

PENGEMBANGAN TES KOGNITIF PADA MATERI SISTEM PERNAPASAN

Wiwin Pramita Arif

Program Studi Pendidikan Biologi Fakultas Keguruan dan Ilmu Pendidikan, Universitas
Muslim Maros, Kampus 1 Jalan Dr Ratulangi No. 62 Maros Sulawesi Selatan Kode Pos 90511,
e-mail: wiwinpramitaarif@gmail.com

Abstrak

Penelitian ini adalah penelitian pengembangan. Tujuan penelitian ini adalah untuk mengembangkan tes kognitif yang valid dan reliabel, dengan tes yang memiliki tingkat kesukaran, daya pembeda, dan efektivitas pengecoh pada materi sistem pernapasan pada siswa kelas XI IPA 4 dan XI IPA 5 SMAN 1 Makassar. Proses pengembangan perangkat tersebut menggunakan model pengembangan tipe *formatif research* oleh Tessmer. Data validitas isi dianalisis secara matematis menggunakan rumus Gregory, sedangkan validitas empiris menggunakan korelasi *point biserial*. Reliabilitasnya dihitung dengan menggunakan program *ITEMAN*, sedangkan tingkat kesukaran, daya beda, dan pengecoh dianalisis dengan menggunakan bantuan analisis *ANATES Ver. 4*. Hasil ujicoba menghasilkan tes kognitif dengan soal tes bentuk pilihan ganda sebanyak 50 nomor, 25 nomor untuk paket A dan 25 nomor untuk paket B. Keseluruhan paket soal dinyatakan valid dan reliabel dengan tingkat kesukaran untuk masing-masing paket dari kategori sulit, sedang, dan mudah yaitu 1:2:1. Daya beda masing-masing paket berada pada kategori baik. Efektivitas pengecoh masing-masing paket berfungsi dengan baik.

Kata kunci: Analisis Butir Soal, Sistem Pernapasan, Tes Kognitif.

Abstract

The study was a developmental research, which aimed to develop cognitive tests that was valid and reliable which has difficulty level, determinant index, and distractors effectiveness on respiratory system course material. The subject of the study was XI IPA 4 and XI IPA 5 students of SMAN 1 Makassar. The process of instrument development used formative research type by Tessmer's. The content validity data analyzed with Gregory's formula and empirical validity data used point biserial correlation. Item reliability analyzed with ITEMAN, while item difficulty level, determinant index, and distractors effectiveness analyzed with ANATES Ver.4. The results of the study obtained cognitive tests of 50 numbers multiple of choice questions which consist of 25 numbers of A package and 25 numbers of B package. All the test packages declared valid and reliable with categories of difficulty level fell into difficult, moderate, and easy 1: 2: 1. The determinant index of each package was in good categories and distractor effectiveness all packages were working properly.

Keywords: Analysis of items, Cognitive test, Respiratory system

PENDAHULUAN

Pendidikan merupakan lembaga dan upaya pembangunan bangsa dan watak bangsa mengarah menjadi dewasa, mandiri dan kepribadian yang matang yang meliputi cipta, rasa, maupun karsa (Yahya, 2003). Undang-Undang Nomor 20 tahun 2003 tentang Sistem Pendidikan Nasional, Pasal 3, menyebutkan tujuan pendidikan nasional adalah “ mengembangkan potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan menjadi warga negara yang demokratis serta bertanggung jawab”

Pencapaian tujuan pendidikan senantiasa diupayakan secara menyeluruh, salah satunya melalui penyempurnaan kurikulum. Kurikulum senantiasa berubah dari waktu ke waktu. Perubahan kurikulum ditandai dengan perubahan pada pola proses belajar mengajar, pemanfaatan media pembelajaran, serta penentuan teknik penilaian. Keseluruhan dari perubahan-perubahan tersebut tidak lepas dari peran guru.

Guru sebagai bagian dari proses pembelajaran memiliki banyak peran. salah satunya adalah melakukan penilaian. Dalam hal ini memiliki otoritas menilai keberhasilan pembelajaran (Halking, 2007). Oleh sebab itu, dibutuhkan sebuah acuan penilaian yang baik dan terencana untuk mengukur proses maupun hasil belajar peserta didik (Ali, 2014).

Guru diharuskan mengetahui kemajuan masing-masing peserta didiknya. Kemajuan tersebut dapat diketahui dengan melakukan pengukuran dan penilaian hasil belajar. Pengukuran dapat dilakukan saat kegiatan pembelajaran maupun pada saat tes akhir. Oleh karena itu, guru perlu menguasai teknik dalam mengevaluasi pembelajaran. Guru juga harus mampu menelaah dan meneliti hasil belajar masing-masing peserta didiknya agar mampu menentukan program perbaikan pada pembelajaran selanjutnya (Nurhayati, 2011).

Guru dalam mempersiapkan instrumen penilaian hasil belajar harus memperhatikan karakteristik atau ciri sebuah tes pengukuran hasil belajar. Adapun karakteristik atau ciri tes hasil belajar yang baik, yaitu: *Ciri pertama* adalah bahwa tes hasil belajar harus valid atau memiliki validitas. Kata “valid” sering diartikan dengan: tepat, benar, shahih, absah; tes hasil belajar mampu mengukur yang seharusnya diukur. *Kedua*; tes hasil belajar harus reliabel. Reliabilitas berhubungan dengan masalah kepercayaan atau tetap. Maksudnya adalah jika tes tersebut diujikan berulang-ulang akan memberikan hasil yang sama. Suatu tes dapat dikatakan mempunyai taraf kepercayaan yang tinggi jika tes tersebut dapat memberikan hasil yang tetap. (Arikunto, 2010). *Ketiga*; tes hasil belajar harus memiliki daya beda dimana tes tersebut mampu membedakan antara peserta didik prestasi tinggi dan yang memiliki prestasi rendah. Dengan kata lain jika soal diberikan pada kelompok siswa yang prestasi tinggi, hasilnya baik. Jika diberikan kepada kelompok siswa yang prestasi rendah, hasilnya jelek. Soal tidak memiliki daya pembeda, jika soal tersebut memberikan hasil yang sama jika diberikan kepada kelompok prestasi rendah maupun kelompok dengan prestasi tinggi (Uno, 2012). *Keempat*; Soal harus memiliki tingkat kesukaran dengan pembagian yang seimbang antara soal sukar, sedang, dan mudah. Disesuaikan dengan tujuan tes tersebut

dibuat. *Lima*; tes hasil belajar harus memiliki pengecoh (khusus tes dalam bentuk pilihan ganda) yang berfungsi dengan baik. Pengecoh diketahui dengan mempelajari pilihan jawaban peserta didik.

Terdapat beberapa permasalahan dalam evaluasi pembelajaran di sekolah yang harus diatasi dengan serius. Salah satunya yaitu instrumen yang dihasilkan langsung digunakan sebelum dilakukan uji empirik untuk mengetahui mutunya. Adapun hal lain yang kurang diperhatikan seperti; (1) soal uraian tidak dibuatkan pedoman penskoran/rubrik; (2) materi yang diujikan tidak sesuai indikator penvapaian (tidak esensial); (3) soal objektif tidak dibatkan pengecoh. Sehingga mutu instrumen/soal yang dibuat seadanya saja (Safari, 2013)

Banyak peneliti telah melakukan observasi dan analisis butir soal seperti yang dilakukan oleh Anwar (2006) di beberapa SMA di Kota Padang menemukan bahwa: (1) secara konten, masih ada soal yang dibuat tidak sesuai dengan tujuan pembelajaran; (2) secara empiris, dari 35 butir soal yang diujikan, hanya 17,14% butir soal yang valid, soal pada taraf kesukaran yang diharapkan ($P = 0,33 - 0,67$) sebesar 25,71% butir soal, 31,42% butir soal memiliki indek daya pembeda dalam taraf penerimaan ($D = \geq 0,30$).

Hasil penelitian yang dilakukan Widodo (2010) lebih mengkhawatirkan. Pada tahun pelajaran 2008/2009 diketahui hampir semua guru SD, SMP, SMA BPK PENABUR Tasikmalaya tidak melakukan analisis butir soal-soal tes formatif. Hanya sebagian kecil guru saja yang melakukan analisis butir soal-soal sumatif atau tes evaluasi akhir semester (2 orang guru SD, 1 orang guru SMP, dan 1 Orang guru SMA). Padahal Analisis butir soal merupakan hal yang sangat penting dan tidak dapat dipisahkan dari tahapan penilaian (evaluasi). Instrumen yang baik akan memberikan hasil pengukuran yang baik pula.

Berdasarkan masalah yang telah diuraikan terkait kegiatan yang sangat penting dalam proses evaluasi adalah dengan menyiapkan instrumen atau soal yang terstandar. Sudah menjadi kewajiban guru untuk membuat dan mengembangkan instrumen yang mampu mengukur keadaan siswa apa adanya. Oleh karena itu, peneliti bermaksud membuat tes yang terstandar dengan mengembangkan tes kognitif pada materi sistem pernapasan.

Berdasarkan hal tersebut di atas, maka rumusan masalah dalam penelitian ini adalah bagaimana mengembangkan tes kognitif pada materi sistem pernapasan yang valid dan reliabel dengan karakteristik butir tes yang meliputi tingkat kesukaran sedang, daya beda baik, dan kualitas pengecoh (*distractor*) yang berfungsi.

METODOLOGI PENELITIAN

Penelitian ini adalah penelitian pengembangan (*research and development*) dengan menggunakan metode riset pengembangan tipe *formative research* (Tessmer, 1993., dalam Rahayu, T., dkk, 2008) untuk menghasilkan tes kognitif pada materi sistem pernapasan yang valid, reliabel, dan memiliki karakteristik butir tes yang meliputi, daya beda baik, dan kualitas pengecoh (*distractor*) baik serta tingkat

kesukaran sedang pada materi sistem pernapasan. Penelitian ini menghasilkan tes yang telah diuji coba di SMAN 1 Makassar.

Langkah-langkah penelitian sebagai berikut sebagai berikut:

Self Evaluation

a. Analisis

Tahap awal dalam penelitian pengembangan adalah dengan melakukan analisis pendahuluan meliputi analisis kurikulum berdasarkan standar isi kurikulum 2013 dan analisis siswa.

b. Desain

Setelah melakukan analisis kurikulum, tahap selanjutnya adalah dengan mendesain instrumen yang akan dikembangkan. Desain instrument disesuaikan dengan indikator dan tujuan pembelajaran yang telah dikembangkan dari kompetensi dasar. Pada tahap ini dibuat desain tes kognitif dengan menentukan format/tipe tes yaitu soal yang berbasis revisi taksonomi Bloom (Taksonomi Anderson), dengan bentuk tes pilihan ganda. Setelah menentukan tipe soal yang akan dibuat maka selanjutnya menyusun kisi-kisi tes berdasarkan indikator dan tujuan pembelajaran. Tahap terakhir adalah menyusun soal dari kisi-kisi yang terdiri 50 butir soal yang terbagi menjadi 2 paket yaitu 25 butir paket A dan 25 butir paket B. Hasil pendesain ini disebut sebagai prototipe pertama.

Prototyping

Hasil pendesainan pada prototipe pertama yang dikembangkan kemudian diberikan pada 2 orang pakar (*expert review*) dan satu peserta didik (*one-to-one*) secara paralel. Telaah dan saran oleh pakar serta masukan dari peserta didik dijadikan bahan revisi.

a. *Expert Review*

Pada tahap *expert review*, produk yang telah dibuat (prototipe pertama) diberikan kepada dua orang pakar yang terdiri dari pakar konten dan pakar evaluasi. Kedua pakar kemudian menilai ketepatan konten, bahasa dan konstruk dari masing-masing butir soal (prototype pertama). Semua tanggapan dan saran dari para pakar dituliskan pada lembar validasi sebagai bahan untuk merevisi tes yang dikembangkan.

b. *One-to-one*

Pada tahap *one-to-one*, soal yang telah dibuat diujicobakan kepada satu orang peserta didik yang menjadi tester. Selain menjawab soal, tester juga diminta menanggapi soal-soal yang dianggap ambigu atau yang tidak dipahami. Mencatat waktu yang diperlukan, serta mengamati kesulitan-kesulitan yang dialami peserta didik selama mengerjakan soal tersebut. Hasil dari pelaksanaan ini digunakan untuk merevisi prototipe yang telah dibuat.

c. *Small Group*

Hasil revisi prototipe kedua berdasarkan telaah para pakar (*expert*) dan uji coba pada *one-to-one* kemudian hasilnya diujicobakan pada *small group* (20 orang peserta didik). Hasil Analisis *small group* kemudian direvisi kembali berdasarkan

saran/komentar peserta didik pada *small group*. Hasil analisis butir soal ini yang selanjutnya dinamakan prototipe ketiga.

Field Test

Hasil analisis dari prototipe ketiga kemudian direvisi kembali sebelum diujicobakan pada subjek penelitian, uji coba terakhir ini merupakan uji lapangan (*field test*). Produk yang telah diujicobakan pada uji lapangan merupakan produk yang telah memenuhi kriteria kualitas.

Data validitas isi tes kognitif diperoleh melalui hasil pandangan dua orang pakar. Sedangkan validitas empirik dianalisis dengan korelasi *Point Biserial*. Reliabilitasnya dihitung dengan menggunakan program *ITEMAN*, sedangkan tingkat kesukaran, daya beda, dan pengecoh dianalisis dengan menggunakan bantuan analisis *ANATES Ver. 4*.

Analisis Data

Analisis validitas isi dilakukan dengan memperhatikan relevansi oleh dua orang pakar. Hasil analisis tersebut dijadikan sebagai pedoman untuk merevisi soal yang masih mendapat penilaian kurang. Relevansi kedua pakar dihitung berdasarkan rumus Gregory sebagai berikut (Ruslan, 2009).

$$\text{Validasi isi: } \frac{D}{(A+B+C+D)}$$

Keterangan :

A = jumlah butir soal dari validator 1 dengan relevansi lemah (butir bernilai 1 atau 2) terhadap relevansi lemah (butir bernilai 1 atau 2) dari validator kedua

B = jumlah butir soal dari validator 1 dengan relevansi lemah (butir bernilai 1 atau 2) terhadap relevansi kuat (butir bernilai 3 atau 4) dari validator kedua

C = jumlah butir soal dari validator 1 dengan relevansi kuat (butir bernilai 3 atau 4) terhadap relevansi lemah (butir bernilai 1 atau 2) dari validator kedua

D = jumlah butir soal dari validator 1 dengan relevansi kuat (butir bernilai 3 atau 4) terhadap relevansi kuat (butir bernilai 3 atau 4) dari validator kedua

Berikut adalah model kesepakatan antar penilai untuk validasi isi:

Tabel 1. Model Kesepakatan Antar Dua Pakar

		Validator I	
		Tidak relevan Skor (1-2)	Relevan Skor (3-4)
Validator II	Tidak relevan Skor (1-2)	A	B
	Relevan Skor (3-4)	C	D

Sumber: (Gregory dalam Ruslan, 2005)

Derajat validitas soal ditentukan dengan memperhatikan relevansi penilaian kedua pakar. Butir soal dikatakan valid jika koefisien validasi >75% (Ruslan, 2009). Jika tidak memenuhi kriteria maka perlu dilakukan revisi berdasarkan saran dari

validator satu dengan melihat kembali aspek-aspek yang nilainya kurang, selanjutnya dilakukan validasi ulang kemudian dianalisis kembali. Demikian seterusnya hingga diperoleh data dalam kategori valid.

Validitas empirik dianalisis dari hasil uji coba small group dengan menggunakan rumus korelasi *point biserial* mempunyai pola rumus :

$$r_{pbis} = \frac{M_p - M_t}{Sd_t} \sqrt{\frac{p}{q}}$$

Keterangan :

- r_{pbis} = Koefisien korelasi point biserial
- M_p = Skor rata-rata hitung untuk butir yang dijawab betul
- M_t = Skor rata-rata dari skor total
- Sd_t = Standar deviasi skor total
- p = Proporsi siswa yang menjawab betul pada butir yang diuji validitasnya
- q = Proporsi siswa yang menjawab salah pada butir yang diuji validitasnya

Soal-soal yang dinyatakan valid kemudian dianalisis derajat reliabilitasnya. dengan menggunakan program *ITEMAN*. Adapun nilai koefisien reliabilitas tes diinterpretasi dengan menggunakan patokan menurut Sukiman (2012) dalam pemberian interpretasi terhadap koefisien reliabilitas instrumen (r) pada umumnya digunakan patokan sebagai berikut:

- a. Jika $r \geq 0,70$ berarti reliabel
- b. Jika $r < 0,70$ berarti tidak reliabel

Uji tingkat kesukaran soal, daya pembeda, dan pengecoh/*distractor* dianalisis menggunakan program *ANATES* Ver. 4. Indeks tingkat kesukaran soal (p) yaitu: $p < 0,3$ (sukar), $0,3 \leq p \leq 0,70$ (sedang) dan $p > 0,7$ (mudah) Purwanto, 2001).

Distractor sudah berfungsi dengan baik jika sudah dipilih oleh lebih dari 5% pengikut tes ($p > 5\%$) jika empat pilihan jawaban dan 3% untuk lima pilihan jawaban (Depdiknas, 2004).

Rentang indeks daya beda berkisar 0,00 sampai 1,00. Semakin tinggi nilai indeks daya beda semakin baik. Crocker & Algina (1986) menyatakan, kriteria besarnya koefisien daya beda yaitu: 0,00 sampai 0,19 buruk; 0,20 sampai 0,29 perlu revisi 0,30 sampai 0,39 dapat diterima (tidak perlu direvisi); 0,40 sampai 0,1 baik.

HASIL PENELITIAN

Tes kognitif pada materi sistem pernapasan yang diujikan pada Kelas XI SMA Negeri 1 makassar dengan model riset pengembangan tipe *formatif research* (Tessmer, 1993., dalam Rahayu, T., dkk, 2008), hasil penelitian yang telah dilakukan sebagai berikut:

Self Evaluation

Analisis

Langkah awal penelitian pengembangan dilakukan dengan melakukan analisis pendahuluan meliputi:

1) Analisis Siswa

Peserta didik yang menjadi subyek penelitian ini adalah peserta didik kelas XI IPA 4 dan kelas XI IPA 5 SMA Negeri 1 Makassar sebanyak 60 orang. Pada analisis siswa, peneliti menelaah tentang latar belakang pengetahuan siswa, bahasa yang digunakan dan perkembangan kognitif siswa sesuai dengan pemahaman guru. Berdasarkan telaah tersebut, penulis menemukan bahwa peserta didik kelas XI IPA 4 dan XI IPA 5 SMA Negeri 1 Makassar memiliki kemampuan akademik yang beragam, yang terdiri atas siswa yang berkemampuan tinggi, sedang, dan rendah. Dilihat dari latar belakang pengetahuannya, peserta didik XI IPA 4 dan XI IPA 5 SMA Negeri 1 Makassar telah mempelajari konsep dasar sistem pernapasan manusia di Sekolah Menengah Pertama (SMP). Bahasa yang digunakan siswa baik dalam kehidupan sehari-hari maupun dalam proses pembelajaran adalah Bahasa Indonesia. Perkembangan kognitif siswa di kelas XI IPA 4 dan XI IPA 5 SMA Negeri 1 Makassar sangat beragam yaitu kemampuan kognitif tinggi, sedang dan rendah, berdasarkan nilai yang diperoleh pada materi-materi sebelumnya dan sesuai dengan pemahaman guru biologi kelas XI IPA 4 dan XI IPA 5 SMA Negeri 1 Makassar. Oleh karena itu sangat tepat bila kelas XI IPA 4 dan XI IPA 5 SMA Negeri 1 Makassar dijadikan sebagai subjek penelitian pengembangan tes kognitif yang membutuhkan peserta didik yang beragam..

2) Analisis kurikulum

Setelah melakukan analisis siswa selanjutnya dilakukan analisis kurikulum berdasarkan Standar Isi Kurikulum 2013 yang diterapkan di SMA Negeri 1 Makassar. Kurikulum yang diterapkan adalah kurikulum 2013. Kompetensi Dasarnya adalah Menganalisis hubungan antara struktur penyusun organ pada sistem respirasi dalam kaitannya dengan bioproses dan gangguan fungsi yang dapat terjadi pada sistem respirasi manusia. Indikatornya meliputi Mengidentifikasi struktur dan fungsi sistem pernapasan manusia: (1) menjelaskan proses pernapasan yang terjadi pada manusia, (2) membandingkan volume dan kapasitas paru-paru, (3) menjelaskan proses pertukaran gas, (4) mengumpulkan informasi dari berbagai sumber tentang gangguan/penyakit yang terdapat dalam sistem pernapasan manusia.

Desain

Desain pengembangan tes kognitif yang dilakukan meliputi pembuatan kisi-kisi (blue print), penggunaan tabel taksonomi, dan menentukan bentuk instrumen penilaian. Hasil dari tahap ini menghasilkan tes kognitif berupa soal tes dengan kategori soal tes pilihan ganda sebanyak 50 butir yang terdiri dari dua paket yaitu paket A 25 butir dan Paket B 25 butir. Hasil pengembangan soal tes ini kemudian dijadikan sebagai prototipe awal (pertama) yang selanjutnya akan divalidasi oleh pakar (validator).

Prototyping

Soal yang dihasilkan pada prototipe pertama yang dikembangkan setelah analisis siswa dan analisis kurikulum pada tahap *self evaluation* kemudian diberikan pada pakar (*expert review*) dan seorang siswa (*one-to-one*) secara paralel.

Expert Review

Produk awal yang telah dikembangkan diberikan kepada dua orang pakar/validator untuk mengukur tingkat validitas. Validasi relevansi antara indikator pembelajaran dengan soal tes serta dimensi pengetahuan dan dimensi proses kognisi pada tabel klasifikasi butir soal dilakukan dengan memberikan kisi-kisi butir tes dan tes yang akan diujikan. Hasil validasi yang diberikan oleh kedua validator yang dihitung tingkat validitasnya berdasarkan rumus validitas isi Gregory menurut Lawshe dan Martuza (dalam Ruslan, 2009) diperoleh nilai sebesar 1,0 Hasil analisis penilaian dua pakar dapat dijelaskan sebagai berikut.

Tabel 2. Rekap Hasil Relevansi Validasi Ahli

Dari penilaian yang diberikan oleh kedua pakar di atas dapat dihitung nilai validitas isi sebagai berikut:

		Validator I	
		Tidak relevan Skor (1-2)	Relevan Skor (3-4)
Validator II	Tidak relevan Skor (1-2)	1	0
	Relevan Skor (3-4)	0	16

$$\text{Validitas Isi} = \frac{16}{1+0+0+16} = \frac{16}{17} = 0,94$$

Berdasarkan nilai validitas/kesahihan yang diperoleh yaitu 0,94 atau $V = 94\%$. Hal ini berarti bahwa hasil penilaian dari kedua validator memiliki “relevansi kuat” dengan koefisien validitas isi lebih dari 0,75 atau $V > 75\%$, maka dapat dikatakan bahwa relevansi antara indikator, jenis soal, dimensi pengetahuan dan dimensi proses kognisis pada tabel klasifikasi butir soal tes kognitif yang dibuat adalah valid. Hal ini berarti perangkat penilaian kinerja tersebut telah layak diuji coba pada *small group*.

a. *One-to-one*

Pada tahap *one-to-one*, soal diujikan kepada seorang peserta didik yang bernama Andi Ray Rangga Y, kelas XI IPA 5 SMA Negeri 1 Makassar diminta untuk mensimulasikan pengerjaan tes kognitif yang akan diujicobakan.

Prosedur yang pertama dalam *one-to-one* adalah dengan menjelaskan kepada peserta didik tentang bahan pembelajaran yang akan digunakan. Peneliti memberikan soal-soal kognitif kepada peserta didik yang telah terpilih. Setelah peserta didik memahami bahan pengajaran, meninjau, mengkaji kembali inti pengajaran dan menjawab tes kognitif. Peneliti menanyakan pendapat peserta didik terhadap materi sistem pernapasan, menanyakan kekurangan materi dan mencatat waktu yang

diperlukan untuk menyelesaikan materi yang dibutuhkan oleh peserta didik tersebut dan mengoreksi kesalahan-kesalahan yang nyata.

Peserta didik menemukan beberapa kesalahan-kesalahan yaitu: berupa kesalahan penulisan rumus kimia H₂O seharusnya H₂O, C₆H₁₂O₆ seharusnya C₆H₁₂O₆, angka pada penunjukkan kurang jelas dan ada beberapa soal yang tidak didahului oleh pengantar soal sehingga soal menjadi sulit dipahami.

Berdasarkan hasil ujicoba desain *one-to-one* peneliti dapat menggunakan semua informasi yang diperoleh untuk melakukan beberapa revisi. Berdasarkan *one-to-one* dan *expert reviews* yang diberikan secara paralel maka prototipe pertama akan direvisi.

b. *Small Group*

Hasil revisi dan saran dari telaah *expert* dan kesulitan yang dialami saat uji coba pada prototipe pertama dijadikan dasar untuk merevisi prototipe tersebut hasil revisi kemudian diujicobakan pada *small group* yang terdiri dari 20 orang peserta didik kelas XI IPA 5 SMA Negeri 1 Makassar yang diminta untuk mengerjakan soal-soal untuk mensimulasikan waktu pengerjaan sesuai banyak jumlah butir soal.. Peneliti terlibat langsung untuk melihat kesulitan-kesulitan yang mungkin terjadi selama proses pengerjaan soal sehingga dapat memberikan gambaran yang jelas apakah tes kognitif tersebut perlu direvisi atau tidak.

Dari hasil *small group* yang dilakukan dapat dilihat validitas, reliabilitas dan analisis butir soal dari perangkat yang dibuat.

1) Analisis validasi *Koefisien korelasi point biserial*

Pada tahap ini instrumen tes kognitif prototipe pertama yang telah direvisi dan menjadi prototipe kedua kemudian diujikan untuk mengukur validitas dan reliabilitas soal yang dilakukan pada siswa kelas XI IPA 5 SMA Negeri 1 Makassar yang berjumlah 20 orang yang telah mempelajari materi sistem pernapasan.

Hasil pengujian pada *small group* untuk validitas dengan menggunakan analisis korelasi Pearson pada software *ITEMAN*, hasil perhitungan dapat dilihat pada tabel di bawah ini.

Tabel 3 Hasil Validasi Isi Butir Soal paket A dan paket B dengan Analisis Koefisien korelasi point biserial

Kategori	Paket A		Paket B	
	Nomor Soal	Jumlah	Nomor Soal	Jumlah
Valid	1, 2, 3, 4, 6, 7,	23 (92%)	1, 2, 3,4, 6, 7,	23 (92%)
	8, 9, 10, 11,		8, 9, 10, 11,	
	13, 14, 15, 16,		13, 14, 15, 16,	
	17, 18, 19, 20,		17, 18, 19, 20,	
	21, 22, 23, 24,		21, 22, 23, 24,	
	25		25	
Tidak Valid	5, 12	2 (8%)	7, 11	(8%)

Sumber: Data Peneliti

Berdasarkan tabel dengan mengetahui nilai *Koefisien korelasi point biserial* maka soal pilihan ganda yang dinyatakan tidak valid ada 2 butir untuk paket A maka soal yang valid setara 92% dan untuk paket B soal tidak valid ada 2 butir, maka soal yang valid setara 92%.

2) Analisis reliabilitas

Hasil pengujian untuk reliabilitas dari soal yang telah dinyatakan valid dengan menggunakan analisis reliabilitas koefisien *Alpha Cronbach* pada *software ITEMAN*. Hasil analisis reliabilitas menunjukkan bahwa nilai reliabilitas yang diperoleh untuk soal pilihan ganda paket A sebesar 0,879 dan paket B sebesar 0,872. Berdasarkan hasil perhitungan reliabilitas koefisien *Alpha Cronbach* maka dapat diinterpretasikan jika nilai reliabilitas (r) yang diperoleh $\geq 0,70$ tes hasil belajar yang sedang diuji reliabilitasnya memiliki tingkat keandalan yang tinggi, dengan demikian dapat dikatakan soal bentuk pilihan ganda baik untuk soal paket A maupun paket B adalah reliabel.

Field Test

Tes kognitif prototipe kedua yang telah diuji validitas dan reliabilitasnya pada tahap *small group* menjadi prototipe ketiga kemudian diujikan kepada *field test*. Pada tahap ini tes kognitif prototipe ketiga diujikan pada siswa kelas XI IPA 4 SMA Negeri 1 Makassar yang berjumlah 40 orang. Setelah tes kognitif diujikan kemudian dianalisis tingkat kesukaran, daya beda, dan pengecoh (*distractor*) tiap butir soal.

a. Hasil Analisis Tingkat Kesukaran

Pengkategorian tingkat kesukaran sesuai dengan ketentuan yang telah ditetapkan yaitu apabila ITK (Indeks Tingkat Kesukaran) 0,00-0,30 soal tergolong sukar, ITK 0,31-0,70 soal tergolong sedang, ITK 0,71-1,00 soal tergolong mudah. Adapun hasil perhitungan tingkat kesukaran soal dapat dilihat pada tabel 4.

Tabel 4. Rekap hasil Analisis Tingkat Kesukaran Soal Paket A dan Paket B

Kategori	Paket A		Paket B	
	Nomor Soal	Jumlah	Nomor Soal	Jumlah
Mudah	2, 9, 12, 15, 18, 19	6 (26%)	1, 5, 16, 17	4 (17%)
Sedang	1, 3, 4, 5, 13, 14, 16, 17, 20, 21, 22, 23	12 (52%)	3, 4, 6, 7, 8, 9, 10, 11, 12, 15, 19, 21, 22, 23	14(61%)
Sukar	6, 7, 8, 10, 11	5 (22%)	2, 13, 14, 18, 20	5 (22%)

Sumber: Data Peneliti

Berdasarkan Tabel 4 diketahui bahwa kedua paket soal yang telah dibuat memiliki proporsi kategori mudah untuk paket A 6% dan untuk paket B 4%, kategori sedang

untuk paket A 52% dan untuk paket B 14%, sedangkan kategori sukar untuk paket A 5% dan paket B 5%.

b. Hasil Analisis Daya Pembeda

Tes kognitif yang telah dibuat memiliki koefisien daya beda yaitu: 0,00 sampai 0,19 buruk; 0,20 sampai 0,29 perlu revisi 0,30 sampai 0,39 dapat diterima (tidak perlu direvisi); 0,40 sampai 0,1 baik. Adapun hasil perhitungan daya pembeda dapat dilihat pada tabel 5.

Tabel 5. Rekap hasil Analisis Daya Pembeda Soal Paket A dan Paket B

Kategori	Paket A		Paket B	
	Nomor Soal	Jumlah	Nomor Soal	Jumlah
Baik	Semua Nomor Soal (1-23)	23 (100%)	1-21, 23	22 (96%)
Perlu Revisi	-	0	22	1(14%)

Sumber: Data Peneliti

Berdasarkan Tabel 5 diketahui bahwa soal paket A telah memiliki daya beda yang baik dan tidak ada butir soal yang memerlukan revisi, sedangkan soal paket B terdapat 1 (14%) soal yang harus direvisi.

c. Hasil Analisis Efektivitas Pengecoh

Analisis pengecoh dilakukan dengan cara menghitung persentase jumlah peserta didik yang memilih setiap jawaban pada setiap soal. Hasil perhitungan pengecoh dapat dilihat pada tabel 6.

Tabel 6. Rekap hasil Analisis Pengecoh/Distractor Paket A dan Paket B

Kategori	Paket A		Paket B	
	Nomor Soal	Jumlah	Nomor Soal	Jumlah
Efektif	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 22	20 (87%)	3, 4, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 22	22 (96%)
Tidak Efektif	3, 12, 21	3 (13%)	22	1(4%)

Sumber: Data Peneliti

Berdasarkan Tabel 6 diketahui bahwa 3 (13%) soal paket A harus direvisi karena memiliki pengecoh yang tidak berfungsi, sedangkan soal paket B terdapat 1 (4%) soal yang harus direvisi.

PEMBAHASAN

Validitas

Hasil penilaian dari validator terhadap tes kognitif prototipe pertama yang dibuat dinyatakan memenuhi kriteria valid berdasarkan hasil analisis validitas isi sebesar 0,94. Nilai tersebut menunjukkan bahwa relevansi antara indikator, jenis soal, dimensi pengetahuan dan dimensi proses kognisis pada tabel klasifikasi butir soal berada dalam kategori valid karena koefisien validitas isi yang dihasilkan $> 0,75$. Hal ini ditunjukkan oleh Gregory dalam Ruslan (2009) bahwa jika hasil dari koefisien validitas isi tinggi ($V > 0,75$), maka dapat dinyatakan bahwa hasil pengukuran yang dilakukan adalah valid meskipun sebelumnya telah dilakukan beberapa revisi sesuai dengan saran yang diberikan oleh validator.

Hasil ujicoba lapangan pada *small group* untuk mengetahui tingkat validitas isi dari tiap butir soal dilakukan dengan menganalisis hasil pemberian tes kognitif dengan menggunakan analisis korelasi *koefisien point biserial*. Hasil analisis validitas isi tiap butir soal diperoleh tes kognitif yang dinyatakan valid sebanyak 46 atau 92% butir soal dari 50 butir soal yang terdiri dari paket A yang tidak valid 2 atau 92% butir soal yang valid dari 25 butir soal dan paket B yang tidak valid 2 atau 92% butir soal yang valid dari 25 butir soal. Soal yang tidak valid kemudian peneliti tidak digunakan lagi pada ujicoba selanjutnya karena soal yang valid sudah memenuhi indikator pencapaian kompetensi pada pembelajaran.

Berdasarkan hasil penelitian maka dapat dikatakan validitas adalah salah satu syarat untuk mendapatkan soal yang baik. Senada dengan dalam hasil penelitian dari Anwar (2006) soal yang bermutu adalah soal yang memenuhi persyaratan mutu soal salah satunya adalah valid. Validitas soal adalah kecocokan atau ketepatan suatu tes dalam mengukur sesuatu yang hendak diukur.

Reliabilitas

Hasil pengujian reliabilitas tes kognitif secara empiris dengan menggunakan koefisien *Alpha Cronbach* diperoleh nilai untuk tes bentuk pilihan ganda paket A sebesar 0,879 dan untuk paket B sebesar 0,872. Hal ini menunjukkan bahawa hasil tes telah menunjukkan hasil yang tetap seperti yang dikemukakan oleh Arikunto (2010) bahwa reliabilitas tes berhubungan dengan ketepatan hasil tes, perubahan yang ada dapat dikatakan tidak berarti.

Berdasarkan hasil yang diperoleh maka dapat dikatakan bahwa tes kognitif yang dikembangkan dapat digunakan karena memiliki nilai reliabilitas yang tinggi. Pernyataan tersebut juga diperkuat oleh Mahaputri, dkk (2013) yang menyatakan bahwa tes yang dikembangkan memiliki kualitas yang baik dan telah memenuhi standar reliabilitas.

Tingkat Kesukaran

Berdasarkan hasil analisis terhadap tingkat kesukaran soal, tes kognitif yang telah dikembangkan memiliki tingkat kesukaran pada kategori mudah, sedang dan

sukar. Tes bentuk pilihan ganda paket A memiliki soal dengan tingkat kesukaran pada kategori mudah 28%, kategori sedang 52%, dan kategori sukar 20%. Hal ini sejalan dengan hasil penelitian Taufiq (2013) bahwa tes bentuk pilihan ganda memiliki tingkat kesukaran, untuk kategori mudah 30%, kategori sedang 48%, kategori sukar 22 %. Sedangkan untuk tes bentuk pilihan ganda paket B dengan tingkat kesukaran pada kategori mudah 24%, kategori sedang 56%, dan kategori sukar 20%. Soal yang berada pada kategori sukar dan mudah kemudian ditelaah kembali.

Menurut Sukiman (2012) tidak lanjut dari hasil analisis tingkat kesukaran butir soal ini adalah sebagai berikut: (1) Mencatat butir soal yang sudah baik (memiliki $ITK=cukup$) dalam buku bank soal, (2) Bagi soal yang sukar ada dua kemungkinan, yaitu: dibuang atau diteliti ulang dimana letak yang membuat soal tersebut sukar, mungkin kalimatnya yang tidak baik atau petunjuk mengerjakannya yang kurang jelas, dan sebagainya, kemudian setelah diperbaiki dipakai kembali; atau disimpan untuk kepentingan yang lain (seperti untuk tes seleksi), (3) Untuk butir yang mudah juga ada tiga kemungkinan dibuang atau diteliti ulang dimana letak yang membuat soal tersebut sukar, mungkin kalimatnya yang tidak baik atau petunjuk mengerjakannya yang kurang jelas, dan sebagainya, kemudian setelah diperbaiki dipakai kembali; atau disimpan untuk kepentingan yang lain (seperti untuk tes seleksi)

Daya Pembeda

Analisis daya merupakan pengkajian butir-butir soal yang dimaksudkan untuk membedakan siswa yang tergolong mampu dengan siswa yang tergolong tidak mampu (Uno, 2012). Berdasarkan hasil analisis terhadap tingkat daya beda soal, tes kognitif yang telah dibuat memiliki koefisien daya beda yaitu besarnya koefisien daya beda yaitu: 0,00 sampai 0,19 buruk; 0,20 sampai 0,29 perlu revisi 0,30 sampai 0,39 dapat diterima (tidak perlu direvisi); 0,40 sampai 0,1 baik. Tes bentuk pilihan ganda untuk paket A dan paket B dengan jumlah soal 50 nomor memiliki soal dengan daya pembeda pada 0,40 sampai 0,1 baik adalah 98% dan kategori 0,20 sampai 0,29 perlu revisi sebesar 2%.

Menurut Mansyur (2009) semakin tinggi daya pembeda suatu butir soal (mendekati 1) maka semakin baik butir soal tersebut dapat membedakan kelompok peserta yang memiliki kemampuan tinggi dengan kelompok siswa yang memiliki kemampuan rendah.

Efektivitas pengecoh (*Distractor*)

Sebuah pengecoh berfungsi disebut berfungsi dengan baik bila pengecoh tersebut dipilih oleh peserta didik yang kurang memahami materi. Analisis efektivitas pengecoh (*Distractor*) atau analisis pola jawaban dilakukan dengan menghitung peserta tes yang memilih tiap alternatif jawaban pada masing-masing *item* (Uno, 2012). Berdasarkan hasil analisis efektivitas pengecoh (*Distractor*) atau analisis pola jawaban, untuk soal paket A terdapat 7 butir soal yang pengecohnya kurang baik dari 25 soal atau sebesar 72% soal yang memiliki pengecoh yang berfungsi dengan baik. Sedangkan untuk soal paket B terdapat 7 butir soal yang pengecohnya kurang baik dari 25 soal atau

sebesar 72%. Pengecoh dikatakan berfungsi efektif apabila paling tidak ada siswa yang terkecoh memilih (Purwanto, 2013).

Dalam rangka meningkatkan mutu tes, perlu dilakukan analisis butir tes sehingga tes menjadi alat yang mampu mengukur sejauh mana pencapaian hasil belajar peserta didik. Surapranata (dalam Mansyur, 2009) menyatakan bahwa “salah satu tujuan dilakukannya analisis adalah untuk meningkatkan kualitas soal, yaitu apakah suatu soal (1) dapat diterima karena telah didukung oleh data statistik yang memadai, (2) diperbaiki, karena terbukti terdapat beberapa kelemahan, atau bahkan (3) tidak digunakan sama sekali karena terbukti secara empiris tidak berfungsi sama sekali”.

Analisis soal (*item analysis*) adalah suatu prosedur yang sistematis,. Analisis soal terutama dapat dilakukan untuk tes objektif. Analisis butir soal antara lain bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang baik, dan soal yang jelek. Dengan analisis soal dapat diperoleh informasi tentang kejelekan sebuah soal atau “petunjuk” untuk mengadakan perbaikan (Arikunto, 1999).

Soal yang berkualitas dapat membantu guru membedakan peserta didik yang sudah memahami materi yang telah diajarkan dengan peserta didik yang belum memahami materi. Kemampuan peserta didik menjawab soal yang diberikan bergantung dari kemampuan peserta didik memahami materi yang diajarkan. Oleh karena itu, menjadi kewajiban seorang guru untuk membuat sebuah tes yang terstandar sehingga tidak terjadi kesalahan dalam menentukan program pengajaran selanjutnya.

KESIMPULAN

Pengembangan Tes Kognitif pada Materi Sistem Pernapasan dengan menggunakan pengembangan model riset tipe *formatif research*. Berdasarkan hasil analisis dan pembahasan penelitian yang telah dilakukan, serta dihubungkan dengan rumusan masalah, maka dapat disimpulkan bahawa pengembangan tes kognitif pada materi sistem pernapasan yang dikembangkan berdasarkan hasil validasi ahli yang dilakukan menunjukkan bahwa tes kognitif pada materi sistem pernapasan tersebut memenuhi kriteria valid dan reliable dan telah memenuhi karakteristik butir tes yang meliputi tingkat kesukaran sedang, daya beda baik, dan kualitas pengecoh (*distractor*) baik pada materi sistem pernapasan.

DAFTAR PUSTAKA

- Ali, Ahmad (2014). Pengembangan Perangkat Asesmen Praktikum Anatomi Fisiologi Manusia Berbasis Keterampilan Proses Sains Mahasiswa Jurusan Pendidikan Biologi UIN Alauddin Makassar. *Jurnal Biotek (online)*, Vol.1.No.1(<http://journal.uinalauddin.ac.id/index.php/biotek/article/view/1686>, diakses 26 Juli 2017..

- Anwar, Syafri. (2006). *Peningkatan Kinerja Guru dalam Membuat Soal Objektif melalui Umpan Balik*. Jurnal Pendidikan dan Kebudayaan. No. 059, Tohun Ke-12, Morel 2006.
- Arikunto, Suharsimi. (2010). *Dasar-dasar Evaluasi Pendidikan (Edisi revisi)*. Jakarta: Bumi Aksara.
- Departemen Pendidikan Nasional. (2004). *Teknik Evaluasi*. Jakarta: Diknas.
- Haling. (2007). *Belajar dan Pembelajaran*. Makassar: UNM Makassar.
- Mahaputri, Ni Luh Putu., Nyoman Dantes, I Wayan Sadia. (2013). Pengembangan tes prestasi belajar berbasis taksonomi Anderson dan krathwohl pada kompetensi dasar fisika Smk kelas x semester ganjil se-kota singaraja. *Jurnal Penelitian Pascasarjana UNDIKSHA (online)*. Vol. 3 (<http://pasca.undiksha.ac.id/e-journal/index.php/jurnal-ep/article/view/763>, diakses 20 Juli 2016).
- Mansyur, Rasyid, H. & Suratno. (2009). *Asesmen Pembelajaran di Sekolah*. Yogyakarta: Multi Pressindo.
- Nurhayati. (2011). *Strategi Belajar Mengajar Biologi*. Makassar: Jurusan Biologi Universitas Negeri Makassar.
- Purwanto. (2013). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Rahayu, T., Purwoko & Zulkardi. (2008). Pengembangan Instrumen Penilaian dalam Pendidikan Matematika Realistik Indonesia (PMRI) di SMPN 17 Palembang. *Jurnal Pendidikan Matematika (Online)*, Vol. 2. No. 2 (<http://ejournal.unsri.ac.id/index.php/jpm/article/view/301>, diakses 20 Juli 2016).
- Republik Indonesia, Undang-Undang No. 20 Tahun 2003 Tentang Sistem Pendidikan Nasional. Cet. II; Jakarta: Sinar Grafika, 2009.
- Ruslan. (2005). *Prinsip Dasar Evaluasi*. Makassar: Disampaikan pada Diklat Guru Sekolah Dasar Mata pelajaran Matematika. Lembaga Penjamin Mutu Pendidikan.
- _____. (2009). Validitas Isi. *Buletin Pa'biritta* No. 10. Tahun VI, 18-19
- Safari. (2013). *Evaluasi Pembelajaran*. Jakarta: Departemen Pendidikan Nasional.
- Sukiman. (2012). *Pengembangan Sistem Evaluasi*. Yogyakarta: Insan Madani.
- Taufiq, Ainul U. (2013). Pengembangan Tes Kognitif Berbasis Revisi Taksonomi Bloom Pada Materi Sistem Reproduksi untuk Siswa SMA. *Jurnal Biotek (online)*, Vol. 3, No. 2 (<http://journal.uin-alauddin.ac.id/index.php/biotek/article/view/1014>, diakses 26 Juli 2017).
- Uno, Hamzah B., Koni, Satria. (2012). *Assessment Pembelajaran*. Jakarta: Bumi Aksara.
- Yahya, Yudrik. (2003). *Wawasan Kependidikan*. Jakarta: Departemen Pendidikan Nasional.
- Widodo. (2010). *Analisis Butir Soal Tes*. Jurnal Pendidikan Penabur - No.14/Tahun ke-9/Juni 2010