

BAB III

METODE PENELITIAN

Bab ini berisi penjabaran rinci mengenai metode penelitian pengembangan instrumen penilaian otentik untuk mengukur keterampilan proses sains pada pembelajaran reaksi eksoterm dan reaksi endoterm, yang termasuk di dalamnya menjabarkan lokasi dan subjek populasi/ sampel, desain penelitian dan metode penelitian, definisi operasional, instrumen penelitian, proses pengembangan instrumen dan teknik pengumpulan data dan analisis data.

A. Lokasi dan Subjek Populasi Sampel

Penelitian dilakukan di kelas XI IPA salah satu SMA Negeri kota Bandung dengan menggunakan 38 orang peserta didik sebagai sampel (responden) untuk memperoleh data validitas empiris, reliabilitas, dan data hasil analisis keterampilan proses sains yang terukur melalui instrumen penilaian otentik yang telah dikembangkan. Penelitian dilaksanakan selama dua jam pelajaran (penelitian melibatkan pembelajaran). Instrumen penilaian yang dikembangkan diantaranya adalah instrumen penilaian tertulis yang berupa tes uraian terbatas dengan tiga butir soal uraian dan instrumen penilaian kinerja dengan sembilan *task* kerja utama. Instrumen penilaian otentik yang dikembangkan digunakan setelah sampel (responden) mendapatkan materi reaksi eksoterm dan reaksi endoterm. Instrumen penilaian uraian yang otentik digunakan oleh seluruh sampel (38 orang responden), sedangkan penilaian kinerja digunakan oleh 33 orang responden yang dikelompokkan menjadi tujuh kelompok kerja (setiap kelompok kerja terdiri dari empat sampai lima orang responden) dan lima orang responden sisanya melakukan kerja secara individu yang dinilai langsung oleh lima orang *rater* untuk menilai keterampilan dalam melakukan praktikum reaksi eksoterm dan reaksi endoterm. Lima sampel dipisahkan dari sampel lain dan dinilai langsung oleh lima *rater* agar diperoleh data untuk menentukan reliabilitas instrumen penilaian otentik kinerja.

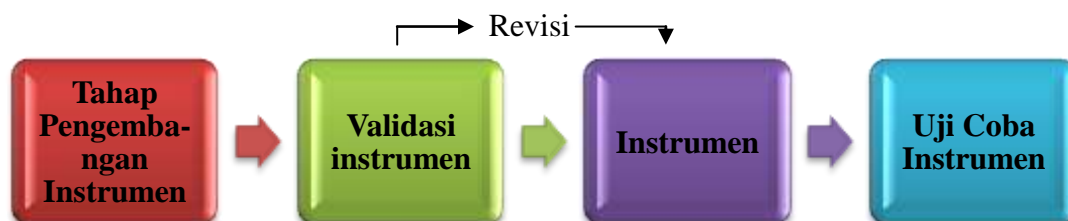
B. Metode dan Desain Penelitian

Sugiyono (2014, hlm. 3) mengungkapkan bahwa metode penelitian dapat diartikan sebagai cara ilmiah untuk mendapatkan informasi atau data dengan tujuan tertentu. Metode penelitian yang digunakan pada penelitian ini adalah metode *research and development (R&D)* yang merupakan metode penelitian dan pengembangan untuk menghasilkan suatu produk atau menyempurnakan produk yang sudah ada sebelumnya. Metode penelitian R&D merupakan metode multi tahap, dimulai dengan *research* yang dilakukan untuk mendapatkan informasi

tentang kebutuhan pengguna dan dilanjutkan dengan *development* untuk menghasilkan perangkat pembelajaran.

Tahapan penelitian R&D ini diadopsi dari Borg & Gall yang terdiri dari: 1) Potensi dan masalah, 2) Pengumpulan data, 3) Desain produk, 4) Validasi desain, 5) Revisi desain, 6) Uji coba produk, 7) Revisi produk I, 8) Uji coba pemakaian (aplikasi), 9) Revisi produk setelah uji aplikasi, 10) Produk masal (Sugiyono, 2014, hlm. 409). Tetapi pada penelitian ini tahap yang dilaksanakan hanya sampai uji coba produk.

Desain penelitian pengembangan instrumen penilaian otentik digambarkan pada gambar di bawah ini:



Gambar 3.1. Desain Penelitian Pengembangan Instrumen Penilaian Otentik

C. Definisi Operasional

Adapun definisi operasional dari istilah-istilah berikut:

1. Penilaian Otentik merupakan penilaian yang digunakan untuk memantau perkembangan kompetensi peserta didik yang sebenarnya (nyata) selama kegiatan pembelajaran dan bukan sesuatu yang dibuat-buat (Arifin, 2009).
2. Keterampilan merupakan kecakapan seseorang dalam menggunakan alat-alat laboratorium dan kecakapan menyelesaikan tugas yang diberikan.
3. Keterampilan proses sains ialah serangkaian kemampuan, yang berguna dalam banyak disiplin ilmu dan merupakan gambaran dari perilaku saintis. Keterampilan ini digunakan ketika seseorang melakukan sains (SAPA, dalam Rezba, 2002).

Keterampilan proses sains juga dapat dikatakan sebagai keterampilan *ber-inquiry* (menurut *Elementary Science Syllabus*) yang digunakan dalam keseharian untuk mempelajari pelajaran apa pun (Carin, 1997).

4. Keterampilan proses sains yang terukur dan teramati pada penelitian ini antara lain:
 - a. Menerapkan konsep : menggunakan konsep yang telah dikuasai untuk memecahkan suatu permasalahan atau untuk menjelaskan suatu peristiwa.
 - b. *Inferring* : memberikan penjelasan untuk objek atau kejadian khusus berdasarkan bukti dan apa yang telah diketahui.
 - c. Menafsirkan data : menganalisis data yang diperoleh dan menyatakan pola hubungan satu data dengan data lain
 - d. Menyimpulkan : menarik suatu kesimpulan dari data hasil percobaan
 - e. Menggunakan alat dan bahan: menggunakan alat dan bahan selama kegiatan ilmiah berlangsung
 - f. Mengamati : Menggunakan panca indera
 - g. Mengukur : Melakukan pengamatan kuantitatif dengan menggunakan suatu alat untuk mengumpulkan data
 - h. Mencatat atau merekam data : Mengumpulkan sedikit informasi mengenai objek atau kejadian yang menggambarkan situasi spesifik
 - i. Mengomunikasi : Menghimpun informasi dari grafik atau gambar dan menggunakan kata-kata (baik lisan maupun tulisan) untuk mendeskripsikan suatu objek, kejadian atau tindakan.

D. Instrumen Penelitian

Penelitian pengembangan instrumen penilaian otentik yang dapat mengukur keterampilan proses sains ini dilakukan melalui tahap pengembangan, tahap validasi dan tahap uji coba. Instrumen penelitian pada tahap pengembangan dan validasi adalah lembar validasi instrumen (format validasi), sedangkan instrumen penelitian pada tahap uji coba adalah instrumen penilaian kinerja (lembar penilaian kinerja) dan instrumen penilaian tertulis (uraian terbatas).

Lana Tantriasa DT, 2014

Pengembangan Instrumen Penilaian Otentik Untuk Mengukur Keterampilan Proses Sains Pada Pembelajaran Termokimia

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

1. Lembar Validasi Instrumen

Lembar validasi instrumen ini dibagi menjadi lembar validasi instrumen penilaian uraian terbatas dan lembar validasi instrumen penilaian kinerja. Lembar validasi instrumen penilaian uraian terbatas digunakan untuk melihat kesesuaian butir soal yang dikembangkan dengan indikator pembelajaran dan kesesuaian butir soal dengan aspek keterampilan proses sains yang terukur. Lembar validasi instrumen penilaian kinerja digunakan untuk melihat kesesuaian *task* dengan rubrik dan kesesuaian *task* dengan aspek keterampilan proses sains yang terukur.

Pengujian validitas instrumen sebelum diuji cobakan, dilakukan melalui analisis rasional atau melalui *professional/ expert judgement* (Azwar, dalam Susetyo, 2011, hlm.89). Ahli yang menjadi validator instrumen penilaian otentik terdiri dari lima orang ahli di bidang *assessment* dan bidang kimia. Format validasi instrumen penilaian uraian terbatas berisi daftar cek (setuju/tidak) kesesuaian butir soal dengan indikator dan daftar cek (setuju/tidak) kesesuaian butir soal dengan aspek keterampilan proses sains yang terukur pada masing-masing butir soal serta dilengkapi kolom catatan perbaikan yang digunakan untuk catatan ahli dalam perbaikan instrumen. Format validasi instrumen penilaian kinerja berisi daftar cek (setuju/tidak) kesesuaian *task* dengan deskriptor atau rubrik dan daftar cek (setuju/tidak) kesesuaian *task* dengan aspek keterampilan proses sains yang terukur pada masing-masing *task* serta dilengkapi kolom catatan perbaikan yang digunakan untuk catatan ahli dalam perbaikan instrumen.

2. Instrumen Penilaian Kinerja

Penilaian kinerja merupakan kegiatan memantau kinerja seseorang dalam menyelesaikan tugas atau permasalahan menggunakan berbagai bentuk instrumen untuk memperoleh informasi tentang kinerja seseorang yang dipantau. Instrumen penilaian kinerja dikembangkan untuk mengukur keterampilan proses sains siswa. Instrumen penilaian kinerja ini berisi rincian tugas atau *task* yang harus dilakukan peserta didik dan dilengkapi deskriptor atau rubrik untuk setiap *task*-nya.

Instrumen penilaian kinerja ini dikembangkan dalam bentuk lembar pengamatan kinerja dengan skala penilaian 0 (untuk kinerja yang dianggap paling buruk) sampai 2 (untuk kinerja yang dianggap paling ideal). Terdapat 11 tugas yang dikembangkan dengan tiga buah rubrik penilaian pada setiap tugasnya. Tugas-tugas yang dikembangkan tersebut divalidasi dan diuji reliabilitasnya sehingga diperoleh instrumen penilaian kinerja yang berisi sembilan butir tugas dengan tiga rubrik penilaian pada setiap tugasnya.

3. Instrumen Penilaian Tertulis (Uraian Terbatas)

Instrumen penilaian tertulis digunakan untuk mengetahui keterampilan proses sains yang dimiliki siswa meskipun dalam bentuk *paper and pencil test*. Setiap butir soal yang dikembangkan mengacu pada klasifikasi indikator keterampilan proses sains. Bentuk instrumen penilaian tertulis ini adalah tes uraian terbatas yang menuntut kemampuan seseorang untuk menyampaikan, menyusun, memadukan gagasan yang dimilikinya dengan menggunakan kata-katanya sendiri. Instrumen penilaian ini terdiri dari tiga butir soal dengan sepuluh pertanyaan beserta pedoman penskoran setiap butir soalnya.

E. Prosedur Penelitian

1. Proses Pengembangan Instrumen

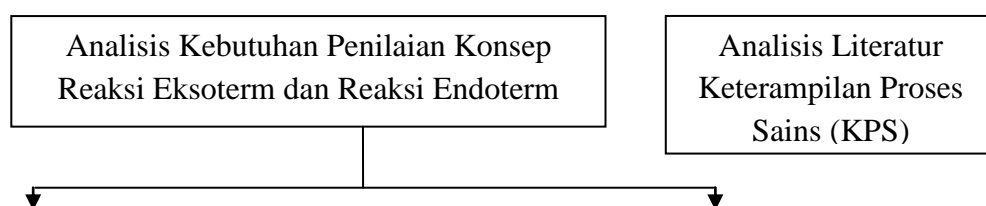
Proses pengembangan instrumen penilaian otentik yang dilakukan pada penelitian ini disesuaikan dengan tahap pengembangan metode *research and development* yaitu dengan melakukan analisis kebutuhan untuk mengembangkan indikator yang diturunkan dari kompetensi inti dan kompetensi dasar serta analisis keterampilan proses sains yang dikembangkan oleh beberapa ahli, membuat kisi-kisi instrumen penilaian dari indikator yang telah dikembangkan, mengembangkan soal tes uraian dan tugas kinerja berdasarkan kisi-kisi yang dibuat dan keterampilan proses sains yang telah dianalisis, serta mengembangkan *task* dan *rubric* penilaian untuk tes kinerja dan pedoman penskoran untuk tes uraian.

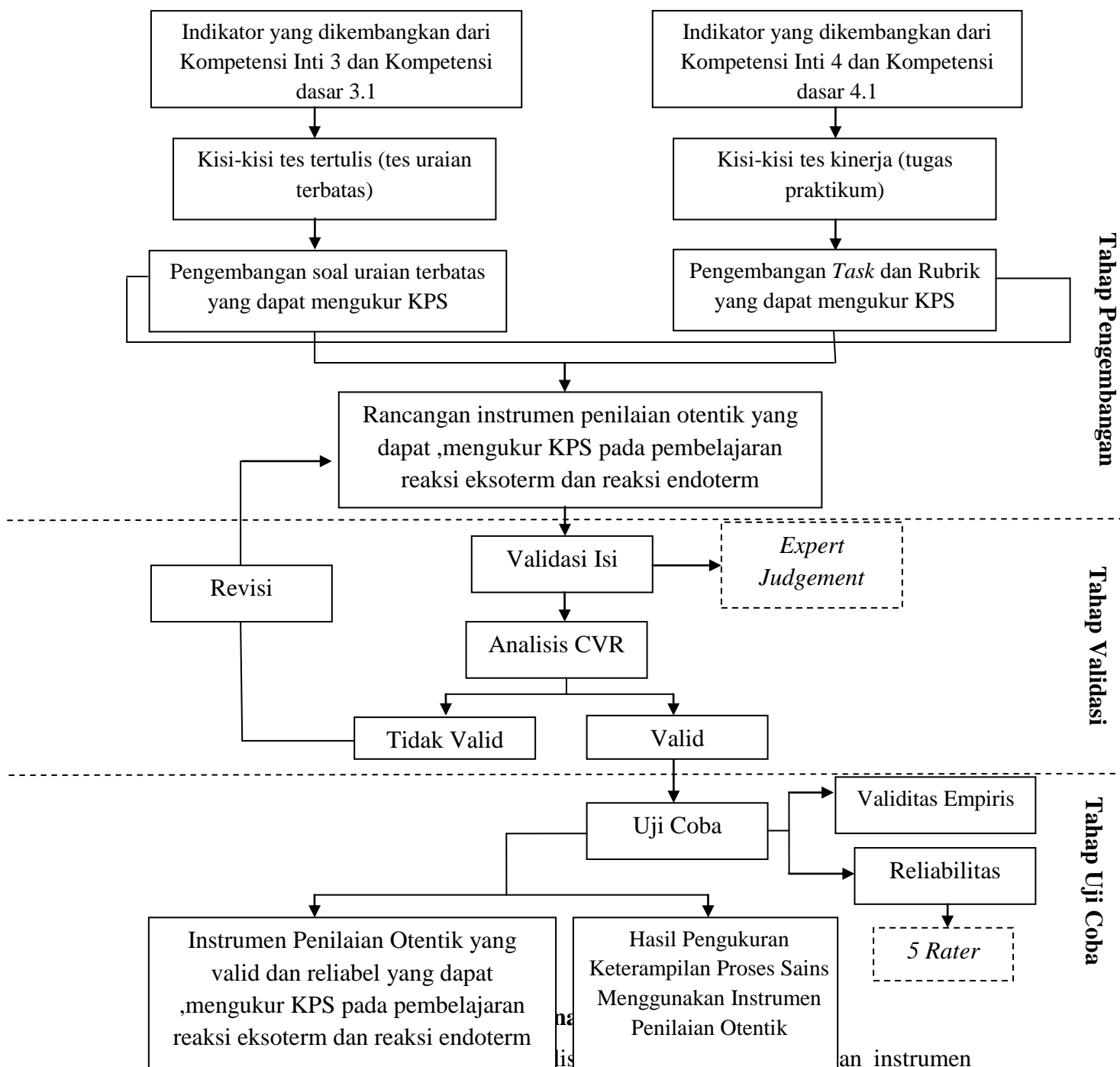
2. Tahap Validasi dan Uji Coba

Tahap validasi dan uji coba instrumen penilaian otentik pada penelitian ini dilakukan dalam beberapa tahap, antara lain:

- 1) Instrumen penilaian otentik (baik instrumen penilaian tertulis maupun instrumen penilaian kinerja) yang telah dikembangkan divalidasi oleh ahli (*expert judgement*) yang terdiri dari lima orang dosen ahli asesmen.
- 2) Hasil validasi isi dianalisis menggunakan analisis CVR kemudian dibandingkan dengan nilai CVR kritis pada lima validator.
- 3) Hasil perbandingan nilai CVR hitung dengan CVR kritis diperoleh informasi ke-valid-an dari butir soal dan *task* yang dikembangkan. Pada tahap ini juga dilakukan revisi instrumen berdasarkan saran yang diberikan oleh validator sehingga didapat instrumen penilaian otentik yang valid.
- 4) Instrumen penilaian otentik yang valid diuji coba ke 38 orang responden untuk memperoleh nilai reliabilitas, validitas empiris, dan informasi keterampilan proses sains yang terukur melalui masing-masing instrumen. Untuk memperoleh data reliabilitas instrumen penilaian kinerja, digunakan lima orang *rater* untuk melakukan penilaian terhadap lima responden yang melaksanakan *task*. Sedangkan data reliabilitas instrumen penilaian tertulis diperoleh melalui pengolahan data hasil uji coba instrumen. Validitas empiris dari instrumen penilaian tertulis dan tes kinerja diperoleh melalui pengolahan data hasil uji coba menggunakan teknik korelasi produk momen.
- 5) Hasil pengukuran keterampilan proses sains diperoleh melalui analisis hasil penilaian yang diperoleh selama pengumpulan data berlangsung.

Tahapan-tahapan yang menjadi acuan dalam penelitian ini diperjelas dalam gambar 3.2.





Gambar 3.2 Diagram alir penelitian pengembangan instrumen penilaian otentik pada pembelajaran kimia meliputi validasi isi yang diperoleh melalui validasi ahli. Sedangkan teknik pengumpulan dan analisis data hasil uji coba instrumen penilaian otentik baik penilaian kinerja maupun uraian terbatas meliputi validitas empiris, reliabilitas dan hasil pengukuran keterampilan proses sains

peserta didik dalam pembelajaran reaksi eksoterm dan reaksi endoterm menggunakan penilaian otentik yang dikembangkan.

1. Validitas Isi

Terdapat banyak jenis validitas, antara lain validitas permukaan (*face validity*), validitas isi (*content validity*), validitas empiris (*empirical validity*), validitas konstruk (*construct validity*) dan validitas faktor (*factorial validity*) (Arifin, 2011, hlm 248). Validitas dilakukan untuk melakukan suatu *judgement* yaitu untuk meminta nasehat dari para ahli dalam membuat soal uji. Validitas permukaan dilakukan untuk melihat sisi muka dari instrumen itu sendiri. Validitas konstruk dilakukan untuk mengetahui apakah instrumen yang dikembangkan dapat mengobservasi dan mengukur apa yang harus diukur. Validitas faktor dilakukan untuk mengetahui homogenitas faktor antar variabel. Validitas isi dilakukan untuk meminta pertimbangan dan saran dari para ahli dalam melihat kesesuaian antara butir soal dengan indikator yang ingin dicapai serta kesesuaian antara *task* dengan indikator yang ingin dicapai. Instrumen penilaian uraian terbatas dan instrumen penilaian kinerja yang dikembangkan dianalisis validitas isi melalui *expert judgement* secara independen pada waktu yang sama. Hasil *expert judgement* yang berisi pertimbangan para ahli di analisis menggunakan analisis *Content Validity Ratio* (CVR) yang dikembangkan oleh Lawshe (1975), dengan rumus:

$$CVR = \frac{(ne - \frac{N}{2})}{\frac{N}{2}}$$

Keterangan: ne : Jumlah ahli yang menyatakan butir soal valid

N : Jumlah anggota tim ahli

Hasil rumusan dari Lawshe (1975):

- 1) Jika validator atau ahli yang menyatakan setuju kurang dari setengah jumlah total validator, maka CVR bernilai negatif
- 2) Jika validator atau ahli yang menyatakan setuju tepat setengah dari jumlah total validator, maka CVR bernilai nol

- 3) Jika validator atau ahli yang menyatakan setuju lebih dari setengah jumlah total validator, maka CVR berada antara 0 sampai dengan 0,99
- 4) Jika semua validator atau ahli yang menyatakan setuju, maka CVR bernilai 1,00

Validator atau ahli yang memvalidasi instrumen penilaian uraian terbatas dan penilaian kinerja yang dikembangkan berjumlah lima orang validator. Nilai CVR kritis untuk lima validator pada tingkat signifikansi 0,05 berdasarkan tabel Schipper adalah 0,736 (Wilson, 2012, hlm.206). Instrumen penilaian, baik instrumen penilaian uraian terbatas maupun instrumen penilaian kinerja, dikatakan valid apabila nilai CVR hitung yang diperoleh lebih tinggi dari pada nilai CVR kritis. Nilai CVR kritis berdasarkan tabel Schipper disajikan dalam Tabel 3.1 berikut.

Tabel 3.1 Nilai CVR Kritis (Wilson, 2012, hlm. 206)

N	Tingkat Signifikansi Tes Satu Sisi					
	.1	.05	.025	.01	.005	.001
	Tingkat Signifikansi Tes Dua Sisi					
	.2	.1	.05	.02	.01	.002
5	.573	.736	.877	.99	.99	.99
6	.523	.672	.800	.950	.99	.99
7	.485	.622	.741	.879	.974	.99
8	.453	.582	.693	.822	.911	.99
9	.427	.548	.653	.775	.859	.99
10	.405	.520	.620	.736	.815	.997

2. Validitas Empiris

Validitas empiris instrumen penilaian yang dikembangkan dilakukan setelah dilaksanakan uji coba atau setelah mendapat pengalaman. Analisis validitas empiris instrumen penilaian, baik instrumen penilaian uraian terbatas

maupun instrumen penilaian kinerja, dilakukan menggunakan teknik analisis korelasi produk momen karena validitas empiris dilakukan dengan tujuan mencari hubungan antara skor tes yang diperoleh dengan kriteria tertentu yang menjadi tolok ukur di luar tes tetapi relevan dengan tes itu sendiri. Pada instrumen penilaian uraian terbatas, validitas empiris dianalisis setiap butir soalnya. Pada instrumen penilaian kinerja, validitas empiris dianalisis secara keseluruhan.

Dalam menguji validitas empiris digunakan jenis statistika korelasi *product-moment* yang dikembangkan oleh Pearson yang menguji perbedaan peringkat atau korelasi diagram pencar. Cara perhitungan korelasi yang dilakukan:

rumus:

$$r_{xy} = \frac{\Sigma xy}{\sqrt{(\Sigma x^2)(\Sigma y^2)}}$$

Keterangan: r = koefisien korelasi

Σxy = jumlah produk xy

Σx^2 = jumlah produk x^2

Σy^2 = jumlah produk y^2

Besarnya koefisien korelasi apabila diperoleh nilai $-1,00$ maka korelasi negatif dan sebaliknya apabila diperoleh nilai $+1,00$ maka nilai korelasi positif.

Untuk menafsirkan koefisien korelasi dapat digunakan kriteria berikut:

0,81 - 1,00 = sangat tinggi

0,61 - 0,81 = tinggi

0,41 - 0,61 = cukup

0,21 - 0,41 = rendah

0,00 - 0,21 = sangat rendah

Apabila nilai koefisien korelasi yang diperoleh bernilai tinggi dapat dikatakan bahwa terdapat hubungan atau korelasi yang kuat antara instrumen yang dikembangkan dengan skor tes sehingga instrumen tersebut mampu mendukung tes untuk mengukur apa yang seharusnya terukur.

3. Reliabilitas Instrumen Penilaian Otentik

Reliabilitas berhubungan dengan tingkat kepercayaan dan keajegan suatu instrumen. Instrumen dikatakan memiliki nilai reliabilitas tinggi apabila instrumen tersebut memberikan hasil yang tetap dan cukup dipercaya untuk digunakan sebagai alat pengumpul data (Arikunto,2013). Perhitungan reliabilitas instrumen penilaian otentik dilakukan dengan cara melakukan analisis tiap butir soal dan *task* berdasarkan indikator kemudian dihitung nilai *cronbach alpha* menggunakan program *IBM SPSS Statistic 20* .

- Reliabilitas instrumen penilaian kinerja setiap *task*

Pada instrumen penilaian kinerja digunakan reliabilitas *inter-rater* yang dihitung berdasarkan tingkat persetujuan setiap *rater*. Metode penilaian ini dilaksanakan satu kali pada beberapa sampel dengan menggunakan lima orang *rater* yang berkerja secara terpisah. Nilai reliabilitas penilaian tes kinerja diperoleh melalui analisis nilai yang diberikan *rater* pada setiap responden untuk *task* tertentu. Selanjutnya nilai dianalisis menggunakan *IBM SPSS Statistic 20* dengan mencari nilai *cronbach alpha* dan mencari matriks *inter-item correlation*. Apabila nilai alpha yang diperoleh bernilai negatif, maka *task* yang bernilai negatif tersebut direduksi. Dan apabila nilai *cronbach's alpha based on Standarized Items* lebih tinggi dari nilai r-tabel maka item/*task* dinyatakan reliabel.

- Reliabilitas instrumen penilaian kinerja keseluruhan

Perhitungan nilai reliabilitas instrumen penilaian kinerja keseluruhan dilakukan dengan cara menghitung rerata nilai yang diberikan *rater* kepada responden untuk semua *task* yang dilakukan responden. Hasil rerata yang diberikan setiap *rater* dihitung nilai *cronbach alpha* dan matriks *inter-item correlation* menggunakan *IBM SPSS Statistik 20*. Nilai maksimum *cronbach alpha* merupakan nilai reliabilitasnya. Dan apabila nilai maksimum *cronbach alpha* lebih tinggi dari nilai r-tabel maka instrumen dinyatakan reliabel.

- Reliabilitas instrumen penilaian uraian terbatas

Perhitungan nilai reliabilitas instrumen penilaian uraian terbatas dilakukan dengan cara menghitung skor total yang diperoleh semua responden yang kemudian skor total tersebut dihitung nilai *cronbach alpha* dan matriks *inter-item correlation* menggunakan IBM SPSS Statistik 20. Nilai maksimum *cronbach alpha* merupakan nilai reliabilitasnya. Dan apabila nilai maksimum *cronbach alpha* lebih tinggi dari nilai r-tabel maka instrumen dinyatakan reliabel.

Nilai reliabilitas hasil perhitungan yang diperoleh kemudian dikategorisasi menggunakan kriteria penafsir koefisien reliabilitas (Arikunto,2007):

0,80 – 1,00	: sangat tinggi
0,60 – 0,80	: tinggi
0,40 – 0,60	: cukup
0,20 – 0,40	: rendah
< 0,20	: sangat rendah

4. Hasil pengukuran keterampilan proses sains

Keterampilan proses sains dianalisis setiap aspeknya yang terukur melalui instrumen penilaian, baik instrumen penilaian uraian terbatas maupun instrumen penilaian kinerja, yang dikembangkan. Aspek keterampilan proses yang terukur melalui butir soal dan/ atau *task* kinerja tertentu diukur dengan cara menghitung skor mentah yang diperoleh masing-masing butir soal dan/ atau *task* kinerja pada satu aspek kemudian skor mentah tersebut dijumlahkan dan dikonfersi ke dalam skor akhir.

$$\text{Rumus : Skor Akhir} = \frac{\Sigma \text{skor yang diperoleh}}{\text{Skor total}} \times 100$$

Skor akhir yang diperoleh selanjutnya dikategorisasi menggunakan kriteria tingkatan berikut (Firman, 2013):

75 – 100	: sangat baik
50 – 74	: baik
25 – 49	: cukup
0 - 24	: kurang

Hasil pengelompokan (kategorisasi) tersebut menunjukkan keterampilan proses sains peserta didik yang terukur dalam satu aspek.