Belajar yang sudah divalidasi oleh ahli validasi. Dan juga data hasil belajar siswa yang mencakup aspek afektif, kognitif, dan psikomotorik.

4. Teknik Pengumpulan Data

Adapun teknik pengumpulan data yang dilakukan dalam penelitian ini antara lain:

a. Catatan Lapangan (*Field Note*)

Catatan lapangan atau *field note* ini dibuat untuk memperoleh data tentang pengembangan instrumen penilaian hasil siswa baik aspek afektif, kognitif, dan psikomotorik.

b. Validasi Ahli

Validasi ahli dilakukan untuk memeriksa butir-butir pertanyaan yang telah disusun apakah butir-butir ini cukup mewakili keseluruhan bidang/variabel yang diteliti¹⁰, dalam hal ini adalah instrumen Penilaian Hasil Belajar.

c. Metode Non Tes Aspek Afektif

Penilaian kompetensi sikap merupakan salah cara mengumpulkan data, cara yang dilakukan melalui observasi, penilaian diri, dan penilaian teman sejawat.

d. Metode Tes Aspek Kognitif

Metode tes digunakan untuk memperoleh data hasil belajar peserta didik.

e. Metode Non Tes Aspek Psikomotorik

Pendidik menilai kompetensi keterampilan melalui praktik, proyek, dan penilaian portofolio.

5. Instrumen Pengumpulan Data

a. Catatan Lapangan (*Field Note*)

Catatan lapangan atau *field note* ini dibuat untuk memperoleh data tentang pengembangan instrumen penilaian hasil belajar siswa baik aspek afektif, kognitif, dan psikomotorik.

b. Lembar Validasi Ahli

Lembar validasi ahli digunakan untuk memperoleh data tentang kevalidan instrumen penilaian.

¹⁰ Sukaesih dan Cahyana, *Metodologi Penelitian Pendidikan*, (jakarta: Rajawali Pers, 2016), 142-143.

c. Lembar Soal Tes

Lembar soal tes sebagai instrumen pengumpul data adalah serangkaian pertanyaan atau latihan yang digunakan untuk mengukur pengetahuan peserta didik.

d. Lembar Non Tes

Lembar non tes sebagai instrumen penilaian hasil belajar peserta didik yang dilakukan tanpa menguji peserta didik, melainkan dengan pengamatan, wawancara, dan meneliti dokumen. Teknik non-tes pada umumnya memegang peranan yang penting dalam rangka mengevaluasi hasil belajar peserta didik dari segi ranah sikap (*affective domain*) dan keterampilan (*psychomotoric domain*).

6. Teknik Analisis Data

Dalam metode penelitian dan pengembangan ini analisis datanya menggunakan analisis data kuantitatif, peneliti hanya melakukan uji validitas dan reliabilitas instrumen Penilaian Hasil Belajar dan analisis data yang dilakukan dalam bagi dalam beberapa tahap, berikut langkah-langkah proses analisis data:

a. *Field Note* (catatan lapangan)

Untuk memperoleh data tentang proses pengembangan instrumen Penilaian Hasil Belajar, peneliti menggunakan *fieldnote* sebagai catatan yang mendeskripsikan tahap-tahap proses pengembangan instrumen ini.

b. Uji Validitas Instrumen

Validitas adalah mengukur apa yang seharusnya diukur Ada tiga macam validitas yaitu, validitas isi, konstruk, dan validitas empirik¹¹.

1. Validitas Isi

Validitas ini disusun untuk mengetahui sampai di mana suatu tes sesuai dengan isi, tujuan, keterampilan yang ingin dicapai. Validitas isi didasarkan pada penilaian ahli.

2. Validitas Konstruk

Disusun berdasarkan dengan teori yang relevan yaitu kesesuaian antara soal dengan level afektif, kognitif, dan psikomotorik. Validitas konstruk ini mengukur sejauh mana dapat diterangkan hubungan antara butir-butir tes dan kelakuan yang bersangkutan. Berikut hasil keputusan dari validitas konstruk.

Tabel 3.1

Kriteria Penilaian Validasi Instrumen Penilaian

¹¹ Sugiyono, *Metode Penelitian & Pengembangan*.....181.

Kode Nilai	Keterangan
A	Dapat digunakan tanpa revisi, jika memperoleh nilai akhir 86-100
B	Dapat digunakan dengan sedikit revisi jika memperoleh nilai akhir 70-85
C	Dapat digunakan dengan banyak revisi jika memperoleh nilai akhir 55-69
D	Tidak dapat digunakan, apabila nilai akhir kurang dari 55

Perhitungan nilai akhir menggunakan rumus:

$$\text{Nilai akhir} = \frac{\text{Jumlah nilai yang diperoleh}}{\text{Jumlah nilai tertinggi}} \times 100$$

3. Validitas Empirik

Validitas empirik ini disusun berdasarkan fakta-fakta uji coba lapangan dan disesuaikan dengan level kompleksitas pada instrumen Penilaian Hasil Belajar. Dalam proses uji validitas ini, ada beberapa teknik atau cara untuk mencari t_{hitung} . Hasil perhitungan t_{hitung} ini dikonsultasikan dengan t_{Tabel} untuk taraf signifikansi $\alpha = 5\%$.

- Jika $t_{hitung} > t_{Tabel}$ maka instrumen tersebut valid.
- Jika $t_{hitung} \leq t_{Tabel}$ maka instrumen tersebut tidak valid.

Uji validitas empirik dilakukan dengan cara menggunakan teknik korelasi *Product Moment* karena data yang diolah berupa data deskriptik atau data dikotomik dengan rumus sebagai berikut:

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

Keterangan : r = koefisien korelasi

$\sum xy$ = jumlah produk x dan y

Sedangkan untuk soal tes uraian menggunakan teknik pearson.

4. Analisis Reliabilitas Instrumen

Reliabilitas adalah derajat konsisten/keajegan suatu instrumen. Instrumen dinyatakan reliabel, apabila instrumen tersebut digunakan untuk

mengukur objek/subjek yang sama oleh orang yang sama atau berbeda dalam waktu yang berbeda, akan menghasilkan data yang relatif sama¹².

Reliabilitas instrumen ada dua yaitu reliabilitas eksternal dan internal. Pengujian reliabilitas eksternal instrumen yang pertama dilakukan dengan *test-retest*, yaitu instrumen digunakan secara berulang-ulang (dalam waktu yang berbeda) untuk mengukur objek yang sama. Bila hasilnya sama maka instrumen tersebut dinyatakan reliabel. Pengujian reliabilitas eksternal instrumen juga dapat juga membuat dua instrumen yang ekuivalen.

Sedangkan reliabilitas internal instrumen merupakan konsistensi antar butir instrumen. Konsistensi internal setiap instrumen dapat dihitung berdasarkan ujicoba satu kali saja¹³. Cara menghitung bisa dengan berbagai cara, dan penelitian ini salah satunya dengan uji reliabilitas koefisien konsisten dengan menggunakan teknik *Kuder-Richardson* yang lebih populer dengan istilah KR_{20} pada soal tes pilihan ganda. Rumus KR_{20} dijelaskan sebagai berikut:

$$KR_{20} = \frac{K}{K-1} \left[\frac{s^2 - \sum p_i \cdot q_i}{s^2_t} \right]$$

Uji reliabilitas KR_{20} ini untuk mencari r_{hitung} . Hasil perhitungan r_{hitung} ini dikonsultasikan dengan r_{Tabel} untuk taraf signifikansi $\alpha = 5\%$. Adapun pengujiannya adalah sebagai berikut:

- a) Jika $r_{hitung} > r_{Tabel}$ maka instrumen tersebut reliabel.
- b) Jika $r_{hitung} \leq r_{Tabel}$ maka instrumen tersebut tidak reliabel.

Sedangkan untuk soal tes uraian uji validitasnya menggunakan teknik *Alpha Cronbach* dengan rumus sebagai berikut:

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum \sigma_b^2}{V_t^2} \right]$$

Dimana:	r_{11}	=	Reliabilitas instrumen
	k	=	Banyaknya butir pertanyaan atau banyaknya soal
	$\sum \sigma_b^2$	=	Jumlah varian butir/item
	V_t^2	=	Varian total

5. Analisis Butir Soal Tes

¹² Ibid, 182

¹³ Ibid, 182

. Tujuan analisis butir soal adalah untuk meningkatkan derajat validitas dan reliabilitas soal secara menyeluruh. Ada empat jenis perhitungan dalam menganalisis butir soal, yaitu:

a) Tingkat Kesukaran

Untuk menghitung tingkat kesukaran soal bentuk objektif pada penelitian dengan cara sebagai berikut:

$$TK = \frac{(WL + WH)}{(nL + nH)} \times 100\%$$

Keterangan :

Tabel 3.2

Perhitungan WL+WH dan WL-WH

WL	Jumlah peserta didik yang menjawab salah dari kelompok bawah
WH	Jumlah peserta didik yang menjawab salah dari kelompok atas
nL	Jumlah kelompok bawah
nH	Jumlah kelompok atas

Sebelum menggunakan rumus di atas, harus terlebih dahulu menempuh langkah-langkah sebagai berikut.

- 1) Menyusun lembar jawaban peserta didik dari skor tertinggi sampai skor terendah.
- 2) Mengambil 27% jawaban tertinggi dan 27% jawaban terendah. Sisa sebanyak 46% disisihkan.
- 3) Membuat tabel untuk masing-masing jawaban kelompok atas dan bawah. Untuk jawaban peserta didik benar diberi tanda (+), jika jawaban salah diberi tanda (-).

b) Daya Pembeda

Perhitungan daya pembeda adalah pengukuran sejauh mana suatu butir soal mampu membedakan peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi berdasarkan kriteria tertentu. Untuk menghitung daya pembeda setiap butir soal dapat digunakan rumus sebagai berikut:

$$DP = \frac{(WL - WH)}{n}$$

Keterangan :

DP = Daya Pembeda.

WL = Jumlah peserta didik yang gagal dari kelompok bawah

WH = Jumlah peserta didik yang gagal dari kelompok atas

$n = 27\% \times N$

Untuk menginterpretasikan koefisien daya pembeda tersebut dapat digunakan kriteria yang dikembangkan oleh Ebel sebagai berikut¹⁴:

<i>Index of diskrimination</i>	<i>Item evaluation</i>
0,40 and up	: <i>Very good item</i>
0,30 – 0,39	: <i>Reasonably good, but possibly subject to improvement</i>
0,20 – 0,29	: <i>Marginal items, usually needing and being subject to improvement</i>
below – 0,19	: <i>Poor item, to be rejected or improved by revision</i>

c) Teknik Analisis Fungsi *Distraktor* (Pengecoh)

Pada tes pilihan ganda untuk setiap butir item yang dikeluarkan dalam tes hasil belajar telah dilengkapi dengan beberapa kemungkinan jawaban, atau yang sering dikenal dengan *option* atau alternatif jawaban. Salah satu di antaranya adalah jawaban betul (= kunci jawaban) sedangkan sisanya adalah jawaban salah. Jawaban salah itulah yang biasa dikenal dengan istilah *distractor* (pengecoh)¹⁵.

E. Proses Pengembangan

1. Deskripsi dan Analisis Data Proses Pengembangan Penilaian Hasil Belajar

Proses pengembangan sampai penyelesaian penelitian ini dilakukan berdasarkan model pengembangan Thiagarajan dengan 3 tahap yaitu *define* (pendefinisian), *design* (perencanaan) dan *develop* (pengembangan). Berikut merupakan analisisnya:

a. *Define* (Pendefinisian)

¹⁴ Mohamad Ali, *Evaluasi Pendidikan*, (Jakarta : Remaja Rosdakarya, 2013), 290-291.

¹⁵ Anas Sudjiono, *Pengantar Evaluasi Pendidikan*. (Jakarta: Raja Grafindo Persada, 2013), 409.

Pada tahapan ini dilakukan analisis kurikulum, analisis siswa, analisis materi dan perumusan indikator pencapaian kompetensi. Keempat analisis ini dapat dijabarkan sebagai berikut:

b. ***Design* (Perancangan)**

1) **Pemilihan Format**

Pemilihan format berfungsi untuk mendesain atau merancang instrumen Penilaian Hasil Belajar. Berdasarkan hasil analisis dari tahap *define* (pendefinisian), kemudian dilanjutkan dengan analisis tentang desain instrumen penilaian yang akan dikembangkan. Analisis ini meliputi analisis instrumen penilaian, perumusan kisi-kisi dan analisis lembar instrumen penilaian.

Deskripsi dan Analisis Validitas Instrumen Penilaian Hasil Belajar

1. Model *Develop* (Pengembangan)

Pada tahap pengembangan ini bersisi kegiatan membuat rancangan menjadi produk yang berupa instrumen Penilaian Hasil Belajar sampai dihasilkan instrumen penilaian yang sesuai dengan spesifikasi yang ditetapkan. Di dalam proses ini ada tiga kegiatan utama, yaitu : 1) uji validitas konstruk/isi dan empirik sesuai dengan rancangan pada bab III, 2) uji reliabilitas, 3) dan analisis butir soal tes. Berikut uraian mengenai dua tahapan tersebut :

a. **Validitas Konstruk/Isi**

Instrumen Penilaian Hasil Belajar hasil belajar aspek afektif, kognitif , dan psikomotorik yang dikembangkan selanjutnya divalidasi kepada ahli validasi. Hal ini untuk mengetahui kelayakan instrumen penilaian untuk digunakan pada langkah selanjutnya. Adapun validator yang dipilih dalam penelitian ini adalah staf pengajar di lingkungan MI Raudlatul Mubtadiin dan salah satu dosen di UINSA.

a) Penilaian Hasil Validasi Aspek Kognitif

Untuk soal tes aspek kognitif instrumen Penilaian Hasil Belajar yang dikembangkan. Butir soal nomor 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, dan 12 masing-masing dengan rata-rata nilai akhir adalah 90. Sesuai pedoman pada bab III, diperoleh bahwa 12 instrumen penilaian dikatakan layak menurut para validator untuk digunakan pada tahap uji coba lapangan.

b) Penilaian Hasil Validasi Instrumen Pengamatan Sikap

Penilaian hasil validasi instrumen pengamatan sikap sosial menurut para ahli secara keseluruhan dapat dilihat bahwa validasi pengamatan sikap sosial memiliki rata-rata nilai 28. Nilai akhir yang diperoleh adalah 86,5 dapat digunakan dengan sedikit revisi.

c) Hasil Validasi Instrumen Penilaian Psikomotorik

Penilaian hasil validasi aspek psikomotorik memiliki rata-rata nilai 26. Nilai akhir yang diperoleh adalah 82,3 dapat digunakan dengan sedikit revisi.

b. Validitas Empirik

Dalam proses uji validitas ini, ada beberapa teknik atau cara untuk mencari t_{hitung} . Hasil perhitungan t_{hitung} ini dikonsultasikan dengan t_{Tabel} untuk taraf signifikansi $\alpha = 5\%$.

- c) Jika $t_{hitung} > t_{Tabel}$ maka instrumen tersebut valid.
- d) Jika $t_{hitung} \leq t_{Tabel}$ maka instrumen tersebut tidak valid.

a) Data Hasil Uji Validitas Soal Tes Pilihan Ganda

Jenis data yang diolah dalam penelitian ini berupa data diskret atau data dikotomik, maka uji validitasnya menggunakan teknik korelasi *product moment*, berikut ini hasil pengolahan dan perhitungan validitas instrumen jenis tes yang berbentuk pilihan ganda dengan menggunakan program microsoft excel.

Tabel 4.43
Hasil Validasi Empirik Untuk Soal Tes Pilihan Ganda

Nomor item soal	T_{hitung}	T_{Tabel}	Interpretasi
1	0,300071	0,404	Tidak valid
2	0,419147	0,404	Valid
3	0,421076	0,404	Valid
4	0,540291	0,404	Valid
5	0,125691	0,404	Tidak valid
6	0,31824	0,404	Tidak valid
7	0,574971	0,404	Valid
8	-0,00267	0,404	Tidak valid
9	-0,33896	0,404	Tidak valid
10	0,165027	0,404	Tidak valid
11	0,469915	0,404	Valid
12	0,421076	0,404	Valid

Dari hasil analisis tersebut di atas, dari 12 butir item yang diuji validitasnya, 6 butir item diantaranya telah dapat dinyatakan sebagai item yang valid, yaitu item nomor 2, 3, 4, 7, 11, dan 12. Sedangkan 6 butir item lainnya, yakni butir item nomor 1, 5, 6, 8, 9, dan 10 merupakan item yang invalid.

b) Data Hasil Uji Validitas Soal Tes Uraian

Selanjutnya untuk validasi soal tes uraian, uji validitasnya menggunakan teknik pearson untuk mencari t_{hitung} . Hasil perhitungan t_{hitung} ini dikonsultasikan dengan t_{Tabel} untuk taraf signifikansi $\alpha = 5\%$.

- a) Jika $t_{hitung} > t_{Tabel}$ maka instrumen tersebut valid.
- b) Jika $t_{hitung} \leq t_{Tabel}$ maka instrumen tersebut tidak valid.

Berikut ini hasil pengolahan dan perhitungan validitas instrumen jenis tes yang berbentuk uraian dengan menggunakan program microsoft excel.

Tabel 4.44
Hasil validasi empirik untuk soal tes uraian

Nomor item soal	T_{hitung}	T_{Tabel}	Interpretasi
1	0.53729	0,404	Valid
2	0.663912	0,404	Valid
3	0.183895	0,404	Tidak valid
4	0.358895	0,404	Tidak valid
5	0.936498	0,404	Valid
6	0.474255	0,404	Valid
7	0.416183	0,404	Valid

Dari hasil analisis tersebut di atas, dari 7 butir item yang diuji validitasnya, 5 butir item diantaranya telah dapat dinyatakan sebagai item yang valid, yaitu item nomor 1, 2, 5, 6, dan 7. Sedangkan 2 butir item lainnya, yakni butir item nomor 3 dan 4 merupakan item yang invalid.

Deskripsi dan Analisis Data Reliabilitas Instrumen Penilaian .

1. Reliabilitas Soal Tes Pilihan Ganda

Uji reliabilitas koefisien konsisten internal ini menggunakan teknik *Kuder-*

Richardson yang lebih populer dengan istilah KR_{20} . Analisis data uji reliabilitas soal tes pilihan ganda akan diuraikan sebagai berikut:

Tabel 4.45
Perhitungan Reliabilitas Koefisien Konsisten Internal

Peserta didik	1	2	3	4	5	6	7	8	9	10	Σ	P	q	pq
Nomor soal														
1.	+	+	+	+	+	+	+	+	+	+	10	1	0	0

2.	+	-	+	-	-	+	-	-	+	+	5	0,5	0,5	0,25
3.	+	+	+	+	+	+	+	+	+	+	10	1	0	0
4.	+	+	+	+	+	-	+	+	+	+	9	0,9	0,1	0,09
5.	+	+	+	-	+	+	+	-	+	-	7	0,7	0,3	0,21
6.	+	+	+	+	+	+	+	+	+	+	10	1	0	0
7.	+	+	+	+	-	+	-	+	+	+	8	0,8	0,2	0,16
8.	+	+	-	+	+	+	+	+	-	+	8	0,8	0,2	0,16
9.	+	+	+	+	+	-	+	+	+	+	9	0,9	0,1	0,09
10.	+	+	+	+	+	+	+	+	+	-	9	0,9	0,1	0,09
11.	+	+	+	+	+	+	+	+	+	+	10	1	0	0
12.	+	+	+	+	-	+	+	+	-	+	8	0,8	0,2	0,16
X	12	11	11	10	9	10	10	10	10	10	103			
X²	144	121	121	100	81	100	100	100	100	100	1067			

Keterangan :

P = proporsi peserta didik yang menjawab betul dari suatu butir soal

q = 1-p

$$s_t^2 = \frac{n \sum X^2 - (\sum X)^2}{n(n-1)} = \frac{12 (1067) - (103)^2}{12 (12-1)} = \frac{12804 - 10609}{132} = \frac{2195}{132} = 16,628$$

k = 12 (jumlah butir soal)

$$\sum p \cdot q = 1,21$$

$$KR_{20} = \frac{K}{K-1} \left[\frac{s^2 - \sum p_i \cdot q_i}{s_t^2} \right] = \frac{12}{12-1} \left[\frac{16,628 - 1,21}{16,628} \right] = 1,09 (0,927) = 1,01$$

Uji reliabilitas dalam penelitian dengan KR_{20} ini untuk mencari r_{hitung} . Hasil perhitungan r_{hitung} ini dikonsultasikan dengan r_{Tabel} untuk taraf signifikansi $\alpha = 5\%$.

Adapun pengujiannya adalah sebagai berikut:

a) Jika $r_{hitung} > r_{Tabel}$ maka instrumen tersebut reliabel.

b) Jika $r_{hitung} \leq r_{Tabel}$ maka instrumen tersebut tidak reliabel.

Dari hasil perhitungan tersebut diperoleh nilai $r_{hitung} = 1,01$. Sedangkan untuk nilai distribusi nilai pada $r_{Tabel} = 0,632$. Jadi kesimpulannya $r_{hitung} > r_{Tabel}$ dan instrumen tes tersebut reliabel.

1. Reliabilitas Soal Tes Uraian

Selanjutnya untuk tes uraian uji reliabilitasnya menggunakan Alfa Cronbach untuk mencari r_{hitung} . Hasil perhitungan r_{hitung} ini dikonsultasikan dengan r_{Tabel} untuk taraf signifikansi $\alpha = 5\%$. Untuk uji reliabilitas ini, yang diuji hanya soal valid yang tidak valid dibuang. Adapun pengujiannya adalah sebagai berikut:

- a) Jika $r_{hitung} > r_{Tabel}$ maka instrumen tersebut reliabel.
- b) Jika $r_{hitung} \leq r_{Tabel}$ maka instrumen tersebut tidak reliabel.

Dalam menguji reliabilitas digunakan uji konsistensi internal dengan menggunakan rumus *Alpha Cronbach* sebagai berikut.

Table 4.45
Jumlah varian tiap item soal dengan menggunakan excel:

No	Varian	Hasil
1	Item 1	4.6875
2	Item 2	4.1232
3	Item 5	24.305
4	Item 6	3.4722
5	Item 7	3.4722

rumus *Alpha Cronbach*.

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum \sigma_b^2}{V_t^2} \right], \text{ (Arikunto, 1999: 193)}$$

Dimana: r_{11} = reliabilitas instrumen

k = banyaknya butir pertanyaan atau banyaknya soal

$\sum \sigma_b^2$ = jumlah varian butir/item

V_t^2 = varian total

Kriteria suatu instrumen penelitian dikatakan reliabel dengan menggunakan teknik ini, bila koefisien reliabilitas (r_{11}) $> 0,6$.

$$k = 5$$

$$\sum \sigma_b^2 = 40,0608$$

$$V_t^2 = 81,2066$$

Dari hasil perhitungan tersebut diperoleh nilai $r_{hitung} = 1,2499$. Sedangkan untuk nilai distribusi nilai pada $r_{Tabel} = 0,632$. Jadi kesimpulannya $r_{hitung} > r_{Tabel}$ dan instrumen tes tersebut reliabel.

A. Analisis Butir Soal

Derajat validitas dan reliabilitas sangat bergantung kepada karakteristik soal-soalnya. Jika soal-soal itu baik, berarti validitas dan reliabilitas soal tersebut tinggi. Tentu tidak semua soal yang dikonstruksi baik. Oleh sebab itu, perlu dianalisis butir-butir soalnya, sehingga dapat diketahui soal-soal mana yang akan diperbaiki, diseleksi, direvisi, atau diganti. Tujuan analisis butir soal adalah untuk meningkatkan derajat validitas dan reliabilitas soal secara menyeluruh. Ada tiga jenis perhitungan dalam menganalisis butir soal, yaitu:

1. Tingkat Kesukaran

Untuk menghitung tingkat kesukaran soal bentuk objektif pada penelitian dengan cara sebagai berikut:

$$TK = \frac{(WL + WH)}{(nL + nH)} \times 100\%$$

Keterangan :

WL	Jumlah peserta didik yang menjawab salah dari kelompok bawah
WH	Jumlah peserta didik yang menjawab salah dari kelompok atas
nL	Jumlah kelompok bawah
nH	Jumlah kelompok atas

Jadi, tingkat kesukaran setiap soal adalah sebagai berikut :

1. Soal nomor 1 : $TK = 1/14 \times 100 = 7\%$
2. Soal nomor 2 : $TK = 8/14 \times 100 = 57\%$
3. Soal nomor 3 : $TK = 1/14 \times 100 = 7\%$
4. Soal nomor 4 : $TK = 1/14 \times 100 = 7\%$
5. Soal nomor 5 : $TK = 5/14 \times 100 = 36\%$
6. Soal nomor 6 : $TK = 1/14 \times 100 = 7\%$
7. Soal nomor 7 : $TK = 2/14 \times 100 = 14\%$
8. Soal nomor 8 : $TK = 3/14 \times 100 = 21\%$
9. Soal nomor 9 : $TK = 3/14 \times 100 = 21\%$
10. Soal nomor 10 : $TK = 2/14 \times 100 = 14\%$
11. Soal nomor 11 : $TK = 4/14 \times 100 = 28\%$
12. Soal nomor 12 : $TK = 4/14 \times 100 = 28\%$

Adapun kriteria penafsiran tingkat kesukaran soal adalah :

- a. Jika jumlah persentase sampai dengan 27% termasuk mudah
- b. Jika jumlah persentase 28% - 72% termasuk sedang
- c. Jika jumlah persentase 73% ke atas termasuk sukar

2. Daya Pembeda

Soal pilihan ganda ini diambil dari tema peristiwa dalam kehidupan di kelas V, dengan jumlah butir soal yang dipilih sebanyak 12 yang disesuaikan dengan mata pelajaran dan telah dirumuskan ke dalam tema tersebut.

$$DP = \frac{(WL-WH)}{n}$$

Keterangan :

DP = daya pembeda.

WL = jumlah peserta didik yang gagal dari kelompok bawah

WH = jumlah peserta didik yang gagal dari kelompok atas

$n = 27\% \times N$

Diketahui :

Jumlah peserta didik (N) = 24

Jumlah sampel (n) = $27\% \times 24 = 6,5 = 7$ (dibulatkan)

WL = 23

WH = 12

Jadi, Daya Pembeda (DP) = $\frac{23-12}{7} = 1,57$

Untuk menginterpretasikan koefisien daya pembeda tersebut dapat digunakan kriteria yang dikembangkan oleh Ebel sebagai berikut:

<i>Index of diskrimination</i>	<i>Item evaluation</i>
0,40 and up	: <i>Very good item</i>
0,30 – 0,39	: <i>Reasonably good, but possibly subject to improvement</i>
0,20 – 0,29	: <i>Marginal items, usually needing and being subject to improvement</i>
below – 0,19	: <i>Poor item, to be rejected or improved by revision</i>

Berdasarkan kriteria di atas, nilai DP = 1,57 termasuk soal yang sangat baik. Ini berarti bahwa soal tersebut dapat membedakan antara peserta didik yang sudah menguasai kompetensi dengan peserta didik yang belum/kurang menguasai kompetensi.

3. Teknik Analisis Fungsi *Distraktor* (Pengecoh)

Dalam proses analisis fungsi pengecoh pada bentuk soal pilihan ganda dengan item sebanyak 12 butir soal yang diujikan ke 24 siswa, dimana setiap butir dilengkapi dengan empat alternatif, yaitu A, B, C, dan D. Selanjutnya diperoleh pola penyebaran jawaban item sebagai berikut:

1. Untuk item nomor 1, kunci jawabannya adalah A, pengecohnya adalah B, C, D. Pengecoh B dipilih 2 anak; $2/24 \times 100 = 8,3\%$ (telah berfungsi dengan baik $8,3\% > 5\%$). C 1; $1/24 \times 100 = 4,2\%$ (belum berfungsi) D 1; $1/24 \times 100 = 4,2$ (belum berfungsi $0\% < 5\%$).
2. Untuk item nomor 2, kunci jawabannya adalah B, pengecohnya adalah A, C, D. Pengecoh A dipilih 9 anak; $9/24 \times 100 = 37,5\%$ (telah berfungsi dengan baik $37,5\% > 5\%$). C 1; $1/24 \times 100 = 4,2\%$ (belum berfungsi dengan baik $0\% < 5\%$). dan D 0; $0/24 \times 100 = 0\%$ (jelek $0\% < 5\%$).
3. Item nomor 3, kunci jawaban A, pengecoh adalah B, C, dan D. Pengecoh B dipilih 0; $0/24 \times 100 = 0\%$ (belum berfungsi $0 < 5\%$), pengecoh C 3 anak; $3/24 \times 100 = 12,5\%$ (telah berfungsi dengan baik $12,5\% > 5\%$). D 1; $1/24 \times 100 = 4,2\%$ (belum berfungsi $4,2\% < 5\%$).
4. Item nomor 4, kunci jawaban A, pengecoh adalah B, C, dan D. Pengecoh B dipilih 0; $0/24 \times 100 = 0\%$ (belum berfungsi $0 < 5\%$), pengecoh C 2 anak; $2/24 \times 100 = 8,3\%$ (telah berfungsi dengan baik $12,5\% > 5\%$). D 1; $1/24 \times 100 = 4,2\%$ (belum berfungsi $4,2\% < 5\%$).
5. Item nomor 5, kunci jawaban A, pengecoh adalah B, C, dan D. Pengecoh B dipilih 3; $3/24 \times 100 = 12,5\%$ (telah berfungsi $12,5 > 5\%$), pengecoh C 2 anak; $2/24 \times 100$

- = 8,3 % (telah berfungsi dengan baik 8,3 % > 5%). D 2; $2/24 \times 100 = 8,3\%$ (telah berfungsi dengan baik 8,3% > 5%).
6. Item nomor 6, kunci jawaban A, pengecoh adalah B, C, dan D. Pengecoh B dipilih 0; $0/24 \times 100 = 0\%$ (belum berfungsi $0 < 5\%$), pengecoh C 6 anak; $6/24 \times 100 = 25\%$ (telah berfungsi dengan baik $25\% > 5\%$). D 1; $1/24 \times 100 = 4,2\%$ (belum berfungsi $4,2\% < 5\%$).
 7. Item nomor 7, kunci jawaban C, pengecoh adalah A, B, dan D. Pengecoh A dipilih 2; $2/24 \times 100 = 8,3\%$ (telah berfungsi dengan baik $8,3\% > 5\%$), pengecoh B 2 anak; $2/24 \times 100 = 8,3\%$ (telah berfungsi dengan baik $8,3\% > 5\%$). D 3; $3/24 \times 100 = 12,5\%$ (telah berfungsi dengan baik $12,5\% > 5\%$).
 8. Item nomor 8, kunci jawaban C, pengecoh adalah A, B, dan D. Pengecoh A dipilih 0; $0/24 \times 100 = 0\%$ (belum berfungsi $0 < 5\%$), pengecoh B 4 anak; $4/24 \times 100 = 16,6\%$ (telah berfungsi dengan baik $16,6\% > 5\%$). D 3; $3/24 \times 100 = 12,5\%$ (telah berfungsi dengan baik $12,5\% > 5\%$).
 9. Item nomor 9, kunci jawaban D, pengecoh adalah A, B, dan C. Pengecoh A dipilih 5; $5/24 \times 100 = 20,8\%$ (telah berfungsi $20,8 > 5\%$), pengecoh B 1 anak; $1/24 \times 100 = 4,2\%$ (belum berfungsi karena $4,2\% < 5\%$). D 3; $3/24 \times 100 = 12,5\%$ (telah berfungsi dengan baik $12,5\% > 5\%$).
 10. Item nomor 10, kunci jawaban A, pengecoh adalah B, C, dan D. Pengecoh B dipilih 4 ; $4/24 \times 100 = 16,6\%$ (telah berfungsi dengan baik $16,6\% > 5\%$), pengecoh C 1 anak; $1/24 \times 100 = 4,2\%$ (belum berfungsi karena $4,2\% < 5\%$). D 3; $3/24 \times 100 = 12,5\%$ (telah berfungsi dengan baik $12,5\% > 5\%$).
 11. Item nomor 11, kunci jawaban C, pengecoh adalah A, B, dan D. Pengecoh A dipilih 1 ; $1/24 \times 100 = 4,2\%$ (belum berfungsi $4,2\% < 5\%$), pengecoh B 0 anak; $0/24 \times 100 = 0\%$ (belum berfungsi karena $0\% < 5\%$). D 4; $4/24 \times 100 = 16,6\%$ (telah berfungsi dengan baik $16,6\% > 5\%$).
 12. Item nomor 12, kunci jawaban C, pengecoh adalah A, B, dan D. Pengecoh A dipilih 1 ; $1/24 \times 100 = 4,2\%$ (belum berfungsi $4,2\% < 5\%$), pengecoh B 3 anak; $3/24 \times 100 = 12,5\%$ (telah berfungsi karena $12,5\% > 5\%$). D 4; $2/24 \times 100 = 8,3\%$ (telah berfungsi dengan baik $8,3\% > 5\%$).

F. Kesimpulan

Berdasarkan penelitian dan pengembangan instrumen penilaian hasil belajar dapat diambil kesimpulan bahwa proses pengembangan instrumen penilaian hasil belajar dikembangkan dengan menggunakan model pengembangan Thiagarajan yang terdiri 3 tiga tahap, yaitu: (a) *Define* (pendahuluan) dilakukan dengan analisis kurikulum, analisis siswa,

analisis materi dan perumusan tujuan pencapaian kompetensi. (b) *Design* (merancang) yakni dengan pemilihan format instrumen penilaian autentik. (c) *Develop* (pengembangan) meliputi proses validasi, reliabilitas dan analisis butir soal.

Uji validitas isi dan konstruk instrumen penilaian autentik aspek afektif, kognitif, dan psikomotorik dinyatakan valid dan layak digunakan. Sedangkan validitas empirik untuk soal tes pilihan ganda yang berjumlah 12 item soal, item 2,3,4,7,11,12 dinyatakan valid. Dan untuk item soal 1,5,6,8,9,10 dinyatakan tidak valid. Dan untuk soal tes uraian yang berjumlah 7 item soal, item soal 1,2,5,6,7 dinyatakan valid, dan item soal 3 dan 4 dinyatakan tidak valid.

Reliabilitas instrumen penilaian soal tes pilihan ganda dengan menggunakan teknik KR_{20} dinyatakan reliabel dengan nilai $r_{hitung} = 1,01$. Sedangkan untuk nilai distribusi nilai pada $r_{Tabel} = 0,632$. Jadi kesimpulannya $r_{hitung} > r_{Tabel}$ dan instrumen tes tersebut reliabel. Sedangkan untuk reliabilitas soal tes uraian dengan rumus *Alpha Cronbach* dinyatakan reliabel. Dari hasil perhitungan tersebut diperoleh nilai $r_{hitung} = 1,2499$. Sedangkan untuk nilai distribusi nilai pada $r_{Tabel} = 0,632$. Jadi kesimpulannya $r_{hitung} > r_{Tabel}$ dan instrumen tes tersebut reliabel.

Untuk karakteristik analisis butir soal tes yang meliputi tingkat kesukaran, daya pembeda, dan teknik analisis fungsi *distractor* (pengecoh) diperoleh hasil sebagai berikut: a) Tingkat kesukaran 12 item soal ada dalam kategori sedang untuk 4 item soal, dan mudah untuk 8 item soal. b) Hasil perhitungan daya Pembeda analisis butir soal adalah 1,57. Berdasarkan kriteria di atas, nilai DP = 1,57 termasuk soal yang sangat baik. c) Fungsi pengecoh pada bentuk soal pilihan ganda rata-rata di atas 5% dan dinyatakan berfungsi dengan baik.

G. Saran

1. Bagi peneliti yang akan mengembangkan penelitian ini, disarankan melakukan penelitian lanjutan mengenai penerapan dan pengaruh instrumen penilaian tersebut terhadap kemampuan tertentu.
2. Instrumen penilaian ini hendaknya diujicobakan pada kelas lain atau sekolah sekolah lain sehingga diperoleh Instrumen penilaian yang jauh lebih baik.

H. Daftar Pustaka

- Abdullah, Ridwan S, *Pembelajaran Sainifik untuk Implementasi Kurikulum 2013*, (Jakarta: Bumi Aksara, 2014)
- B Uno, Hamzah dan Satria Koni, *Assessment Pembelajaran*, (Jakarta: Bumi Aksara, 2014)
- Emzir, *Metodologi Penelitian Pendidikan*, (Jakarta: Rajawali Pers, 2012)
- Kamus Besar Bahasa Indonesia Dalam Tesis Hernik Ferisia, *Pengembangan Modul Kompetensi Berbicara Mahasiswa Prodi PGMI LAIN Sunan Ampel Surabaya*
- Kurniasih ,Imas, dan Sani , Berlin. 2014. *Implementasi kurikulum 2013 Konsep dan penerapan*, Surabaya : Kata Pena.
- Mungin,Burhan. 2011. *Metodologi Penelitian Kuantitatif: Komunikasi, Ekonomi, dan Kebijakan Publik serta Ilmu-Ilmu Sosial Lainnya*. Jakarta: Kencana Prenada Media Group.
- Maman Abdullah, *Sistem Penilaian dalam Kurikulum 2013: Kajian Dokumen terhadap Kurikulum 2013*, Diakses pada 18 maret 2020.
- Mardapi, Djemari. 2012. *Pengukuran Penilaian dan evaluasi Pendidikan*. Yogyakarta: Nuha Litera.
- Modul KEMENAG Kab. Sidoarjo, BIMTEK Implementasi Kurikulum 2013
- Muhammad Surya , Djumhur. *Bimbingan dan Penyuluhan di Sekolah Jilid 3* (Bandung: Erlangga, 1976)
- Sudjana , Nana, *Penilaian Hasil Proses Belajar Mengajar*, (Bandung: Rosdakarya, 1999)
- Sugiyono, *Metode Penelitian Pendidikan dalam tesis mashajir model pengembangan penilaian LAIN Sunan Ampel Surabaya*
- Trianto, *Pengantar Penelitian Pendidikan bagi Pengembang*,(Jakarta: Prenada Media Group, 2011)
- Sugiyono. *Metode Penelitian Kuantitatif Kualitatif dan R & D* (Bandung: Alfabeta. 2011)
- Sanjaya, Wina, *Penelitian Pendidikan*, (Jakarta: Prenada Media Group, 2014)
- Syaodih S , Nana, *Metode Penelitian Pendidikan*, (Bandung: Remaja Rosdakarya, 2012)
- Selly Rahmawati dan Sunarti *penilaian dalam kurikulum 2013: Membantu Guru dan Calon Guru mengetahui Langkah-langkah Penilaian Pembelajaran*. Yogyakarta: Penerbit Andi.
- Thiagarajan, S., Semmel, D.S. & Semmel, M.I., *Instructional Development for Training Teachers of Exceptional Children*. Indiana: Indiana University Bloomington, 1974.
- Thobroni, *Belajar dan Pembelajaran: Teori dan Praktik*. Ar-Ruzz Media: Yogyakarta, 2015.