

## ABSTRAK

**Tondowala , Sulvia Fery Hanry (2012), Pengembangan tes objektif pilihan ganda Berbasis taksonomi anderson dan krathwohl Untuk kemampuan membaca bahasa inggris Kelas viii smp di Kabupaten Poso, Propinsi Sulawesi Tengah.**

Tesis, Penelitian dan Evaluasi Pendidikan, Program Pascasarjana, Universitas Pendidikan Ganesha.

Tesis ini sudah disetujui dan diperiksa oleh: Pembimbing I : Prof. Dr. Nyoman Dantes dan Pembimbing II : Prof. Dr. AAIN. Marhaeni, M.A.

*Kata Kunci:* Tes Objektif Pilihan ganda, Taksonomi Anderson dan Krathwohl, dan membaca pemahaman bahasa inggris.

Penelitian ini merupakan model penelitian pengembangan yang diadopsi dari model pengembangan Sugiyono. Penelitian bertujuan untuk mengembangkan tes objektif pilihan ganda berbasis taksonomi Anderson dan Krathwohl. Populasi penelitian yaitu SMP kelas VIII di Kabupaten Poso, Propinsi Sulawesi Tengah, dengan sampel penelitian berjumlah 250 orang siswa kelas VIII yang diambil dengan menggunakan tehnik *cluster sampling* dari 12 kelas pada 7 sekolah, dengan rincian sebagai berikut: 4 kelas dari SMP di kabupaten, 4 kelas dari SMP di Kecamatan dan 4 kelas dari SMP pedesaan.

Hasil penelitian menunjukkan, bahwa instrumen tes yang dikembangkan telah mempunyai nilai validitas dan uji empirik yang cukup baik. Berdasarkan hasil uji lapangan dan analisis data ditarik simpulan sebagai berikut; (1) Kisi-kisi tes objektif pilihan ganda yang berdasarkan taksonomi Anderson dan Krathwohl memungkinkan pembuatan butir tes yang bervariasi untuk setiap jenis proses kognitif. (2) Validasi tes berbasis taksonomi Anderson dan Krathwohl yang meliputi: (a) Validitas isi sudah memenuhi syarat dengan melakukan uji judges, judges menyatakan instrumen relevan untuk diujicobakan. (b) Uji empirik instrumen pun dinyatakan sudah memenuhi syarat karena semua ketentuan yang harus ditempuh seperti; uji validitas butir, uji tingkat kesukaran butir, daya beda dan uji efektivitas pengecoh, semuanya sudah dilakukan seperti yang telah ditentukan. (3) Reliabilitas tes objektif pilihan ganda berdasarkan taksonomi Anderson dan Krathwohl dinyatakan sudah memenuhi syarat karena konsistensi atau keajegan tes masih terhitung tinggi.

## ABSTRACT

Tondowala, Sulvia Fery Hanry (2012), *The Development of Anderson and Krathwohl Taxonomy Based Objective Test for English Reading Ability of Eight Grade Students of Junior High Schools in Poso Regency, Central Sulawesi Province*.

*Thesis*. Educational research and Evaluation, Postgraduate Program, Ganesha University of Education.

This thesis has been corrected and approved by

Supervisor I : Prof. Dr. Nyoman Dantes

Supervisor II : Prof. Dr. A.A.I.N. Marhaeni, M. A.

Keywords: objective test, Anderson and Krathwohl Taxonomy, reading comprehension.

This research is a developmental research model which is adopted from Sugiyono. It aims at developing objective test which is based on Anderson and Krathwohl Taxonomy. The population of this research was eight grade students of junior high schools in Poso Regency, Central Sulawesi Province, while the sample was 250 eight grade students of which was taken using cluster sampling technique from twelve classes of seven schools: four classes from the junior high schools in the regency, four classes from the junior high schools in the subdistric, and four classes from the junior high schools in village area.

The findings of this research show that the instrument of the developed test has had adequate validity value and empirical test. Based on the field test and data analysis, the conclusion can be made as follows; (1) the blueprint of objective test which is based on Anderson and Krathwohl Taxonomy enables the development of the varied objective test for every type of cognitive process, (2) the validation tests which are based on Anderson and Krathwohl Taxonomy which include: (a) the content validity has fulfilled the requirement through the process of judges test, the judges stated that the instrument is relevant to be tried out, (b) the empirical test proved that the instrument has fulfilled the requirement since all of the requirements; item validity test, index of difficulty test, index of discrimination test, and distracter effectiveness test have been done properly, (3) the reliability test for the objective test proved that the instrument has fulfill the requirement since the consistency of the test can be categorized as high.

## **I. PENDAHULUAN**

Dalam khazanah profesionalnya sehari-hari seorang guru tidak mungkin melepaskan diri dari kegiatan penilaian. Untuk melakukan suatu penilaian seorang guru membutuhkan sebuah instrumen untuk mengetahui serta mengukur tingkat kemampuan kognitif keberhasilan seorang siswa. Biasanya Instrumen yang digunakan untuk hal tersebut dikenal dengan nama tes dan kegiatan itu dilakukan pada waktu-waktu tertentu. Kadang kala dilakukan secara teratur setiap satu bulan. Ada pula yang dilakukan secara teratur pada setiap akhir suatu unit atau satuan pelajaran tertentu. Juga kadang kala dilakukan pada setiap akhir suatu pertemuan kelas. Berdasarkan hasil yang diperoleh dari kegiatan ulangan atau tes tersebut kemudian guru membuat keputusan-keputusan. Keputusan tentang siswa, keputusan tentang proses belajar-mengajar, keputusan tentang rencana pelajaran, keputusan tentang materi pelajaran, metode pengajaran dan sebagainya. Sering memang bahwa keputusan yang diambil tidaklah sebanyak daftar yang ada atau yang seharusnya dilakukan. Tetapi tak dapat dibantah bahwa keputusan demi keputusan diambil oleh guru setelah melakukan tes.

Menurut Koyan penilaian, tes, evaluasi dan pengukuran merupakan empat istilah yang berbeda namun saling berhubungan (Koyan, 2011). Banyak orang tidak mengetahui secara jelas perbedaan dan hubungan diantara ketiganya, sehingga istilah tersebut sering tidak tepat penggunaannya. Agar jelas berikut ini akan diuraikan perbedaan dan hubungan antara Penilaian, tes, evaluasi dan pengukuran.

Menurut Linn dan Gronlund (dalam Koyan, 2011:6), asesmen (*assessment*) adalah istilah umum yang melibatkan semua rangkaian prosedur yang digunakan untuk mendapatkan informasi tentang hasil belajar peserta didik (misalnya: observasi, skala bertingkat tentang kinerja, tes tertulis) dan pelaksanaan penilaian mengenai kemajuan belajar peserta didik. Sedangkan *test* adalah tipe khusus dari asesmen yang secara khusus terdiri atas seperangkat pertanyaan yang dilaksanakan pada periode waktu tertentu sampai dengan dapat membandingkan semua peserta didik. Di lain pihak, *measurement* atau pengukuran ialah pemberian tanda atau angka pada hasil sebuah tes atau bentuk lain dari asesmen menurut aturan tertentu. Linn dan Gronlund (dalam Koyan, 2011:6) pengertian asesmen hampir sama dengan pengertian evaluasi (*evaluation*), tetapi asesmen memberi penekanan yang lebih besar pada kinerja tugas-tugas pada bentuk nyata dan kompleks. Penggunaan istilah asesmen tampaknya lebih ramah dan saat ini cenderung digunakan secara bergantian atau bersama-sama dengan istilah evaluasi.

Dengan demikian, jelaslah bahwa *assessment* memiliki pengertian yang lebih luas daripada pengertian *evaluation*, *measurement*, dan *test*. Hubungan dengan ini, Brown (dalam Koyan, 2011:6) menyatakan bahwa pengukuran adalah pemberian tanda dengan angka terhadap perilaku menurut aturan tertentu. Sedangkan Kerlinger (dalam Koyan, 2011:6) menyatakan bahwa pengukuran adalah pemberian angka pada objek-objek atau kejadian-kejadian menurut suatu aturan. Dengan demikian, dapat disimpulkan bahwa pengukuran adalah proses

kuantifikasi atau pemberian tanda dengan bilangan atau angka kepada objek atau perilaku tertentu menurut aturan-aturan tertentu.

Mengenai tes sendiri, Gronlund (dalam Koyan, 2011:7) menyatakan bahwa tes hasil belajar adalah suatu prosedur yang sistematis untuk mengukur sampel perilaku yang representatif tentang tugas-tugas pembelajaran peserta didik. Pendapat lain menyatakan bahwa tes adalah seperangkat pertanyaan atau tugas-tugas untuk menentukan bentuk-bentuk respon yang berkenaan dengan perilaku peserta didik. Sedangkan Nitko (dalam Koyan, 2011:7) menyatakan bahwa tes adalah suatu instrumen atau prosedur yang sistematis untuk mengobservasi dan menggambarkan satu atau lebih ciri-ciri peserta didik dengan menggunakan skala numerik atau klasifikasi. Dan menurut Koyan (2011) tes adalah instrumen atau alat atau prosedur yang sistematis yang terdiri atas seperangkat pertanyaan atau tugas-tugas untuk mengukur suatu perilaku tertentu pada peserta didik dengan menggunakan bantuan skala numerik atau kategori tertentu. Tes hasil belajar adalah alat atau instrumen untuk mengukur hasil belajar, baik hasil belajar pada ranah kognitif, afektif, maupun psikomotorik. Untuk melaksanakan asesmen terhadap proses dan hasil belajar, terlebih dahulu perlu dilakukan pengukuran terhadap sasaran ukur (atribut orang, objek, peristiwa) (Koyan, 2011:7).

Berdasarkan pendapat-pendapat tersebut di atas disimpulkan bahwa tes merupakan sebuah alat atau instrumen yang dijadikan sebagai sarana dalam mengukur suatu perilaku tertentu pada peserta didik dengan menggunakan skala pengukuran tertentu.

Menurut Koyan, perlu dibedakan pengertian antara "prestasi belajar" (*achievement*) dan "hasil belajar" (*learning outcome*). Hasil belajar meliputi aspek pembentukan watak atau karakter afektif, kognitif, dan psikomotor, sedangkan prestasi belajar lebih mengacu pada aspek kognitif atau pengetahuan saja. Tes untuk mengukur prestasi belajar, baik untuk aspek pengetahuan maupun aspek keterampilan, dapat dibagi menjadi dua jenis, yaitu "tes kemampuan" (*power test*) dan "tes kecepatan" (*speed test*). Prinsip dari "power test" adalah tidak adanya batasan waktu didalam pengerjaan tes. Sebaliknya, pada "speed test", yang diukur adalah kecepatan di dalam memikirkan atau mengerjakan tes (Koyan, 2011).

Menurut aturan (Kerlinger dalam Koyan, 2011:8) mengenai pengukuran (*measurement*) adalah pemberian angka pada objek atau peristiwa. Pendapat yang hampir sama dikemukakan oleh Wiersma dan Jurs (dalam Koyan, 2011:8) bahwa pengukuran adalah pemberian angka pada objek atau peristiwa menurut aturan yang memberikan arti kuantitatif kepada angka itu. Gronlund (dalam Koyan, 2011:8) menyatakan bahwa pengukuran adalah proses untuk memperoleh deskripsi angka tentang derajat karakteristik tertentu yang dimiliki oleh individu. Selanjutnya Koyan menyimpulkan bahwa pengukuran adalah pemberian bilangan atau proses kuantifikasi kepada atribut orang, objek, atau peristiwa menurut aturan tertentu. Dalam kaitan ini sasaran ukurnya adalah atribut orang, objek, atau suatu peristiwa, diukur dengan alat ukur tertentu, cara mengukurnya, sasaran ukurnya atau respondenya, sehingga diperoleh skor atau bilangan. Alat ukurnya dapat berupa tes atau non tes dan menurut aturan tertentu. Skor ditransformasi menjadi nilai (Koyan, 2011).

Berkaitan dengan proses belajar peserta didik untuk mengetahui sejauh mana kompetensi yang telah mereka capai salah satu cara untuk melihat peningkatan kemampuan tersebut adalah dengan melakukan tes. Tes yang berkaitan dengan tujuan ini sering disebut tes prestasi hasil belajar (TPHB). Azwar (dalam Sudiono, 2011) menyatakan bahwa tes prestasi hasil belajar adalah tes yang disusun secara terencana untuk mengungkap informasi subyek atas bahan-bahan yang telah diajarkan. Sudijono juga menyatakan juga bahwa tes prestasi hasil belajar adalah tes yang digunakan untuk mengungkap tingkat pencapaian belajar (Sudiono, A, 2011).

Dari beberapa pengertian di atas, ada satu benang merah yang sepertinya disepakati yaitu bahwa tes prestasi hasil belajar merupakan salah satu cara untuk menelusuri kemampuan-kemampuan yang telah dimiliki siswa setelah mengikuti proses belajar mengajar selama waktu tertentu. Meskipun tes bukanlah satu-satunya cara untuk mengungkap hasil belajar siswa, tetapi ia merupakan alat yang paling sering digunakan karena kepraktisan penggunaannya serta biaya yang murah. Tidak seperti alat pengukur ilmu alam yang tunggal, alat pengukur dalam ilmu-ilmu sosial dapat terdiri lebih dari satu macam.

Tes sendiri menurut Koyan, jika ditinjau dari bentuk soalnya dapat dibedakan menjadi dua yaitu tes hasil belajar dalam bentuk uraian (non obyektif dan tes hasil belajar bentuk obyektif). Disebut tes obyektif karena siapapun yang memeriksa hasil tes akan menghasilkan skor yang sama sedangkan tes uraian hasilnya dipengaruhi oleh pemberi skor (Koyan, 2011).

Menurut Koyan Tes objektif adalah salah satu jenis tes hasil belajar yang terdiri dari butir-butir soal (items) yang dapat dijawab oleh testee dengan jalan memilih salah satu atau lebih jawaban di antara beberapa kemungkinan jawaban yang telah dipasangkan pada masing-masing items, atau dengan jalan menuliskan (mengisikan) jawaban berupa kata-kata atau simbol-simbol tertentu pada tempat yang telah disediakan untuk masing-masing butir item yang bersangkutan (Koyan, 2011). Tes objektif dapat dibedakan menjadi lima golongan, yaitu sebagai berikut, a) *tes objektif bentuk benar-salah ( True-False test)*, b) *tes objektif bentuk menjodohkan (Matching Test)*, c) *tes objektif bentuk melengkapi (Completion Test)*, d) *tes objektif bentuk isian (Fill in Test)*, e) *tes objektif bentuk pilihan ganda (Multiple choice Item Test)* (Koyan, 2011). Tes objektif yang banyak dipakai dalam evaluasi hasil belajar siswa di sekolah adalah tes objektif pilihan ganda. Tes pilihan ganda memiliki semua persyaratan sebagai tes yang baik, yakni dilihat dari segi objektivitas, reliabilitas, dan daya pembeda antara siswa yang berhasil dengan siswa yang gagal atau bodoh. Sebagian besar guru merasakan bahwa tes objektif tipe pilihan ganda juga efektif dalam mengungkap materi pembelajaran dengan cakupan pengetahuan yang lebih kompleks, dengan tingkat pengetahuan yang lebih tinggi (Koyan, 2011).

Tes objektif bentuk pilihan ganda merupakan tes yang memiliki satu pemberitahuan tentang suatu materi tertentu yang belum sempurna serta beberapa alternatif jawaban yang terdiri dari kunci jawaban dan pengecoh. Tugas peserta tes adalah memilih jawaban dari pilihan yang tersedia dan paling sesuai dengan pernyataan yang ada dalam soal (Koyan, 2011).



Bagi guru dan praktisi pendidikan, kata “Jenjang Kognitif atau sering disingkat “C” (dari kata Cognitive) merupakan istilah yang sangat akrab dalam merumuskan tujuan pembelajaran. Istilah kognitif dimaksud diambil dari buku *“The Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook I: Cognitive Domain”* yang dikenal dengan taksonomi belajar yang dikemukakan oleh Bloom, Engelhart, Furst, Hill dan Krathwohl (Widodo, 2006). Selama hampir setengah abad buku itu banyak menjadi rujukan di berbagai negara, termasuk Indonesia. Sangat banyak manfaat dari taksonomi belajar versi Bloom ini, banyak ide yang digunakan mengacu dari taksonomi belajar versi Bloom inilah, pada perkembangan dan temuan baru dalam dunia pendidikan. Salah satu hasil dari pengembangan itu adalah edisi revisi buku tersebut yang berjudul *“A Taxonomy for Learning and Teaching and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives”* oleh Anderson dan Krathwohl (Widodo, 2006)

Dalam tulisannya Widodo mencoba mengemukakan prinsip perubahan taksonomi Bloom menjadi taksonomi Anderson, menurut Widodo perubahan dari kata benda menjadi kata kerja itulah yang salah satu perubahan yang sangat signifikan (Widodo, 2006). Menurut Anderson dan Krathwohl (dalam Widodo, 2006) taksonomi perlu mencerminkan berbagai bentuk atau cara berpikir dalam suatu proses yang aktif, oleh karena itu kata kerja lebih lebih sesuai daripada kata benda. Misalnya (menurut Bloom) merupakan hasil berpikir bukan cara berpikir sehingga diperbaiki menjadi mengingat yang menunjukkan proses berpikir tingkat awal (Widodo, 2006).

Taksonomi kognitif Bloom terdiri dari enam tingkatan kognitif, yaitu pengetahuan (*knowledge*), pemahaman (*comprehension*), aplikasi (*apply*), analisis (*analysis*), sintesis (*synthesis*), dan evaluasi (*evaluation*) yang kemudian Anderson dan Krathwohl merevisinya dari satu dimensi menjadi dua dimensi, yaitu dimensi proses kognitif (*cognitive process*) dan dimensi pengetahuan (*types of knowledge*) (Widodo, 2006).

Anderson dan Krathwohl (dalam Widodo, 2006) mengklasifikasikan proses kognitif menjadi enam kategori, yaitu ingatan (*remember*), pemahaman (*understand*), aplikasi (*apply*), analisis (*analyze*), evaluasi (*evaluate*), dan kreatifitas (*create*). Dimensi pengetahuan diklasifikasi menjadi empat kategori, yaitu pengetahuan faktual (*factual knowlwdge*), pengetahuan konseptual (*conceptual knowledge*), pengetahuan prosedural (*procedural knowledge*), dan pengetahuan metakognisi (*metacognitive knowledge*).

Pada pendidikan formal dalam mengukur aspek kognitif siswa semuanya telah dirumuskan dalam kurikulum KTSP. Walaupun dalam kurikulum KTSP telah diberikan kewenangan kepada pihak sekolah untuk merumuskan indikator-indikator pembelajaran dari Standar Kompetensi dan Kompetensi Dasar yang telah ditetapkan, namun selama ini dalam merumuskan setiap indikator-indikator dari setiap standar kompetensi dan kompetensi dasar itu masih ada kesulitan dan ketidakpuasan guru pada taksonomi bloom yang menjadi pedoman umum KTSP, hal ini disebabkan ketidakleluasaan guru untuk merumuskan butir-butir tes dari setiap indikator menjadi lebih bervariasi karena dimensi pengetahuan masih disatukan dengan setiap jenjang proses kognitif, bahkan jenjang pengetahuan

jarang sekali dimunculkan dan dimasukkan pada jenjang paling bawah. Sementara menurut Anderson dan Krathwohl dalam Widodo (2006) dimensi pengetahuan dan dimensi proses kognitif merupakan dua hal yang berbeda namun sama pentingnya dalam proses pembelajaran, sebab dimensi pengetahuan merupakan kata benda sedangkan dimensi proses kognitif merupakan kata kerja.

Berkaitan dengan penilaian dan evaluasi, instrumen tes merupakan elemen yang memiliki peranan penting. Oleh karena itu instrumen tes sebagai alat ukur haruslah memenuhi syarat-syarat yang memang telah ditetapkan. Namun sungguh disayangkan, sudah bukan rahasia lagi jika instrumen tes yang sering digunakan untuk penilaian oleh guru-guru secara umum hampir diseluruh indonesia hanya dibuat secara dadakan, tanpa mempertimbangkan kelayakan tes sebagai alat ukur. Tahapan perumusan instrumen tes dimulai dari perumusan kisi-kisi dan validasi tes kurang diperhatikan dan bahkan tidak dilakukan sama sekali sehingga jarang sekali didapatkan instrumen tes yang sudah memenuhi syarat seperti yang telah ditentukan.

Melihat fenomena ini, muncullah sebuah keinginan untuk melakukan penelitian pengembangan untuk menghasilkan sebuah instrumen tes yang layak dan telah memenuhi ketetapan-ketetapan yang sesuai aturan yang berlaku. Penelitian pengembangan ini akan dicobakan dengan menggunakan taksonomi Anderson dan Krathwohl dalam merumuskan Standar kompetensi dan kompetensi dasar melalui kisi-kisi yang untuk menghasilkan butir-butir tes. Diharapkan rumusan kisi-kisi dengan taksonomi Anderson dan Krathwohl lebih memungkinkan pembuatan butir tes yang bervariasi untuk setiap jenis proses kognitif. Apabila

dalam taksonomi bloom, hanya dikenal jenjang C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, dan seterusnya. Dalam taksonomi yang baru tiap jenjang menjadi empat kali lipat sebab ada empat macam pengetahuan. Guru yang membuat soal jenjang C<sub>1</sub>, kini bisa memvariasikan soalnya, menjadi C<sub>1</sub>-faktual, C<sub>1</sub>-konseptual, C<sub>1</sub>-prosedural, C<sub>1</sub>-metakognitif, dan seterusnya.

Untuk mempermudah penelitian ini, fokus penelitian hanya pada standar kompetensi dan kompetensi dasar membaca siswa dalam bahasa Inggris yang dirumuskan dalam ke dalam indikator-indikator butir tes. Hasil penelitian ini diharapkan menjadi *prototype* dalam menggunakan taksonomi Anderson dan Krathwohl dalam merumuskan standar kompetensi dan kompetensi dasar kedalam indikator-indikator butir tes yang diinginkan terutama tes pilihan ganda.

Dalam penelitian ini fokus peneliti hanya membahas terkait pada pengembangan tes objektif pilihan ganda yang dibedah secara empiris konseptual mengenai, proses pengembangan tes objektif pilihan ganda yang valid dan reliabel berdasarkan taksonomi Anderson dan Krathwohl pada SMP kelas VIII di Kabupaten Poso, Sulawesi Tengah.

### **Perumusan Masalah**

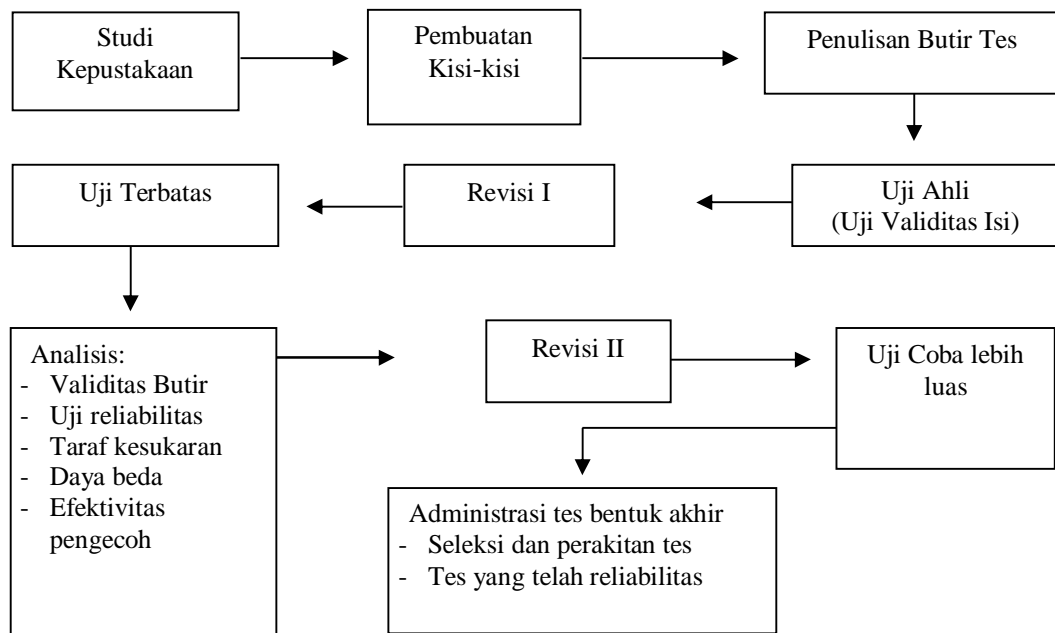
Berdasarkan latar belakang yang telah diuraikan di atas, maka dirumuskan masalah utamanya yaitu, bagaimanakah mengembangkan tes objektif pilihan ganda berbasis taksonomi Anderson dan Krathwohl untuk kemampuan membaca Bahasa Inggris kelas VIII SMP di Kabupaten Poso Propinsi Sulawesi Tengah?. Masalah utama tersebut dijabarkan menjadi tiga masalah khusus yaitu:

- 1) Bagaimanakah kisi-kisi tes objektif pilihan ganda yang berdasarkan taksonomi Anderson dan Krathwohl?
- 2) Apakah Validasi tes berdasarkan taksonomi Anderson dan Krathwohl sudah terpenuhi? yang meliputi:
  - a) Validitas isi dengan melakukan uji judges
  - b) Uji empirik dengan melakukan uji validitas butir, uji tingkat kesukaran butir, daya beda dan uji efektivitas pengecoh, semuanya sudah memenuhi syarat seperti yang telah ditentukan?
- 3) Apakah reliabilitas tes objektif pilihan ganda berdasarkan taksonomi Anderson dan Krathwohl sudah memenuhi syarat?

## **II. METODE PENELITIAN**

### **Rancangan Penelitian**

Penelitian ini dikategorikan ke dalam studi penelitian dan pengembangan, langkah-langkah yang akan dijadikan acuan dalam penelitian ini yaitu pendapat sugiyono. Mengacu dari rancangan yang dikemukakan oleh Sugiyono inilah dibuat sebuah rancangan penelitian untuk pengembangan tes objektif pilihan ganda ini dengan gambar seperti berikut ini.



Gambar 3.1 Prosedur Penelitian Pengembangan Tes

### Populasi dan Sampling

Dalam penelitian dan pengembangan penentuan sampel harus disesuaikan dengan butir instrumen yang akan dikembangkan. menurut Nunnally banyaknya sampel untuk penelitian adalah 5 sampai 10 kali lipat banyaknya butir yang hendak dikembangkan atau dianalisis (dalam Rajeg, 2006:85). Sementara Comrey (dalam Rajeg, 2006: 85) membuat patokan besar sampel seperti berikut.

Tabel. 3.1 Patokan Besar Sampel

Besar sampel	Kualitas
50	Sangat Kurang
100	kurang
200	Cukup
300	Baik
500	Sangat Baik
1000	Unggul

(Sumber: Nunnaly dalam Rajeg, 2006)

Objeknya pada penelitian yaitu pada perancangan tes pilihan ganda dalam mata pelajaran bahasa Inggris dikhususkan pada membaca pemahaman siswa di semester genap dengan menganalisis Standar Kompetensi dan Kompetensi Dasar dalam mata pelajaran bahasa Inggris semester genap dan yang akan menjadi subjek penelitian hanya difokuskan pada SMP kelas VIII yang berada di kabupaten Poso, Propinsi Sulawesi Tengah. Penelitian akan dilakukan 2 kali dalam hal ini untuk pengujian instrumen tes (uji empirik pertama dengan sampel terbatas dan empirik kedua dengan jumlah sampel lebih besar).

Penelitian ini pengambilan sampel berdasarkan teknik *cluster sampling*. Yang akan dijadikan sampel penelitian yaitu siswa di tingkat SMP kelas VIII yang mewakili populasi penelitian dari yang telah dipilih. Adapun pengambilan sampel pada populasi penelitian dipilih berdasarkan klasifikasi sebagai berikut, SMP di pedesaan, kecamatan dan Kabupaten.

Penentuan besaran sampel dalam penelitian ini mengacu pada pendapat Nunnally oleh karena itu karena dalam penelitian ini jumlah butir tes yang akan dikembangkan atau dianalisis adalah 50, jadi jumlah sampel pada penelitian 5-10 x 50. Untuk memudahkan penelitian ini, angka standar akan digunakan dikalikan jumlah butir,  $5 \times 50 = 250$ . Jadi jumlah sampel pada penelitian ini adalah 250 orang siswa yang termasuk dalam populasi.

*Tabel 4. 1 Sampel Pada Uji Empirik Pertama*

No	Nama	Pedesaan	Kecamatan	Kabupaten	Jumlah kelas	Jumlah siswa
A	B	C	D	E	F	G
	SMP Negeri 1 Pamona Barat		√		1	25

Tabel 4. 2 Sampel Pada Uji Empirik Kedua

No	Nama	Pedesaan	Kecamatan	Kabupaten	Jumlah kelas	Jumlah siswa
A	B	C	D	E	F	G
	SMP Satu Atap Matia lemba	√			1	21
	SMP Satu Atap Masewe	√			1	20
	SMP Satu Atap Tampemadoro	√			1	20
	SMP Satu Atap Barati	√			1	21
	SMPN 2 Pamona Timur		√		2	42
	SMP 2 Kr. GKST Pamona Utara		√		2	42
	SMPN 1 Poso Kota			√	4	84

### Studi Kepustakaan

Aspek kognitif pada penelitian ini hanya pada aspek membaca yang akan di ukur, mengacu pada standar kompetensi dan kompetensi dasar membaca yang ada dalam bahasa inggris SMP kelas VIII semester genap.

Ada begitu banyak pendapat para ahli mengenai metode untuk mengukur membaca pemahaman siswa, agar pada penelitian ini tidak terjadi tumpang tindih pemahaman oleh karena itu dalam mengukur membaca pemahaman hanya akan berpedoman pada pendapat yang dikemukakan oleh Farr.

Adapun Farr (dalam Djiwandono, 1998:98) mengemukakan bahwa untuk mengukur pemahaman bacaan di antaranya haruslah berisi pertanyaan tentang



pandangan atau maksud pengarang dan pertanyaan tentang kesimpulan bacaan.

Secara terinci Farr membagi pertanyaan itu menjadi sembilan, yaitu :

- 1) Pengetahuan tentang makna kata;
- 2) Kemampuan memilih makna yang dimiliki kata atau frasa dalam latar kontekstual khusus;
- 3) Kemampuan untuk memilih atau memahami susunan dari bacaan dan identitas sebelumnya dan kesimpulan-kesimpulan di dalamnya.
- 4) Kemampuan menyeleksi gagasan pokok melalui bacaan;
- 5) Kemampuan menjawab pertanyaan-pertanyaan yang dijawab khusus dalam suatu bacaan;
- 6) Kemampuan menjawab pertanyaan-pertanyaan yang dijawab dalam bacaan, tetapi tidak pada setiap kata-kata yang mana pertanyaan dijawab;
- 7) Kemampuan menyimpulkan dari bacaan tentang isinya;
- 8) Kemampuan mengingat apa yang ditulis dalam bacaan dan maksud dan suara hati pengarang, dan
- 9) Kemampuan menentukan tujuan-tujuan pengarang, maksud pengarang, dan pandangan pengarang, yaitu membuat kesimpulan-kesimpulan tentang suatu tulisan.

Untuk mengukur membaca pemahaman siswa disesuaikan dengan standar kompetensi dan kompetensi dasar membaca dalam bahasa inggris pada siswa SMP kelas VIII semester genap dalam silabus.

Penelitian ini bertujuan sebagai berikut: (1) untuk mendeskripsikan tahapan pengembangan tes objektif pilihan ganda berbasis taksonomi Anderson dan Krathwohl untuk kemampuan membaca Bahasa Inggris kelas VIII SMP di Kabupaten Poso Propinsi Sulawesi Tengah. 2) untuk mendeskripsikan *validity content* pengembangan tes objektif pilihan ganda berbasis taksonomi Anderson dan Krathwohl untuk kemampuan membaca Bahasa Inggris kelas VIII SMP di Kabupaten Poso Propinsi Sulawesi Tengah, (3) untuk mendeskripsikan validitas dan reliabilitas empiris pengembangan tes objektif pilihan ganda berbasis taksonomi Anderson dan Krathwohl untuk kemampuan membaca Bahasa Inggris kelas VIII SMP di Kabupaten Poso Propinsi Sulawesi Tengah, dan (4) untuk mendeskripsikan uji taraf kesukaran butir, daya beda dan uji efektifitas pengecoh pada tes objektif pilihan ganda berbasis taksonomi Anderson dan Krathwohl untuk kemampuan membaca Bahasa Inggris kelas VIII SMP di Kabupaten Poso Propinsi Sulawesi Tengah.

### **III. HASIL PENELITIAN DAN PEMBAHASAN**

#### **Pengolahan dan Analisis Data**

*Pertama*, Uji validitas isi dimaksudkan untuk mengetahui kesesuaian antara teori dan butir tes yang dibuat.

Dari penilaian pakar terhadap tes objektif pilihan ganda yang berjumlah 54 butir tes, semua butir tes dinyatakan relevan oleh kedua pakar dari butir tes nomor

1 – 54. Walaupun semua butir dinyatakan relevan oleh kedua pakar, namun terdapat beberapa masukan dari kedua pakar, yakni: (1) tiap indikator, minimal diwakili 3 butir pertanyaan yang setara, (2) menutup kata diakhir kalimat menggunakan empat titik, di tengah kalimat tiga titik, (3) perhatikan penggunaan huruf kapital dalam penulisan butir soal (*lihat lampiran 4*).

Dengan adanya koreksi dan saran dari kedua pakar pada butir tes, dilakukan sedikit revisi untuk memenuhi validitas isi seperti yang diharapkan, sesudah itu dilanjutkan dengan analisis validitas isi untuk mengetahui tingkat efisiensinya dengan menggunakan tehnik yang dikembangkan oleh Gregory

*Tabel 4.3 Uji Pakar Dengan Formula Gregory*

		Pakar I: Prof. Dr. I Wayan Koyan, M.Pd	
		KR	SR
Pakar II: Prof. Dr. AAIN Marhaeni, MA	KR	A (0)	B (0)
	SR	C (0)	D (54)

Keterangan:

1. KR : Kurang Relevan
2. SR : Sangat Relevan

Dari rekapitulasi di atas diketahui A= 0, B= 0, C = 0, D = 54

$$\text{Validitas isi} = \frac{D}{A+B+C+D} = \frac{54}{0+0+0+54} = 1,00$$

Koefisien bergerak dari + s/d 1, dengan kriteria :

- 0,9 – 1,0      → sangat tinggi
- 0,6 – 0,89    → tinggi
- 0,4 – 0,59    → sedang
- 0,2 – 0,39    → rendah
- 0,0 – 0,19    → sangat rendah

Dengan demikian mengacu dari kriteria yang ditetapkan dapat dikatakan bahwa validitas isi yang koefisiennya 1,00 tergolong sangat baik.

### **Uji Empirik Pertama**

Uji empirik pertama dilakukan untuk mendapatkan data awal seberapa dalam tingkat validitas butir dan reliabilitas butir. Pada penelitian ini, uji empirik pertama ini diberikan pada sampel terbatas yang berjumlah 25 orang. Sesudah butir tes diujikan, butir tes akan dianalisis dengan langkah-langkah sebagai berikut.

### **Uji Validitas Butir Pada Uji Pertama**

Validitas butir tes dianalisis dengan mengkorelasikan sekor butir dengan sekor total yang diperoleh respon. Rumus yang akan digunakan sebagai berikut.

$$r_{pbi} = \frac{M_p - M_t}{s_t} \sqrt{\frac{p}{q}}$$

#### Keterangan:

- $M_p$  = Rerata sekor total dari subjek yang menjawab betul butir yang dicari validitasnya.
- $M_t$  = rerata sekor total
- $S_t$  = Standar deviasi sekor total
- $p$  = proporsi siswa yang menjawab benar butir yang dicari validitasnya
- $q$  = proporsi siswa yang menjawab salah butir yang dicari validitasnya ( $q = 1-p$ )

Dalam uji validitas butir tes ini, tiap butir dianalisis dengan menggunakan korelasi point biserial melalui program *microsoft office excel*, dengan  $n= 25$  pada taraf signifikansi 0,05 didapatkan  $r$  tabel =0,396. Hasil analisis dapat di lihat pada tabel di bawah ini.

Tabel 4.4 Butir Tes yang Valid/Drop pada Empirik Pertama

<b>Butir</b>	<b>r-pbi</b>	<b>Pada r-tabel 25 (0,05)</b>	<b>Ket</b>
1	0,544	0,396	Valid
2	0,609	0,396	Valid
3	0,764	0,396	Valid
4	0,548	0,396	Valid
5	0,461	0,396	Valid
6	0,514	0,396	Valid
7	0,208	0,396	Drop
8	0,489	0,396	Valid
9	0,466	0,396	Valid
10	0,764	0,396	Valid
11	0,594	0,396	Valid
12	0,583	0,396	Valid
13	0,071	0,396	Drop
14	0,556	0,396	Valid
15	0,764	0,396	Valid
16	0,474	0,396	Valid
17	0,664	0,396	Valid
18	0,454	0,396	Valid
19	0,501	0,396	Valid
20	0,446	0,396	Valid
21	0,559	0,396	Valid
22	0,461	0,396	Valid
23	0,734	0,396	Valid
24	0,594	0,396	Valid
25	0,525	0,396	Valid
26	0,508	0,396	Valid
27	0,550	0,396	Valid
28	0,474	0,396	Valid
29	0,474	0,396	Valid
30	0,594	0,396	Valid
31	0,466	0,396	Valid
32	0,533	0,396	Valid
33	0,462	0,396	Valid
34	0,746	0,396	Valid
35	0,507	0,396	Valid
36	0,694	0,396	Valid
37	0,474	0,396	Valid

*Tabel Lanjutan*

<b>Butir</b>	<b>r-pbi</b>	<b>Pada r-tabel 25 (0,05)</b>	<b>Ket</b>
<b>38</b>	0,120	0,396	Drop
<b>39</b>	0,396	0,396	Valid
<b>40</b>	0,483	0,396	Valid
<b>41</b>	0,432	0,396	Valid
<b>42</b>	0,415	0,396	Valid
<b>43</b>	0,452	0,396	Valid
<b>44</b>	0,427	0,396	Valid
<b>45</b>	0,412	0,396	Valid
<b>46</b>	0,489	0,396	Valid
<b>47</b>	0,746	0,396	Valid
<b>48</b>	0,616	0,396	Valid
<b>49</b>	0,491	0,396	Valid
<b>50</b>	0,462	0,396	Valid
<b>51</b>	0,715	0,396	Valid
<b>52</b>	0,612	0,396	Valid
<b>53</b>	-0,199	0,396	Drop
<b>54</b>	0,605	0,396	Valid

Hasil analisis menunjukkan bahwa uji validitas butir dari 54 butir tes terdapat 50 butir tes yang valid yakni: 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 54, dan terdapat 4 butir tes yang drop atau gugur yakni; 7, 13, 38 dan 53.

### **Uji Reliabilitas tes**

Langkah berikut dari rangkaian uji empirik pertama, yaitu uji reliabilitas butir. Reliabilitas hakikatnya bertujuan untuk mengetahui sejauh mana hasil suatu pengukuran dapat dipercaya atau memiliki keajegan. Uji reliabilitas ini dilakukan

setelah didapatkan butir-butir yang dinilai valid, dari 54 butir tes setelah di uji validitasnya didapatkan 50 butir tes yang valid.

Untuk menganalisis tingkat reliabilitas butir digunakan formula KR-20,

Formula KR-20 adalah seperti berikut:  $r_{1.1} = \left( \frac{k}{k-1} \right) \left( \frac{SDt^2 - \sum pq}{SDt^2} \right)$

Untuk menganalisis reliabilitas butir, dilakukan perhitungan secara manual dengan mengambil beberapa harga tertentu dalam perhitungan *microsoft office excel*, dapat diketahui,  $n = 25$ ,  $k = 50$ ,  $\sum pq = 11,6896$

Jadi,

$$SD = \sqrt{\frac{25 \sum X^2 - (\sum X)^2}{n(n-1)}} = \sqrt{\frac{25 \sum 15270 - (581)^2}{25(24-1)}}$$

$$SD = \sqrt{\frac{381750 - 268324}{600}} = \sqrt{\frac{113426}{600}}$$

$$SD = \sqrt{189,035} = 13,749$$

$$SD^2 = (13,749)^2$$

$$SD^2 = 189,035$$

$$r_{1.1} = \left( \frac{k}{k-1} \right) \left( \frac{SDt^2 - \sum pq}{SDt^2} \right) = \left( \frac{50}{50-1} \right) \left( \frac{189,035 - 11,6896}{189,035} \right)$$

$$r_{1.1} = 0,957308 = 0,96$$

Keterangan:

$r_{1.1}$  = koefisien reliabilitas tes

$n$  = banyaknya testee

$k$  = banyak butir tes

$SDt$  = varian total tes

$p$  = proporsi testee yang menjawab benar

$q$  = proporsi testee yang menjawab salah

$pq$  =  $pxq$

Kriteria derajat reliabilitas tes atau instrumen evaluasi dapat digunakan kriteria yang dibuat oleh Guiford dalam Candi (2010:107), seperti berikut ini:

$r_{1.1} \leq 0,20$ ,    derajat reliabilitas sangat rendah

$0,20 \leq 0,40$ ,    derajat reliabilitas rendah

$0,40 \leq 0,60$ ,    derajat reliabilitas sedang

$0,60 \leq 0,80$ ,    derajat reliabilitas tinggi

$0,80 \leq 1,00$ ,    derajat reliabilitas sangat tinggi,

Mengacu dari kriteria yang dibuat oleh Guiford, pada uji reliabilitas instrumen telah didapatkan  $r_{1.1} = 0,95$ , jadi instrumen yang telah diujicobakan terbatas ini memiliki reliabilitas atau keajegan atau konsistensi yang sangat tinggi (*lihat lampiran 6*).

### **Uji Empirik Kedua**

Pada Uji kedua ini dilakukan untuk lebih mengetahui seberapa lebih dalam tingkat kevalidan dan reliabilitas butir tes, oleh karena itu pada tahap ini pengujian butir tes diberikan pada sampel yang berjumlah cukup besar disesuaikan dengan pendapat Nunnaly bahwa banyaknya sampel untuk penelitian adalah 5 sampai 10 kali lipat banyaknya butir yang hendak dikembangkan atau dianalisis (Koyan, 2011). Dalam penelitian ini jumlah butir tes yang akan dikembangkan atau dianalisis adalah 50 setelah 4 butir tes dinyatakan gugur pada uji empirik pertama, dengan jumlah sampel pada penelitian 5-10 x 50. Pada penelitian ini, angka standar akan digunakan untuk dikalikan jumlah butir, 5 x 50 = 250. Jadi jumlah sampel pada penelitian ini adalah 250 orang siswa yang



termasuk dalam populasi. Setelah dilakukan pengujian butir tes pada sampel, pengujian empirik kedua ini akan dianalisis melalui lima tahapan, meliputi: 1) uji validitas butir, 2) uji reliabilitas butir, 3) uji taraf kesukaran butir, 4) uji daya beda, dan 5) efektivitas pengecoh.

### **Uji Validitas Butir**

Proses analisis uji validitas butir pada uji empirik kedua ini sama halnya dengan proses analisis uji validitas butir pada uji empirik pertama, validitas butir tes ini pun tiap butir dihitung dengan mengkorelasikan skor butir dengan skor total yang diperoleh respon. Skor butir tes objektif berupa skala dikotomi, sedangkan skor totalnya berupa skala interval yakni jumlah skor butir. Oleh karena itu, tehnik korelasi yang digunakan untuk menghitung validitas butir tes objektif (Candiasa, 2010:96). Rumus yang akan digunakan untuk menghitung validitas tiap butir, sebagai berikut.

$$r_{pbi} = \frac{M_p - M_t}{s_t} \sqrt{\frac{p}{q}}$$

Keterangan:

- $M_p$  = Rerata skor total dari subjek yang menjawab betul butir yang dicari validitasnya.
- $M_t$  = rerata skor total
- $S_t$  = Standar deviasi skor total
- $p$  = proporsi siswa yang menjawab benar butir yang dicari validitasnya
- $q$  = proporsi siswa yang menjawab salah butir yang dicari validitasnya ( $q = 1-p$ )

Uji validitas butir tes ini dihitung dengan menggunakan korelasi point biserial melalui program *microsoft office excel*, dengan  $n= 200$  pada taraf signifikansi 0,05 didapatkan  $r\text{-tabel}=0,138$ , hasil butir yang valid dapat dilihat seperti di bawah ini.

*Tabel 4.5 Butir Tes yang Valid/Drop Pada Uji Empirik Kedua*

<b>Butir</b>	<b>r-pbi</b>	<b>Pada r-tabel 200 (0,05)</b>	<b>Ket</b>
<b>1</b>	0.351	0.138	Valid
<b>2</b>	0.330	0.138	Valid
<b>3</b>	0.047	0.138	Drop
<b>4</b>	0.478	0.138	Valid
<b>5</b>	0.220	0.138	Valid
<b>6</b>	0.412	0.138	Valid
<b>7</b>	0.083	0.138	Drop
<b>8</b>	0.420	0.138	Valid
<b>9</b>	0.571	0.138	Valid
<b>10</b>	0.355	0.138	Valid
<b>11</b>	0.373	0.138	Valid
<b>12</b>	0.210	0.138	Valid
<b>13</b>	0.056	0.138	Drop
<b>14</b>	0.033	0.138	Drop
<b>15</b>	0.096	0.138	Drop
<b>16</b>	0.325	0.138	Valid
<b>17</b>	0.004	0.138	Drop
<b>18</b>	0.365	0.138	Valid
<b>19</b>	0.041	0.138	Drop
<b>20</b>	0.314	0.138	Valid
<b>21</b>	0.215	0.138	Valid
<b>22</b>	0.301	0.138	Valid
<b>23</b>	0.323	0.138	Valid
<b>24</b>	0.331	0.138	Valid
<b>25</b>	0.501	0.138	Valid
<b>26</b>	0.504	0.138	Valid
<b>27</b>	0.355	0.138	Valid
<b>28</b>	0.202	0.138	Valid
<b>29</b>	0.560	0.138	Valid

*Tabel Lanjutan*

<b>Butir</b>	<b>r-pbi</b>	<b>Pada r-tabel 200 (0,05)</b>	<b>Ket</b>
<b>30</b>	0.455	0.138	Valid
<b>31</b>	0.139	0.138	Valid
<b>32</b>	0.214	0.138	Valid
<b>33</b>	0.411	0.138	Valid
<b>34</b>	0.077	0.138	Drop
<b>35</b>	0.211	0.138	Valid
<b>36</b>	0.254	0.138	Valid
<b>37</b>	0.316	0.138	Valid
<b>38</b>	0.417	0.138	Valid
<b>39</b>	0.078	0.138	Drop
<b>40</b>	0.493	0.138	Valid
<b>41</b>	0.425	0.138	Valid
<b>42</b>	0.365	0.138	Valid
<b>43</b>	0.539	0.138	Valid
<b>44</b>	0.255	0.138	Valid
<b>45</b>	0.451	0.138	Valid
<b>46</b>	-0.071	0.138	Drop
<b>47</b>	0.174	0.138	Valid
<b>48</b>	0.226	0.138	Valid
<b>49</b>	0.280	0.138	Valid
<b>50</b>	0.103	0.138	Drop

Hasil analisis uji validitas butir pada uji empirik II ini menunjukkan dari 50 butir tes didapatkan 39 butir tes yang valid yakni: 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 16, 18, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 35, 36, 37, 38, 40, 41, 42, 43, 44, 45, 47, 48 dan 49. Kemudian terdapat 11 butir tes yang gugur atau drop yakni; dan 7, 13, 14, 15, 17, 19, 34, 39, 46 dan 50.

### **Uji Taraf Kesukaran Butir**

Uji taraf kesukaran butir dilakukan dengan maksud untuk dapat membedakan mana butir yang masuk dalam kriteria mudah, sedang dan sukar.

Untuk menganalisis uji taraf kesukaran butir dilakukan dengan memperhitungkan banyak yang menjawab butir tersebut dengan benar (candiasa, 2010). Rumus yang

digunakan yaitu,  $I = \frac{B}{N}$

Keterangan:

*I = Indeks Kesukaran Butir*

*B = Banyaknya jumlah siswa yang menjawab butir tersebut dengan benar*

*N = Jumlah siswa yang mengikuti tes*

*Tabel 4.6 Taraf Kesukaran Butir*

<b>Butir</b>	<b>B</b>	<b>N</b>	<b>I</b>	<b>Keterangan</b>
<b>1</b>	120	250	0,48	Sedang
<b>2</b>	186	250	0,74	Mudah
<b>3</b>	119	250	0,48	Sedang
<b>4</b>	91	250	0,36	Sedang
<b>5</b>	130	250	0,52	Sedang
<b>6</b>	146	250	0,58	Sedang
<b>7</b>	99	250	0,40	Sedang
<b>8</b>	101	250	0,40	Sedang
<b>9</b>	68	250	0,27	Sukar
<b>10</b>	169	250	0,68	Sedang
<b>11</b>	130	250	0,52	Sedang
<b>12</b>	60	250	0,24	Sukar
<b>13</b>	35	250	0,14	Sukar
<b>14</b>	45	250	0,18	Sukar
<b>15</b>	80	250	0,32	Sedang
<b>16</b>	151	250	0,60	Sedang
<b>17</b>	60	250	0,24	Sukar
<b>18</b>	161	250	0,47	Sedang
<b>19</b>	45	250	0,18	Sukar
<b>20</b>	104	250	0,42	Sedang
<b>21</b>	52	250	0,21	Sukar
<b>22</b>	108	250	0,43	Sedang
<b>23</b>	59	250	0,24	Sukar
<b>24</b>	173	250	0,69	Sedang

*Tabel Lanjutan*

<b>Butir</b>	<b>B</b>	<b>N</b>	<b>I</b>	<b>Keterangan</b>
25	126	250	0,50	Sedang
26	135	250	0,54	Sedang
27	131	250	0,52	Sedang
28	140	250	0,56	Sedang
29	94	250	0,38	Sedang
30	110	250	0,44	Sedang
31	158	250	0,63	Sedang
32	65	250	0,26	Sukar
33	121	250	0,48	Sedang
34	109	250	0,44	Sedang
35	58	250	0,23	Sukar
36	39	250	0,56	Sedang
37	75	250	0,30	Sukar
38	82	250	0,33	Sedang
39	40	250	0,16	Sukar
40	88	250	0,35	Sedang
41	94	250	0,38	Sedang
42	66	250	0,26	Sukar
43	78	250	0,31	Sedang
44	87	250	0,34	Sedang
45	93	250	0,37	Sedang
46	47	250	0,19	Sukar
47	71	250	0,28	Sukar
48	49	250	0,20	Sukar
49	94	250	0,38	Sedang
50	86	250	0,34	Sedang

Kriteria yang digunakan untuk dapat mengidentifikasi kelas butir adalah sebagai berikut:

Butir dengan P 0,00 – 0,30 tergolong sukar

Butir dengan P 0,31 – 0,70 tergolong sedang

Butir dengan P 0,71 – 1,00 tergolong mudah

Mengacu pada kriteria yang telah ditentukan dilihat pada tabel di atas terdapat 1 butir tes yang kategorinya mudah (butir soal no. 2), 33 butir tes kategori sedang (butir 1, 3, 4, 5, 6, 7, 8, 10, 11, 15, 16, 18, 20, 22, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 36, 38, 40, 41, 43, 44, 45, 49 & 50) dan 16 butir kategori sukar (butir 9, 12, 13, 14, 17, 19, 21, 23, 32, 35, 37, 39, 42, 46, 47 & 48),

### **Uji Daya Beda**

Menurut Candiasa (2010:85) analisis daya beda butir merupakan pengkajian butir-butir tes yang dimaksudkan untuk mengetahui kesanggupan butir tes untuk membedakan peserta tes yang tergolong mampu dengan peserta tes yang tergolong tidak mampu. Butir tes yang baik harus mampu membedakan peserta tes yang mampu dan peserta tes yang tidak mampu. Dengan kata lain, butir tes yang baik seharusnya dijawab benar oleh peserta tes yang dianggap mampu dan dijawab salah oleh peserta tes yang tidak mampu.

Sebelum dicari indeks kesukaran butir, data dibagi dalam 2 kelompok yaitu kelompok atas dan kelompok bawah. butir tes yang diambil untuk di bagi dalam kedua kelompok hanya butir tes yang valid. Kelompok atas dan bawah diperoleh dengan merangking peserta berdasarkan sekor total yang diperoleh, kemudian 27% dari siswa yang memperoleh sekor tinggi diambil sebagai kelompok tinggi responden dan 27% dari siswa yang memperoleh sekor rendah diambil sebagai kelompok rendah. Responden (peserta tes) sebanyak 250 orang, sehingga kelompok tinggi dan rendah masing-masing di ambil berdasarkan  $27\% \times 250 = 68$  responden.

Indeks daya beda butir dinyatakan dengan  $d$  dan dihitung dengan

menggunakan rumus,  $d = \frac{U-L}{N}$

Keterangan

$d$  = indeks daya beda

$U$  = banyak kelompok atas yang menjawab butir dengan benar

$L$  = banyak kelompok bawah yang menjawab butir dengan benar

$N$  = banyak peserta tes

Tabel 4.7 Hasil Uji Daya Beda

<b>NOMOR BUTIR</b>	<b>U</b>	<b>L</b>	<b>N</b>	<b>D</b>	<b>Keterangan</b>
1	49	17	250	0,13	Lemah
2	64	36	250	0,11	Lemah
4	50	11	250	0,16	Lemah
5	40	26	250	0,07	Lemah
6	58	23	250	0,14	Lemah
8	49	16	250	0,13	Lemah
9	44	5	250	0,16	Lemah
10	61	33	250	0,11	Lemah
11	54	17	250	0,15	Lemah
12	24	13	250	0,04	Lemah
16	59	26	250	0,13	Lemah
18	50	14	250	0,14	Lemah
20	46	19	250	0,11	Lemah
21	13	7	250	0,02	Lemah
22	43	17	250	0,10	Lemah
23	29	5	250	0,10	Lemah
24	62	31	250	0,12	Lemah
25	67	11	250	0,22	Cukup
26	64	11	250	0,21	Cukup
27	52	24	250	0,11	Lemah
28	46	28	250	0,07	Lemah
29	59	6	250	0,21	Cukup
30	53	13	250	0,17	Lemah
31	51	37	250	0,06	Lemah
32	25	13	250	0,05	Lemah

*Tabel Lanjutan*

<b>NOMOR BUTIR</b>	<b>U</b>	<b>L</b>	<b>N</b>	<b>D</b>	<b>Keterangan</b>
<b>33</b>	52	15	250	0,15	Lemah
<b>35</b>	25	8	250	0,07	Lemah
<b>36</b>	16	5	250	0,04	Lemah
<b>37</b>	31	9	250	0,10	Lemah
<b>38</b>	43	8	250	0,14	Lemah
<b>40</b>	47	9	250	0,15	Lemah
<b>41</b>	48	8	250	0,16	Lemah
<b>42</b>	33	9	250	0,10	Lemah
<b>43</b>	53	6	250	0,19	Lemah
<b>44</b>	35	22	250	0,05	Lemah
<b>45</b>	47	10	250	0,15	Lemah
<b>47</b>	21	8	250	0,05	Lemah
<b>48</b>	22	6	250	0,06	Lemah
<b>49</b>	44	18	250	0,10	Lemah

Untuk dapat mengidentifikasi uji daya beda digunakan kriteria yang dikemukakan oleh Sudijono (2011:189).

Kurang dari 0,20       $\longleftrightarrow$       Butir memiliki daya beda yang lemah sekali

0,20 – 0,40           $\longleftrightarrow$       Butir memiliki daya beda yang cukup

0,40 – 0,70           $\longleftrightarrow$       Butir memiliki daya beda yang baik

0,70 – 1,00           $\longleftrightarrow$       Butir memiliki daya beda yang baik sekali

Berdasarkan kriteria yang ditetapkan Sudijono tersebut, dapat ditarik kesimpulan sebagai berikut. Terdapat 36 butir yang daya bedanya kurang yakni



butir; 1, 2, 4, 5, 6, 8, 9, 10, 11, 12, 16, 18, 20, 21, 22, 23, 24, 27, 28, 30, 31, 32, 33, 35, 36, 38, 40, 41, 42, 43, 44, 45, 47, 48, 49. Dan terdapat 3 butir yang berkategori cukup yakni butir; 25, 26 dan 29 (*lihat lampiran 10*).

### Uji Efektivitas Pengecoh

Uji efektifitas pengecoh seperti yang dikemukakan oleh Candiasa (2010), analisis efektifitas pengecoh (*distractor*) atau analisis pola jawaban dilakukan dengan menghitung peserta tes yang memilih tiap alternatif jawaban pada masing-masing butir. Kriteria pengecoh yang baik adalah apabila pengecoh tersebut dipilih oleh paling sedikit 5% dari peserta tes.

Pada penelitian ini adapun banyak peserta tes 250 orang, 5% dari 250 orang responden adalah 12,50 dibulatkan menjadi 13 orang. Jadi tiap butir pengecoh minimal dipilih oleh 13 orang peserta tes. Berdasarkan jawaban siswa didapatkan hasil seperti berikut ini.

*Tabel 4.8 Hasil Uji Efektivitas Pengecoh*

NOMOR BUTIR YANG VALID	PILIHAN JAWABAN				KUNCI	$\frac{5}{100} \times 250$	NOMOR BUTIR YANG VALID
	A	B	C	D			
<b>B1</b>	59	112	32	47	B	12,5 = 13	Semua pengecoh baik
<b>B2</b>	186	13	17	34	A	12,5 = 13	Semua pengecoh baik
<b>B4</b>	59	94	34	63	B	12,5 = 13	Semua pengecoh baik
<b>B5</b>	24	74	131	21	C	12,5 = 13	Semua pengecoh baik
<b>B6</b>	148	38	30	34	A	12,5 = 13	Semua pengecoh baik
<b>B8</b>	100	55	51	44	A	12,5 = 13	Semua pengecoh baik
<b>B9</b>	76	48	59	67	D	12,5 = 13	Semua pengecoh baik
<b>B10</b>	26	35	173	16	C	12,5 = 13	Semua pengecoh baik
<b>B11</b>	49	127	25	49	B	12,5 = 13	Semua pengecoh baik
<b>B12</b>	44	83	63	60	D	12,5 = 13	Semua pengecoh baik

*Tabel Lanjutan*

NOMOR BUTIR YANG VALID	PILIHAN JAWABAN				KUNCI	$\frac{5}{100} \times 250$	NOMOR BUTIR YANG VALID
	A	B	C	D			
B16	149	41	38	22	A	12,5 = 13	Semua pengecoh baik
B18	114	49	40	47	A	12,5 = 13	Semua pengecoh baik
B20	73	30	43	104	D	12,5 = 13	Semua pengecoh baik
B21	84	49	46	71	B	12,5 = 13	Semua pengecoh baik
B22	74	110	35	31	B	12,5 = 13	Semua pengecoh baik
B23	66	71	57	56	D	12,5 = 13	Semua pengecoh baik
B24	28	19	173	30	C	12,5 = 13	Semua pengecoh baik
B25	47	38	36	129	D	12,5 = 13	Semua pengecoh baik
B26	48	136	46	20	B	12,5 = 13	Semua pengecoh baik
B27	29	47	130	44	C	12,5 = 13	Semua pengecoh baik
B28	31	30	49	140	D	12,5 = 13	Semua pengecoh baik
B29	35	87	91	37	B	12,5 = 13	Semua pengecoh baik
B30	33	61	109	47	C	12,5 = 13	Semua pengecoh baik
B31	161	32	25	32	A	12,5 = 13	Semua pengecoh baik
B32	65	58	66	61	C	12,5 = 13	Semua pengecoh baik
B33	47	119	54	30	B	12,5 = 13	Semua pengecoh baik
B35	64	72	72	42	A	12,5 = 13	Semua pengecoh baik
B36	51	105	42	52	C	12,5 = 13	Semua pengecoh baik
B37	68	44	74	64	C	12,5 = 13	Semua pengecoh baik
B38	78	86	49	37	A	12,5 = 13	Semua pengecoh baik
B40	90	73	40	47	A	12,5 = 13	Semua pengecoh baik
B41	44	88	60	58	B	12,5 = 13	Semua pengecoh baik
B42	95	34	48	73	D	12,5 = 13	Semua pengecoh baik
B43	64	63	44	79	D	12,5 = 13	Semua pengecoh baik
B44	37	82	92	39	C	12,5 = 13	Semua pengecoh baik
B45	47	38	70	95	D	12,5 = 13	Semua pengecoh baik
B47	68	77	68	37	A	12,5 = 13	Semua pengecoh baik
B48	82	53	61	54	B	12,5 = 13	Semua pengecoh baik
B49	68	49	96	37	C	12,5 = 13	Semua pengecoh baik

Jika dilihat pada tabel diatas dapat dipastikan semua pengecoh yang ada pada butir berkategori sangat baik karna tiap pengecoh dipilih oleh lebih dari 5% peserta tes.

## Uji Reliabilitas Tes

Uji reliabilitas butir pada uji empirik kedua ini sama halnya dengan proses analisis pada uji reliabilitas butir pada uji empirik pertama yaitu dilakukan setelah didapatkan butir-butir yang dinilai valid, dari 50 butir tes setelah di analisis validitasnya didapatkan 39 butir tes yang valid.

Untuk menganalisis tingkat reliabilitas butir digunakan formula KR-20

Formula KR-20 adalah seperti berikut:  $r_{1.1} = \left( \frac{k}{k-1} \right) \left( \frac{SD_t^2 - \sum pq}{SD^2} \right)$

Perhitungan secara manual dengan mengambil beberapa harga tertentu dalam pertungan *microsoft office excel*, diketahui  $n = 250$ ,  $k = 39$ ,  $\sum pq = 11,6896$ , jadi

$$SD = \sqrt{\frac{n \sum X^2 - (\sum X)^2}{n(n-1)}} = \sqrt{\frac{250 \sum 75664 - (4018)^2}{250(250-1)}}$$

$$SD = \sqrt{\frac{18916000 - 16144324}{6000}} = \sqrt{\frac{2771676}{6000}}$$

$$SD = \sqrt{46,1946} = 6,797$$

$$SD_t^2 = (6,797)^2$$

$$SD_t^2 = 46,1946$$

$$r_{1.1} = \left( \frac{k}{k-1} \right) \left( \frac{SD_t^2 - \sum pq}{SD_t^2} \right) = \left( \frac{39}{39-1} \right) \left( \frac{46,1946 - 8,606272}{46,1946} \right)$$

$$r_{1.1} = 0,835108 = 0,84$$

Keterangan:

$r_{1.1}$	= koefisien reliabilitas tes
$n$	= banyaknya testee
$k$	= banyak butir tes
$SDt$	= varian total tes
$p$	= proporsi testee yang menjawab benar
$q$	= proporsi testee yang menjawab salah
$pq$	= $pxq$

Kriteria derajat reliabilitas tes atau instrumen evaluasi dapat digunakan kriteria yang dibuat oleh Guiford (1951) (Candi, 2010:107), seperti berikut ini:

$r_{1.1} \leq 0,20$ ,	derajat reliabilitas sangat rendah
$0,20 \leq 0,40$ ,	derajat reliabilitas rendah
$0,40 \leq 0,60$ ,	derajat reliabilitas sedang
$0,60 \leq 0,80$ ,	derajat reliabilitas tinggi
$0,80 \leq 1,00$ ,	derajat reliabilitas sangat tinggi,

Mengacu dari kriteria yang dibuat oleh Guiford, pada uji reliabilitas instrumen telah didapatkan  $r_{1.1} = 0,84$ , jadi instrumen yang telah diujicobakan terbatas ini memiliki reliabilitas atau keajegan atau konsistensi yang sangat tinggi (lihat lampiran. 8 ).

Dari keseluruhan hasil analisis di atas, dalam penelitian pengembangan ini dirumuskan simpulan sebagai berikut, terdapat 11 butir yang harus dipertimbangkan untuk digunakan yakni butir 2, 9, 12, 21, 32, 35, 37, 42, 47 dan 48. Terdapat 25 butir yang memadai untuk digunakan yakni butir 1, 4, 5, 6, 8, 10, 11, 16, 18, 20, 22, 24, 27, 28, 30, 31, 33, 36, 38, 40, 41, 43, 44, 45, 49 dan terdapat 3 butir yang sangat memadai untuk digunakan yakni butir 25, 26, dan 29. Penarikan kesimpulan dapat dilihat pada tabel di bawah ini.

*Tabel 4.9 Kesimpulan Hasil Analisis*

<b>Nomor Butir</b>	<b>Validitas Butir</b>	<b>Tingkat Kesukaran Butir</b>	<b>Daya Beda Butir</b>	<b>Tingkat efektivitas Pengecoh</b>	<b>Keterangan</b>
<b>B. 1</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 2</b>	Valid	Mudah	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 4</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 5</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 6</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 8</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 9</b>	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 10</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 11</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 12</b>	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 16</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 18</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 20</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 21</b>	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 22</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 23</b>	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 24</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 25</b>	Valid	Sedang	Cukup	Semua pengecoh baik	Sangat memadai untuk digunakan
<b>B. 26</b>	Valid	Sedang	Cukup	Semua pengecoh baik	Sangat memadai untuk digunakan
<b>B. 27</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 28</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 29</b>	Valid	Sedang	Cukup	Semua pengecoh baik	Sangat memadai untuk digunakan
<b>B. 30</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 31</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 32</b>	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 33</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 35</b>	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 36</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 37</b>	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 38</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 40</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 41</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 42</b>	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
<b>B. 43</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 44</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan
<b>B. 45</b>	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan

*Tabel Lanjutan*

Nomor Butir	Validitas Butir	Tingkat Kesukaran Butir	Daya Beda Butir	Tingkat efektivitas Pengecoh	Keterangan
B. 47	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
B. 48	Valid	Sukar	Lemah	Semua pengecoh baik	Dipertimbangkan untuk digunakan
B. 49	Valid	Sedang	Lemah	Semua pengecoh baik	Memadai untuk digunakan

Dantes (2012) mengusulkan kriteria untuk penarikan kesimpulan dalam menentukan layak tidaknya butir tes untuk digunakan, seperti berikut ini.

*Tabel 4.10 Kriteria Penarikan Kesimpulan*

Kriteria butir	Deskripsi
Tidak digunakan	Jika butir tidak valid, tingkat kesukaran butir sukar atau mudah, daya beda lemah sekali, dan pengecoh tidak efektif dan atau dengan kata lain, hanya salah satu dari empat faktor dalam validasi tes yang terpenuhi dan atau empat faktor dalam validasi tes tidak memenuhi syarat sama sekali.
Dipertimbangkan untuk digunakan	Jika dua kriteria dari empat faktor dalam validasi tes memenuhi syarat
Memadai untuk digunakan	jika tiga dari faktor dari empat faktor dalam validasi tes memenuhi syarat
Sangat memadai untuk digunakan	jika semua faktor dalam validasi tes memenuhi syarat

#### **IV. PENUTUP**

##### **Simpulan**

Penelitian pengembangan tes objektif pilihan ganda dilakukan dengan melalui beberapa tahapan, diupayakan untuk menjawab rumusan masalah serta tujuan penelitian dapat disimpulkan seperti berikut ini.

- 1) Kisi-kisi tes objektif pilihan ganda yang berdasarkan taksonomi Anderson dan Krathwohl memungkinkan pembuatan butir tes yang bervariasi untuk setiap jenis proses kognitif. Apabila dalam taksonomi bloom, hanya dikenal jenjang C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, dan seterusnya. Dalam taksonomi yang baru tiap jenjang menjadi empat kali lipat sebab ada empat macam pengetahuan. Guru yang membuat soal jenjang C<sub>1</sub>, kini bisa memvariasikan soalnya, menjadi C<sub>1</sub>-faktual, C<sub>1</sub>-konseptual, C<sub>1</sub>-prosedural, C<sub>1</sub>- metakognitif, dan seterusnya.
- 2) Validasi tes berbasis taksonomi Anderson dan Krathwohl yang meliputi:
  - a) Validitas isi sudah memenuhi syarat dengan melakukan uji judges, judges menyatakan instrumen relevan untuk diujicobakan.
  - b) Uji empirik instrumen pun dinyatakan sudah memenuhi syarat karena semua ketentuan yang harus ditempuh seperti; uji validitas butir, uji tingkat kesukaran butir, daya beda dan uji efektivitas pengecoh, semuanya telah dilakukan seperti yang telah ditentukan.
- 3) Reliabilitas tes objektif pilihan ganda berdasarkan taksonomi Anderson dan Krathwohl dinyatakan sudah memenuhi syarat dengan deskripsi berikut ini, dari 50 butir tes yang diuji empirik tahap II pada 250 sampel hasil analisis menunjukkan 39 butir tes yang valid, dengan tingkat reliabilitas = 0,84 itu berarti butir tes yang telah diujikan ini masih memiliki reliabilitas atau keajegan atau konsistensi yang sangat tinggi.

Pada penelitian ini, dari keseluruhan hasil analisis butir didapatkan ada 3 butir tes yang didapatkan sangat memadai untuk digunakan yakni butir; 25, 26 29, dan 25 butir tes yang dianggap memadai untuk digunakan yaitu butir : 1, 4, 5,

6, 8, 10, 11, 16, 18, 20, 22, 24, 27, 28, 30, 31, 33, 36, 38, 40, 41, 43, 44, 45, dan 49 serta 11 butir tes yang harus dipertimbangkan untuk digunakan yakni butir; 2, 9, 12, 21, 23, 32, 35, 37, 42, 47 dan 48.

## **Saran**

### **Untuk Guru**

Di sarankan kepada guru, dalam pembuatan butir tes yang akan diberikan kepada siswa diharapkan benar-benar tes telah teruji validitas dan reliabilitasnya, bila perlu dalam membuat tes objektif pilihan ganda dapat melakukan tahapan yang ada pada penelitian ini.

### **Untuk Sekolah**

Untuk menambah kompetensi guru yang ada di masing-masing sekolah, pihak sekolah diharapkan dapat memfasilitasi guru-gurunya seperti menghadirkan para pakar yang berkompeten, workshop mandiri, secara khusus dalam mempelajari tentang taksonomi-taksonomi diluar taksonomi bloom.

### **Kepada Peneliti Lain**

Jika dilihat dari proporsi pembagian, 20% mudah, 60%sedang dan 20% sukar, instrumen ini belum dapat mengikuti hal itu. Untuk memenuhi persentase tersebut diharapkan bagi peneliti berikut untuk melakukan pengembangan lebih lanjut.

Penelitian pengembangan tes objektif pilihan ganda ini hal yang diukur hanya terbatas pada salah satu aspek kognitif yang ada pada mata pelajaran bahasa inggris, yaitu hanya mengukur membaca pemahaman siswa kelas VIII. Diharapkan kepada peneliti lain untuk mengembangkan penelitian ini dapat mengukur semua aspek yang ada pada setiap mata pelajaran secara kolektif.



## DAFTAR PUSTAKA

- Anderson, O. W dan David R. Krathwohln (eds.). 1976. *A Taxonomy for learning, Teaching and Assessing (A Revision of Bloom's Taxonomy of Educational Objectives)*. New York: San Fransisco Boston.
- Annisa. N. 2009. The Correlation Between Eight Grade Students' Reading Comprehension Achievement and Their Performance in Translating Descriptive Text at SMP Negeri 1 Palembang. *Tesis*. Program Pascasarja Universitas Sriwijaya.
- Arikunto. Suharsimi. 2000. *Manajemen Penelitian*. Jakarta: Rineka Cipta.
- Ayu, I Gusti. 2006. Pengembangan Instrumen Asesmen Portofolio Untuk Mengukur Kemampuan Menulis Siswa Pada Mata Pelajaran Bahasa Inggris (Studi Pada Siswa Kelas VIII SMP Negeri di Denpasar). *Tesis*. (Tidak Diterbitkan). Singaraja: Program Pascasarjana Universitas Pendidikan Ganesha Singaraja.
- Bungin, Burhan. 2010. *Metodologi Penelitian Kuantitatif dan Kualitatif (Komunikasi, Ekonomi, dan Kebijakan Publik Serta Ilmu-ilmu Sosial Lainnya)*. Jakarta: Kencana.
- Candiasa, I Made. 2010, *Pengujian Instrumen Penelitian Disertai Aplikasi ITEMAN dan BIGSTEPS*. Singaraja: Unit Penerbitan Undiksha.
- Dinas Pendidikan, Pemuda dan Olahraga. 2012. Data Tentang Keadaan Sekolah-Sekolah Di Kabupaten Poso. Poso: Dinas Pendidikan, Pemuda dan Olahraga
- Djiwandono, P.I. 2001. *Strategi Membaca Bahasa Inggris*. Jakarta: PT Gramedia Pustaka Utama.
- Esty, Dkk, 2009. Metode Pembelajaran Keterampilan Membaca Pemahaman Untuk Guru-Guru SMP Se-Kabupaten Ngawi Sebagai Upaya Meningkatkan Kompetensi Pedagogik Guru Bahasa Indonesia. *Usulan PPM Kompetitif*. Disampaikan pada materi *Treaning of Treanir*, tanggal 29 September 2009 di Ngawi.

- Glass, A. L. et al. 1998, *Cognition*. Philippines Copyright: by Addison-Wesley Publishing Company, inc.
- Hafni. 1981. Pemilihan dan Pengembangan Bahan Pengajaran Membaca. *Makalah*. Disajikan pada Penataran dan Lokakarya Tahap II Proyek Pengembangan Pendidikan Guru (P3G) Depdikbud. Jakarta, 29 Maret 1981.
- Harjasujana, A. S. Dan Yeti. M. 1997. *Membaca 2*. Jakarta: Depdikbud Ditjen Dikdasmen Bagian Proyek Penataran Guru SLTP Setara D-III.
- Haryadi, dan Zamsani. 1997. *Peningkatan Keterampilan Berbahasa Indonesia*. Jakarta: Depdikbud Ditjen Dikti Bagian Proyek Pengembangan PGSD.
- Hillerich, R. L. 1983. *The Principle's Guide to Improving Reading Instruction*. Massachussetts: Allyn and Bacon, Inc
- Koyan, I Wayan. 2011. Langkah-langkah Mengkontruksi Tes dan Non Tes. *Buku Ajar*. (tidak diterbitkan) Singaraja: Undiksha
- ..... 2011, *Asesmen Dalam Pendidikan*. Singaraja: Unit Penerbitan Undiksha.
- Kridalaksana, H. 1993. *Kamus Linguistik*. Jakarta: Gramedia Pustaka Utama.
- Sudijono, A. 2011. *Pengantar Evaluasi Pendidikan*. Jakarta: Persada Rajawali.
- Suyanu. 2006. Profil Kemampuan Membaca Pemahaman Siswa Kelas II SMA Negeri di Lombok Barat, Nusa Tenggara Barat Ditinjau Dari Lokasi dan Jurusan. *Tesis*. (Tidak Diterbitkan). Singaraja: Program Pascasarjana Universitas Pendidikan Ganesha.
- Rajeg, I Nyoman. 2006. Pengembangan Alat Pengukur Sikap Sosial dalam konteks pembelajaran Pendidikan Kewarganegaraan dan pendidikan Sosial Siswa SD kelas VI di Kota Denpasar. *Tesis*. (Tidak Diterbitkan). Singaraja: Program Pascasarjana Universitas Pendidikan Ganesha.
- Soedarso, 2002. *Speed Reading*. Jakarta: Gramedia Pustaka Utama.

- Sudiana, I Nyoman. 1996. Pengajaran Membaca Di Sekolah Menengah Pertama. *Disertasi*. (Tidak Diterbitkan). Malang: Program Pascasarjana IKIP Malang.
- Suastika, I Nyoman. 2009. Pengembangan Instrumen Asesmen Otentik Dalam Pendidikan Kewarganegaraan Yang Bermuatan Multikultural Pada SMP Negeri Di Kota Singaraja. Tesis. (Tidak Diterbitkan). Singaraja: Program Pascasarjana Universitas Pendidikan Ganesha.
- Sumaratih. 2006. Tentang Pengaruh Asesmen Portofolio Terhadap Kemampuan Menulis Teks Bahasa Inggris (Eksperimen Pada Siswa SMA Negeri 1 Singaraja Tahun Pelajaran 2005/2006). *Tesis*. (Tidak Diterbitkan). Singaraja: Program Pascasarjana Universitas Pendidikan Ganesha.
- Tarigan, H.G. 1979. *Membaca Sebagai Suatu Keterampilan Berbahasa*. Bandung: Angkasa.
- Trissyana. 2011. The Development Of Performance Of Writting Skill Of Viii Grade Of Junior High Scholl In Singaraja. *Tesis*. (Tidak Diterbitkan). Singaraja: Program Pascasarjana Universitas Pendidikan Ganesha.
- Alexander J. E. et. al. 1998. (eds.) *Teaching Reading*. USA: Scot, Foresman and Company.
- Widodo, A. 2005. Taksonomi Tujuan Pembelajaran. *Didaktis*. Bandung: Buletin Puspendik Univeristas Pendidikan Bandung.
- ..... 2006. Revisi Taksonomi Bloom dan Pengembangan Butir Soal. *Didaktis*. Bandung: Buletin Bandung Univeristas Pendidikan Bandung.