

BAB II

LANDASAN TEORI

A. Deskripsi Teori

1. Instrumen Tes

a. Pengertian tes

Kata “tes” berasal dari bahasa Perancis kuno yaitu “testum” yang berarti piring untuk menyisihkan logam-logam mulia.¹ maksudnya yaitu dengan menggunakan alat berupa piring akan diperoleh jenis-jenis logam mulia yang nilainya sangat tinggi.

Beberapa istilah yang berkaitan dengan tes yaitu *test*, *testing*, *tester*, dan *testee*. *Test* adalah alat atau prosedur yang digunakan untuk mengetahui atau mengukur sesuatu dalam suasana, dengan cara dan aturan-aturan yang sudah ditentukan.² *Testing* merupakan waktu pada saat tes dilaksanakan. *Tester* yaitu orang yang melaksanakan pengambilan tes atau pembuat tes atau eksperimentor. *Testee* yaitu responden yang mengerjakan tes. *Testee* tersebut yang dinilai atau diukur baik mengenai kemampuan, minat, bakat, pencapaian prestasi belajar dan sebagainya.

¹ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 66.

² Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, (Jakarta: Bumi Aksara, 2011), hlm. 53.

Menurut Anas Sudijono dalam buku *pengantar evaluasi pendidikan*,

tes adalah cara yang dapat dipergunakan atau prosedur yang perlu ditempuh dalam rangka pengukuran dan penilaian dibidang pendidikan yang berbentuk pemberian tugas atau serangkaian tugas baik berupa pertanyaan-pertanyaan yang harus dijawab atau perintah-perintah yang harus dikerjakan oleh *testee*, sehingga atas dasar data yang diperoleh dari hasil pengukuran tersebut dapat dihasilkan nilai yang melambangkan tingkah laku atau prestasi *testee*.³

Menurut S. Eko Widoyoko dalam buku *Penilaian hasil Pembelajaran di Sekolah*,

Tes yaitu sejumlah pertanyaan yang membutuhkan jawaban atau sejumlah pernyataan yang harus diberi tanggapan atau respons dengan tujuan mengukur tingkat kemampuan seseorang atau mengungkap aspek tertentu dari orang yang dikenai tes (*testee*).⁴

Menurut Sumarna Surapranata dalam buku *Panduan Penulisan Tes Tertulis Implementasi Kurikulum 2004*,

Tes yaitu sehimpunan pertanyaan yang harus dijawab atau pernyataan-pernyataan yang harus dipilih, ditanggapi, atau tugas-tugas yang harus dilakukan oleh orang yang dites (*testee*) dengan

³ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 67.

⁴ S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, (Yogyakarta: Pustaka Pelajar, 2014), hlm. 50.

tujuan untuk mengukur suatu aspek tertentu dari *testee*.⁵

Menurut Purwanto dalam buku *Evaluasi Hasil Belajar*,

Tes merupakan instrumen alat ukur untuk pengumpulan data dimana dalam memberikan respons atas pertanyaan dalam instrumen, peserta didorong untuk menunjukkan penampilan maksimalnya. Penampilan maksimum yang ditunjukkan memberikan kesimpulan mengenai kemampuan atau penguasaan yang dimiliki.⁶

Menurut Wiliam Wiersma Stephen dalam buku *Educational Measurement and Testing* mengartikan “*the test is the stimulus to which the response is made*, artinya tes adalah suatu rangsangan yang membuat orang untuk merespon (menanggapi)”.⁷

Berdasarkan beberapa pengertian tes diatas dapat disimpulkan bahwa tes adalah alat pengukur pengumpul data berupa sejumlah pertanyaan atau perintah yang memerlukan tanggapan dari *testee* untuk mengukur tingkat kemampuan, prestasi dan penguasaan yang dimiliki oleh *testee*.

⁵ Sumarna Surapranata, *Panduan Penulisan Tes Tertulis Implementasi Kurikulum 2004*, (Bandung: PT Remaja Rosdakarya, 2007), hlm. 19.

⁶ Purwanto, *Evaluasi Hasil Belajar*, (Yogyakarta: Pustaka Pelajar, 2011), hlm. 63.

⁷ Wiliam Wiersma Stephen G. Jurs, *Educational Measurement and Testing*, (United States: A Division of Simon & Schuster, 1990), hlm. 9.

b. Macam-Macam Tes

Tes sebagai alat ukur dibedakan menjadi beberapa jenis tergantung pada segi atau alasan perbedaan penggolongan tersebut.

1) Ditinjau dari segi kegunaannya, dibagi menjadi:

a) Tes Diagnostik

Tes diagnostik yaitu “tes yang digunakan untuk mengetahui kelemahan-kelemahan siswa”.⁸ Berdasarkan kelemahan-kelemahan tersebut dapat dilakukan pemberian perlakuan yang tepat.

b) Tes Formatif

Tes formatif merupakan “tes akhir atau post test yang diberikan pada akhir setiap program”.⁹ Evaluasi formatif digunakan untuk mengetahui sejauh mana kemampuan siswa setelah mengikuti suatu program tertentu. Umumnya, tes formatif disamakan dengan ulangan harian.

c) Tes Sumatif

Evaluasi sumatif atau “tes sumatif merupakan tes yang dilaksanakan setelah berakhirnya pemberian sekelompok program atau

⁸ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, (Jakarta: Bumi Aksara, 2011), hlm. 34.

⁹ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, hlm. 36.

sebuah program yang lebih besar”.¹⁰ Tes sumatif dapat disamakan dengan ulangan umum yang biasanya dilaksanakan pada tiap akhir caturwulan atau akhir semester.

2) Berdasarkan Pelaksanaan

a) *Paper Based Test* (PBT)

“*Paper based test* atau tes tertulis adalah bentuk tes yang dalam pelaksanaannya menggunakan kertas dan tulisan sebagai alat bantu, baik untuk soal tes maupun jawaban tes”.¹¹

b) *Oral Based Test* (OBT)

“*Oral based test* atau tes lisan merupakan bentuk tes yang pelaksanaannya dilakukan secara langsung”.¹² Salah satu bentuk tes lisan yaitu wawancara atau tatap muka secara langsung antara penguji dengan orang yang diuji.

c) *Computer Based Test* (CBT)

“*Computer based test* merupakan tes yang dalam pelaksanaannya menggunakan alat

¹⁰ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, (Jakarta: Bumi Aksara, 2011), hlm. 38.

¹¹ S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, (Yogyakarta: Pustaka Pelajar, 2014), hlm.51.

¹² S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, hlm.53.

bantu computer”.¹³ Hal yang membedakan dengan tes tertulis maupun lisan terletak pada teknik penyampaian butir soal yaitu naskah soal maupun lembar jawaban menggunakan komputer. Sistem skoring atau koreksi langsung dilakukan oleh komputer.

- 3) Berdasarkan Sistem Penskoran
 - a) Tes Objektif

Tes objektif memiliki arti siapa saja yang memeriksa lembar jawaban tes akan menghasilkan skor yang sama. Skor tes ditentukan oleh jawaban yang diberikan oleh peserta tes. “Tes objektif adalah tes yang penskorannya bersifat objektif, yaitu hanya dipengaruhi oleh objek jawaban atau respons yang diberikan oleh peserta tes”.¹⁴ Secara umum ada tiga tipe tes objektif, yaitu: benar salah (*true false*), menjodohkan (*matching*), dan pilihan ganda (*multiple choice*).

- (1) Benar salah (*true false*)

“Benar salah adalah tes yang butir soalnya terdiri dari pernyataan yang disertai dengan

¹³ S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, (Yogyakarta: Pustaka Pelajar, 2014), hlm.54.

¹⁴ S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, hlm.55.

alternatif jawaban yaitu jawaban atau pernyataan yang benar dan yang salah”.¹⁵

(2) Menjodohkan (*matching*)

Tipe tes menjodohkan yaitu

Butir soal ditulis dalam dua kolom atau kelompok. Kelompok pertama disebelah kiri adalah pertanyaan / pernyataan atau premis. Kelompok kedua di sebelah kanan adalah kelompok jawaban.¹⁶

Tugas peserta tes adalah mencari dan menjodohkan jawaban-jawaban, sehingga sesuai dengan pernyataan.

(3) Pilihan ganda (*multiple choice*)

“Tes pilihan ganda adalah tes dimana setiap butir soal memiliki jumlah alternatif jawaban lebih dari dua. Jumlah alternatif jawaban berkisar antara 3 (tiga) atau 5 (lima)”. Berdasarkan beberapa alternatif jawaban yang ditampilkan *testee* hanya diperkenankan memilih satu jawaban.

b) Tes Subjektif

“Tes subjektif merupakan tes yang penskorannya dipengaruhi oleh jawaban peserta

¹⁵ S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, (Yogyakarta: Pustaka Pelajar, 2014), hlm.94.

¹⁶ S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, , hlm.97.

tes dan pemberi skor”.¹⁷ Jawaban yang sama dapat memiliki skor yang berbeda oleh pemberi skor yang berlainan. Ciri-ciri tes subjektif yaitu didahului dengan kata-kata seperti: uraikan, jelaskan, bandingkan, mengapa, bagaimana, simpulkan, dan sebagainya.

- 4) Berdasarkan aspek psikis yang ingin diungkap, dibagi menjadi lima yaitu tes intelegensi, tes kemampuan, tes sikap, tes kepribadian, dan tes hasil belajar.¹⁸

Adapun penjelasannya sebagai berikut:

- a) Tes Intelegensi (*intellegency test*)

“Tes Intelegensi yaitu tes yang dilaksanakan dengan tujuan untuk mengungkap atau mengetahui tingkat kecerdasan seseorang”.¹⁹

- b) Tes Kemampuan (*aptitude test*)

“Tes kemampuan yaitu tes yang dilaksanakan untuk mengungkap kemampuan dasar atau bakat khusus yang dimiliki oleh *testee*”.²⁰

¹⁷ S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, (Yogyakarta: Pustaka Pelajar, 2014), hlm.57.

¹⁸ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 73.

¹⁹ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, hlm. 73.

²⁰ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, , hlm. 73.

c) Tes Sikap (*attitude test*)

Tes sikap yaitu jenis tes yang digunakan untuk mengungkap kecenderungan seseorang dalam melakukan suatu respon tertentu terhadap dunia sekitarnya, baik berupa individu maupun obyek tertentu.²¹

d) Tes Kepribadian (*personality test*)

“Tes kepribadian yaitu tes yang bertujuan mengungkap ciri khas dari seseorang berdasarkan sifat lahiriah”.²²

e) Tes Hasil Belajar atau Tes Pencapaian (*achievement test*)

“Tes yang digunakan untuk mengungkap tingkat pencapaian atau prestasi belajar”.²³ Tes hasil belajar digunakan untuk pengukuran dan penilaian hasil belajar berbentuk tugas atau perintah yang harus dikerjakan oleh *testee*, sehingga dapat dihasilkan nilai sebagai lambang tingkah laku atau prestasi.

²¹ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 73.

²² Anas Sudijono, *Pengantar Evaluasi Pendidikan*, hlm. 74.

²³ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, , hlm. 74.

5) Berdasarkan cara mengajukan pertanyaan dan cara memberikan jawaban, tes dibagi menjadi dua yaitu:

a) Tes Tertulis (*pencil and paper test*)

Tes tertulis yaitu jenis tes dimana *tester* dalam mengajukan butir-butir pertanyaan atau soal dilakukan secara tertulis dan *testee* memberikan jawaban secara tertulis.

b) Tes Lisan (*nonpencil and paper test*)

Tes lisan yaitu jenis tes dimana *tester* dalam mengajukan pertanyaan-pertanyaan atau soal dilakukan secara lisan dan *testee* memberikan jawaban secara lisan.

6) Berdasarkan bentuk responnya dibedakan menjadi dua, yaitu: *verbal test* dan *nonverbal test*.²⁴ Adapun penjelasannya sebagai berikut:

a) *Verbal Test*

“*Verbal test* yaitu suatu tes yang menghendaki respon dalam bentuk ungkapan kata-kata atau kalimat, baik secara lisan maupun tertulis”.²⁵

b) *Nonverbal Test*

“*Nonverbal test* yaitu tes yang menghendaki respon dari *testee* berupa tindakan atau tingkah laku, sehingga respon yang dikehendaki terbentuk

²⁴ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 75.

²⁵ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, hlm. 75.

melalui perbuatan atau gerakan-gerakan".²⁶

Salah satu bentuk *nonverbaltest* yaitu tes praktikum.

c. Ciri-ciri Tes yang Baik

Tes yang baik harus memenuhi persyaratan tes, yaitu memiliki validitas, reliabilitas, objektivitas, praktikabilitas, dan ekonomis.²⁷

1) Validitas

Tes hasil belajar yang baik adalah tes hasil belajar yang bersifat valid atau memiliki validitas. Valid sering diartikan dengan tepat, shahih, benar, dan absah.

Sebuah tes dikatakan valid apabila tes tersebut dengan secara tepat, benar, shahih, atau secara absah telah dapat mengukur apa yang seharusnya diukur atau mengungkap hasil belajar yang telah dicapai oleh peserta didik setelah peserta didik menempuh proses belajar mengajar dalam jangka waktu tertentu.²⁸

2) Reliabilitas

Reliabilitas merupakan keajegan atau kemantapan.

²⁶Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 75.

²⁷ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, (Jakarta: Bumi Aksara, 2011), hlm. 57.

²⁸ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, hlm. 93.

Sebuah tes dinyatakan reliabel apabila hasil-hasil pengukuran yang dilakukan dengan menggunakan tes tersebut secara berulang kali terhadap subyek yang sama menunjukkan hasil yang tetap sama atau bersifat ajeg dan stabil.²⁹

3) Objektivitas

“Tes hasil belajar dikatakan objektif apabila tes tersebut disusun dan dilaksanakan menurut apa adanya”.³⁰ Berdasarkan segi materi tes, istilah apa adanya bermakna bahwa materi tes tersebut diambil atau bersumber dari materi atau bahan pelajaran yang telah diberikan sesuai dengan tujuan instruksional khusus yang telah ditentukan. Sedangkan dari segi pemberian skor dan penentuan nilai hasil tes, istilah apa adanya mempunyai maksud pekerjaan koreksi, pemberian skor, dan penentuan nilainya terhindar dari unsur-unsur subyektivitas pada diri penyusun tes.

4) Praktikabilitas

Tes dapat dikatakan memiliki praktikabilitas yang tinggi apabila tes tersebut bersifat praktis, mudah pengadministrasiannya. Menurut Suharsimi Arikunto,

Tes yang praktis adalah tes yang memenuhi tiga kriteria, yaitu: (a.) mudah dilaksanakan, tidak menuntut peralatan banyak dan memberi

²⁹ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 95.

³⁰ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, hlm. 96.

kebebasan pada siswa untuk mengerjakan bagian yang mudah terlebih dahulu, (b.) mudah pemeriksaannya, tes tersebut dilengkapi dengan kunci jawaban maupun pedoman skoringnya, (c.) dilengkapi dengan petunjuk-petunjuk yang jelas sehingga dapat diberikan atau diawali oleh orang lain.³¹

5) Ekonomis

“Tes yang memiliki nilai ekonomis yaitu pelaksanaan tes tersebut tidak membutuhkan biaya yang mahal, tenaga yang banyak, dan waktu yang lama”.³² Maksudnya yaitu tes tersebut hemat dalam biaya tenaga dan waktu.

2. Analisis butir item soal (*Item Analysis*)

a. Pengertian Analisis Butir Soal

“Analisis butir soal adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai”.³³ Analisis butir soal merupakan kegiatan evaluasi terhadap alat pengukur yang telah digunakan untuk mengukur keberhasilan belajar dari para peserta didik.³⁴ Alat pengukur yang dimaksud yaitu

³¹ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, (Jakarta: Bumi Aksara, 2011), hlm. 62.

³² Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, hlm. 63.

³³ Nana Sudjana, *Penilaian Hasil Belajar Mengajar*, (Bandung: PT Remaja Rosdakarya, 2014), hlm. 135.

³⁴ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 367.

tes hasil belajar yang terdiri dari kumpulan butir-butir soal (item).

b. Tujuan Analisis Butir Soal

Tujuan khusus dari analisis butir soal ialah untuk mencari soal tes yang baik, soal tes yang tidak baik dan alasan soal tersebut dikatakan baik atau tidak baik.³⁵ Menurut Anas Sudijono dalam buku *pengantar evaluasi pendidikan*,

Analisis butir soal bertujuan untuk mengetahui apakah butir-butir item yang membangun tes hasil belajar tersebut sudah dapat menjalankan fungsinya sebagai alat pengukur hasil belajar yang memadai atau belum.³⁶

Analisis butir soal juga bertujuan untuk mengidentifikasi soal-soal yang baik, kurang baik, dan soal yang jelek. Kegiatan analisis soal menghasilkan informasi tentang kejelekan sebuah soal dan petunjuk untuk mengadakan perbaikan. Penganalisisan terhadap butir-butir soal tes hasil belajar dapat dilakukan dari tiga segi, yaitu segi derajat kesukaran itemnya, segi daya pembeda itemnya, dan segi fungsi distraktornya.

Menurut Shodiq Abdullah, item tes yang baik adalah item yang memenuhi syarat karakteristik item tes yang baik. Karakteristik item yang

³⁵ Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, (Bandung: PT Remaja Rosdakarya, 2002), hlm. 118.

³⁶ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 369.

dimaksud yaitu tingkat kesukaran, daya beda dan efektivitas pengecoh.³⁷

3. Validitas tes

a. Pengertian Validitas

“Validitas berasal dari kata *validity*, diartikan sebagai ketepatan atau kesahihan, yaitu sejauh mana sebuah instrumen atau alat ukur mampu atau berhasil mengukur apa yang seharusnya diukur”.³⁸ Validitas juga dapat diterjemahkan sebagai kesahihan atau ketepatan, yaitu sejauhmana sebuah instrumen atau alat ukur mampu atau berhasil mengukur apa yang hendak diukurnya, atau sejauhmana sebuah instrumen memenuhi fungsi ukurnya.

Menurut Sukardi dalam buku *evaluasi pendidikan prinsip dan operasionalnya*, “validitas yaitu derajat yang menunjukkan di mana suatu tes mengukur apa yang hendak diukur”.³⁹

³⁷ Shodiq Abdullah, *Evaluasi Pembelajaran (Konsep dasar, Teori dan Aplikasi)*, (Semarang: Pustaka Rizki Putra, 2012), hlm. 98.

³⁸ Shodiq Abdullah, *Evaluasi Pembelajaran (Konsep dasar, Teori dan Aplikasi)*, hlm. 76.

³⁹ Sukardi, *evaluasi pendidikan prinsip dan operasionalnya*, (Jakarta: PT Bumi Aksara, 2009), hlm. 46.

Menurut S. Eko Widoyoko “validitas berkaitan dengan ketepatan dengan alat ukur, dengan instrumen yang valid akan menghasilkan data yang valid pula”.⁴⁰

Menurut Shodiq Abdullah, validitas diterjemahkan sebagai kesahihan atau ketepatan, yaitu sejauhmana alat ukur mampu atau berhasil mengukur apa yang hendak diukur atau sejauhmana alat ukur memenuhi fungsinya.⁴¹

Menurut William Wiersma Stephen G. Jurs mendefinisikan “*validity is the extend to which a test measures what it is intended to measure.*”⁴² Artinya validitas adalah sejauh mana sebuah tes mengukur apa yang dimaksudkan untuk diukur”.

Menurut Journal on Development Disabilities mendefinisikan “*validity refers to whether the instrument mesures what it intends to measure and describe*”.⁴³ Artinya validitas mengacu pada apakah instrumen mengukur apa yang hendak diukur dan dijelaskan (dideskripsikan).

⁴⁰ S. Eko Putro Widoyoko, *Penilaian Hasil Pembelajaran Di Sekolah*, (Yogyakarta: Pustaka Pelajar, 2014), hlm.172.

⁴¹ Shodiq Abdullah, *Evaluasi Pembelajaran (Konsep dasar, Teori dan Aplikasi)*, (Semarang: Pustaka Rizki Putra, 2012), hlm. 77.

⁴² Wiliiam Wiersma Stephen G. Jurs, *Educational Measurement and Testing*, (United States: A Division of Simon & Schuster, 1990), hlm. 183.

⁴³ Anonim, “*Revisiting the Concept of Reliability and Validity*”, *Journal on Development Disabilities*, (Vol. 13, No. 3/ 2007), hlm. 7.

Berdasarkan beberapa definisi diatas, dapat disimpulkan bahwa validitas tes disebut sebagai kesahihan atau ketepatan tes yaitu sejauhmana sebuah tes dapat mengukur dengan tepat atau sah terhadap apa yang seharusnya diukur.

b. Macam-macam validitas

Secara garis besar ada dua macam validitas, yaitu validitas soal dan validitas butir soal (validitas item)

1) Validitas Soal

Secara garis besar ada dua macam validitas, yaitu validitas logis dan validitas empiris

a) Validitas Logis

Istilah validitas logis mengandung kata logis berasal dari kata logika, yang berarti penalaran. Validitas logis sebuah instrumen evaluasi menunjuk pada kondisi bagi sebuah instrumen yang memenuhi persyaratan valid berdasarkan hasil penalaran.⁴⁴

Ada dua macam validitas logis yaitu validitas isi dan validitas kontrak (*construct validity*)

⁴⁴ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, (Jakarta: Bumi Aksara, 2011), hlm. 65.

(1) Validitas Isi

Sebuah tes dikatakan memiliki validitas isi apabila mengukur tujuan khusus tertentu yang sesuai dengan materi atau isi pelajaran yang diberikan.⁴⁵ Butir-butir tes hasil belajar dinyatakan valid (*logically valid*) apabila setelah mencermati isi butir-butir yang ditulis telah menunjukkan kesesuaian dengan kisi-kisi.⁴⁶

(2) Validitas Konstrak

Secara terminologi, tes hasil belajar dinyatakan sebagai tes yang memiliki validitas konstruksi apabila tes hasil belajar tersebut dapat secara tepat mencerminkan suatu konstruksi dalam teori psikologis. Jiwa seorang peserta didik dapat dirinci ke dalam beberapa aspek atau ranah tertentu. Benjamin S. Bloom membagi aspek kejiwaan menjadi tiga bagian yaitu aspek kognitif (*cognitive domain*), aspek afektif (*affective domain*), dan aspek psikomotorik (*psychomotorik*

⁴⁵ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan (Edisi Revisi)*, (Jakarta: Bumi Aksara, 2011), hlm. 67.

⁴⁶ Purwanto, *Evaluasi Hasil Belajar*, (Yogyakarta: Pustaka Pelajar, 2011), hlm. 121.

domain).⁴⁷ Adapun rincian domain tersebut adalah sebagai berikut:

(a) Ranah Kognitif

Ranah kognitif berkaitan dengan tujuan pendidikan yang meliputi fungsi-fungsi mental seperti pemanggilan kembali informasi dan kemampuan intelektual . Ranah kognitif dibagi menjadi enam tingkatan utama, dari yang paling sederhana (*simple*) sampai dengan yang paling rumit (*complex*).⁴⁸ Enam tingkatan Taksonomi Bloom ranah kognitif yang telah direvisi Anderson dan Krathwohl yakni: mengingat (*remember*), memahami/mengerti (*understand*), menerapkan (*apply*), menganalisis (*analyze*), mengevaluasi (*evaluate*), dan menciptakan (*create*).⁴⁹ Tingkatan pada

⁴⁷ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 166.

⁴⁸ Shodiq Abdullah, *Evaluasi Pembelajaran (Konsep dasar, Teori dan Aplikasi)*, (Semarang: Pustaka Rizki Putra, 2012), hlm. 20.

⁴⁹ Anderson, L.W., dan Krathwohl, D.R, *A Taxonomy for Learning, Teaching, and Assesing: A Revision of Bloom's Taxonomy of Educatioanl Objectives*, (New York: Addison Wesley Longman, Inc, 2001)hlm. 61-88.

ranah kognitif dapat dilihat pada lampiran VIII.

(b) Ranah Afektif

Ranah afektif berkenaan dengan sikap dan nilai. Ranah afektif dibagi menjadi lima sub bab ranah, yaitu: *receiving, responding, valuing, organization, dan characterization by a value kompleks*.⁵⁰

(c) Ranah Psikomotorik

“Ranah psikomotorik yaitu ranah yang berkaitan dengan keterampilan (skill)”⁵¹

Ranah psikomotorik dibagi menjadi tujuh sub ranah, yaitu persepsi, kesiapan, respon terbimbing, gerakan terbiasa, respon kompleks, adaptasi, dan originasi.

b) Validitas Empiris

Validitas empiris mengandung kata “empiris” yang berarti pengalaman. Sebuah instrumen dapat dikatakan memiliki validitas empiris apabila sudah diuji dari pengalaman. Ada dua macam validitas empiris, yaitu:

⁵⁰ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 54.

⁵¹ Shodiq Abdullah, *Evaluasi Pembelajaran (Konsep dasar, Teori dan Aplikasi)*, (Semarang: Pustaka Rizki Putra, 2012), hlm. 35.

(1) Validitas Kongkuren (*concurrent validity*)

Validitas kongkuren adalah derajat dimana skor dalam suatu tes dihubungkan dengan skor lain yang telah dibuat. Tes dengan validasi konkuren biasanya diadministrasi dalam waktu yang sama atau dengan kriteria valid yang sudah ada.

(2) Validitas Prediksi (*predictive validity*)

“Validitas prediktif adalah jika kriteria standar yang digunakan adalah untuk meramalkan prestasi belajar murid masa yang akan datang”.⁵² Validitas prediktif bertujuan untuk melihat sejauh mana suatu tes dapat memprakirakan perilaku peserta didik pada masa yang akan datang.

2) Validitas Butir Soal

Menurut Anas Sudijono dalam buku Pengantar Evaluasi Pendidikan,

Validitas butir soal atau validitas item yaitu ketepatan mengukur yang dimiliki oleh sebutir item (sebutir soal tes) dalam mengukur apa yang seharusnya diukur melalui butir item tersebut. Sebutir soal tes merupakan bagian tak terpisahkan dari tes sebagai suatu totalitas.⁵³

⁵² Zainal Arifin, *Evaluasi Pembelajaran*, (Bandung: PT Remaja Rosdakarya, 2012), hlm. 249.

⁵³ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 182.

Hubungan antara item tes dengan tes hasil belajar dapat diketahui dari semakin banyak butir-butir item yang dapat dijawab dengan betul oleh *testee*, maka skor total hasil tes tersebut akan semakin tinggi. Sebaliknya semakin sedikit butir soal yang dapat dijawab dengan betul oleh *testee*, maka skor total hasil tes tersebut akan semakin rendah.

c. Teknik Pengujian Validitas Butir Soal

Sebutir item dikatakan memiliki validitas yang tinggi atau dikatakan valid, jika skor-skor pada butir item yang bersangkutan memiliki kesesuaian atau korelasi positif yang signifikan antara skor item dengan skor totalnya. Skor total berkedudukan sebagai variabel terikat (dependent variabel), sedangkan skor item berkedudukan sebagai variabel bebas (independent variabel).

Pada tes dengan bentuk soal objektif memiliki dua kemungkinan jawaban yaitu benar dan salah, umumnya untuk setiap item yang dijawab benar diberikan skor 1(satu) dan item yang dijawab salah diberikan skor 0 (nol). Dalam ilmu statistik, jenis data seperti benar-salah, ya-tidak atau sejenisnya disebut dengan nama data diskret murni atau data dikotomik. Sedangkan skor total yang dimiliki oleh setiap individu *testee* yang merupakan hasil penjumlahan dari setiap skor yang dimiliki oleh setiap butir item disebut dengan data kontinyu.

Teknik korelasi yang dapat digunakan dalam mencari korelasi antara variabel I dengan variabel II adalah teknik korelasi point biserial.⁵⁴

4. Reliabilitas Tes

a. Pengertian Reliabilitas Tes

Reliability berasal dari kata *rely* yang artinya percaya dan reliabel yang artinya dapat dipercaya. Menurut Arif Furchan “reliabilitas suatu alat pengukur adalah derajat keajegan alat tersebut dalam mengukur apa saja yang diukurnya”.⁵⁵ Suatu tes dapat dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda.

Menurut Nana Sudjana, ”reliabilitas alat penilaian adalah ketepatan atau keajegan alat tersebut dalam menilai apa yang dinilainya”.⁵⁶ Artinya, kapan pun alat penilaian tersebut digunakan akan memberikan hasil yang relatif sama.

Menurut Chabib Thoha dalam bukunya yang berjudul *Teknik Evaluasi Pendidikan*,

⁵⁴ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 185.

⁵⁵ Arief Furchan, *Pengantar Penelitian Dalam Pendidikan*, (Yogyakarta: Pustaka Pelajar, 2004), hlm. 310.

⁵⁶ Nana Sudjana, *Penilaian Hasil Belajar Mengajar*, (Bandung: PT Remaja Rosdakarya, 2014), hlm. 16.

Reliabilitas diartikan dengan keajegan bilamana tes tersebut diujikan berkali-kali hasilnya relatif sama, artinya setelah hasil tes pertama dengan hasil tes berikutnya dikorelasikan terdapat hasil korelasi yang signifikan.⁵⁷

Menurut William Wiersma Stephen G. Jurs mendefinisikan “ *reliability of measurement is consistency – consistency in measuring whether the instrument is measuring*”.⁵⁸ Artinya reabilitas pengukuran yaitu konsistensi – konsistensi dalam mengukur apapun instrumen yang mengukur (instrumen yang digunakan untuk mengukur) .

Menurut Ellen A. Drost mendefinisikan “*reliability is stability of measurement over a variety of conditions in which basically the same results should be obtained*”.⁵⁹ Artinya stabilitas pengukuran atas berbagai kondisi di mana pada dasarnya hasil yang sama harus diperoleh.

Berdasarkan beberapa definisi diatas, dapat disimpulkan bahwa reliabilitas tes adalah keajegan atau

⁵⁷ Chabib Thoha, *Teknik Evaluasi pendidikan*, (Jakarta: PT Raja Grafindo Persada, 2001), hlm. 118.

⁵⁸ Wiliiam Wiersma Stephen G. Jurs, *Educational Measurement and Testing*, (United States: A Division of Simon & Schuster, 1990), hlm. 155.

⁵⁹ Ellen A. Drost, “*Validity and Reliability in Social Science Research*”, *Education Research and Perspectives*, (Vol.38, No.1), hlm. 106.

konsistensi atau stabilitas sebuah alat ukur (berupa tes) dalam mengukur apa yang diukurinya. Sehingga apabila tes tersebut diujikan beberapa kali dalam kondisi yang berbeda maka akan diperoleh hasil yang sama.

b. Teknik Pengujian Reliabilitas

Penentuan reliabilitas pada tes hasil belajar bentuk objektif dilakukan dengan menggunakan tiga macam pendekatan yaitu *Single Test-Single Trial (Single Test-Single Trial Method)*, *Single Test- Single Retest (Single Test- Single Retest Method)*, dan *Alternate Form (Double Test – Double Trial Method)*.

1) Pendekatan *Single Test-Single Trial (Single Test-Single Trial Method)*

Menurut Anas Sudijono dalam buku *Pengantar Evaluasi Pendidikan*,

Pendekatan *Single Test-Single Trial* yaitu penentuan reliabilitas dengan cara melakukan pengukuran terhadap satu kelompok subyek, pengukuran dilakukan dengan menggunakan satu jenis alat ukur, dan pelaksanaan pengukuran dilakukan sebanyak satu kali.⁶⁰

Pendekatan ini memiliki beberapa teknik untuk menentukan besarnya reliabilitas yaitu teknik pembelahan ganjil genap dan pembelahan awal-akhir

⁶⁰ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 213.

dengan menggunakan rumus Spearman-Brown, penggunaan rumus Flanagan, penggunaan rumus Rulon, penggunaan rumus K-R. 20, penggunaan rumus K-R. 21, dan penggunaan rumus Hoyt. Penentuan reliabilitas tes hasil belajar bentuk objektif menggunakan rumus Spearman-Brown, rumus Flanagan, dan rumus Rulon dilakukan dengan membelah dua tes, sehingga sering disebut dengan teknik belah dua. Sedangkan dengan menggunakan rumus Kuder-Richardson (K-R) dan rumus Hoyt penentuan reliabilitas tes hasil belajar bentuk objektif tidak menggunakan teknik belah dua tes, akan tetapi langsung dimasukkan kedalam rumus.

2) Pendekatan *Single Test-Single Retest (Single Test-Single Retest Method)*

Pendekatan Single Test- Single Retest dikenal dengan istilah pendekatan bentuk ulangan, sehingga penentuan reliabilitas tes dilakukan dengan menggunakan teknik ulangan dimana *tester* hanya menggunakan satu seri tes, tetapi percobaannya dilakukan sebanyak dua kali.⁶¹

Setelah dilaksanakan *testing*, keseluruhan skor hasil tes pertama dikorelasikan dengan

⁶¹ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 268.

keseluruhan skor hasil tes kedua. Jika terdapat korelasi positif yang signifikan antara skor hasil tes pertama dengan skor hasil tes kedua, maka tes hasil belajar tersebut dikatakan reliabel, karena antara skor-skor hasil tes pertama dengan skor-skor hasil tes kedua memperlihatkan adanya keajegan atau kestabilan.

3) Pendekatan *Alternate Form (Double Test – Double Trial Method)*

Pendekatan *Alternate Form* yaitu penentuan reliabilitas dengan cara memberikan dua buah tes yang diberikan kepada sekelompok subyek yang dilakukan secara berbarengan dengan ketentuan bahwa kedua tes tersebut harus sejenis, maksudnya yaitu meskipun butir soal tesnya tidak sama, namun butir soal tes tersebut mengukur hal yang sama, baik dari segi isi, proses mental yang diukur, derajat kesukaran, ataupun jumlah butir butir soal tesnya.

Pengujian reliabilitas tes dengan menggunakan pendekatan *alternate form* atau bentuk paralel, skor-skor yang diperoleh dari kedua seri tes dicari korelasinya. Apabila terdapat korelasi positif yang signifikan maka dikatakan bahwa tes hasil belajar tersebut dikatakan reliabel.⁶² Teknik

⁶² Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 273.

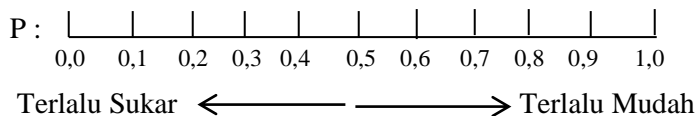
pengujiannya dapat menggunakan rumus korelasi *product moment* atau rumus korelasi *rank order Spearman*.

5. Tingkat Kesukaran

Menurut C. Boopathiraj dan Dr. K. Chellamani mendefinisikan “*Item difficulty may be defined as the proportion of the examinees that marked the item correctly*, yang artinya item kesulitan dapat didefinisikan sebagai proporsi peserta ujian yang menandai item dengan benar”.⁶³ Tingkat kesukaran (*difficulty index*) dalam evaluasi hasil belajar umumnya dilambangkan dengan huruf P (proporsi), didefinisikan sebagai proporsi siswa peserta tes yang menjawab benar. Butir-butir item tes hasil belajar dinyatakan sebagai butir-butir item yang baik apabila butir-butir item tersebut tidak terlalu sukar dan tidak terlalu mudah, dengan kata lain derajat kesukaran item tersebut adalah sedang atau cukup.⁶⁴ Besarnya indeks kesukaran antara 0,0 sampai dengan 1,0. Besarnya angka indeks kesukaran menunjukkan taraf kesukaran soal, seperti pada skala dibawah ini:

⁶³ C. Boopathiraj and Dr. K. Chellamani, “*Analysis of Test Items On Difficulty Level and Discrimination Index in The Test For Research in Education*”, *International Journal of Social Science & Interdisciplinary Research*, (Vol. 2, No. 2, February/2013), hlm. 190.

⁶⁴ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 370.



Angka indeks kesukaran sebesar 0,0 ($P = 0,0$) menunjukkan bahwa butir soal tes tersebut termasuk dalam kategori item yang terlalu sukar, karena seluruh *testee* tidak dapat menjawab dengan betul (benar) butir soal tes yang bersangkutan, sebaliknya angka indeks kesukaran sebesar 1,0 menunjukkan bahwa butir soal tes terlalu mudah, sebab seluruh *testee* dapat menjawab dengan betul butir soal tes yang bersangkutan.

6. Daya Pembeda

Menurut Ngalim Purwanto dalam buku *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*

Daya pembeda suatu soal tes yaitu bagaimana kemampuan soal tersebut untuk membedakan siswa-siswa yang termasuk kelompok pandai (*upper group*) dengan siswa-siswa yang termasuk kelompok kurang (*lower group*).⁶⁵

Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi, disingkat D yaitu *discriminatory power*. Indeks diskriminasi berkisar antara 0,00 sampai 1,00

⁶⁵ Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, (Bandung: PT Remaja Rosdakarya, 2002), hlm. 120.

dengan adanya tanda positif-negatif. Ada tiga titik pada daya pembeda yaitu:

-1,00 \longleftrightarrow 0,00 \longleftrightarrow +1,00

Daya pembeda (-) Daya pembeda nihil Daya pembeda (+)

Sebutir soal tes yang memiliki indeks diskriminasi dengan tanda positif (+) menunjukkan bahwa butir tes tersebut telah memiliki daya pembeda, karena *testee* yang termasuk kategori *upper group* lebih banyak menjawab dengan betul terhadap butir tes yang bersangkutan, sedangkan *testee* yang termasuk dalam kategori *lower group* lebih banyak yang menjawab salah. Sebutir soal tes yang memiliki indeks diskriminasi 0,0 (nihil), maka hal tersebut menunjukkan bahwa butir soal tes tersebut tidak memiliki daya pembeda dan butir soal tes yang memiliki indeks diskriminasi dengan tanda negatif (-) mengandung arti bahwa butir tes yang bersangkutan lebih banyak dijawab betul oleh *testee* dalam kategori *lower group* dibandingkan *testee* yang dikategorikan *upper group*.

7. Fungsi Distraktor

Menurut Anas Sudijono dalam buku *Pengantar Evaluasi Pendidikan*,

Distraktor yang berfungsi dengan baik yaitu distraktor yang memiliki daya tarik atau daya rangsang, sehingga *testee* merasa bimbang, dan ragu-ragu yang

pada akhirnya mereka menjadi terkecoh untuk memilih distraktor sebagai jawaban betul.⁶⁶

Distraktor atau pengecoh yaitu beberapa alternatif jawaban yang salah atau jawaban yang bukan merupakan kunci jawaban yang terdapat pada setiap butir tes. Tujuan dari pemasangan distraktor pada setiap butir tes yaitu agar dari banyaknya *testee* yang mengikuti tes hasil belajar ada yang tertarik untuk memilih distraktor tersebut, sebab *testee* menyangka bahwa distraktor yang dipilih merupakan jawaban betul.

8. Mata Pelajaran Biologi

a. Deskripsi Mata Pelajaran Biologi

Biologi sebagai salah satu bidang IPA menyediakan berbagai pengalaman belajar untuk memahami konsep dan proses sains. Keterampilan proses ini meliputi keterampilan mengamati, mengajukan hipotesis, menggunakan alat dan bahan secara baik dan benar dengan selalu mempertimbangkan keamanan dan keselamatan kerja, mengajukan pertanyaan, menggolongkan dan menafsirkan data, serta mengkomunikasikan hasil temuan secara lisan atau tertulis, menggali dan memilah informasi faktual yang relevan untuk menguji gagasan-gagasan atau memecahkan

⁶⁶ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Pers, 2009), hlm. 410.

masalah sehari-hari. Mata pelajaran Biologi dikembangkan melalui kemampuan berpikir analitis, induktif, dan deduktif untuk menyelesaikan masalah yang berkaitan dengan peristiwa alam sekitar. Penyelesaian masalah yang bersifat kualitatif dan kuantitatif dilakukan dengan menggunakan pemahaman dalam bidang matematika, fisika, kimia dan pengetahuan pendukung lainnya.

b. Tujuan Mata Pelajaran Biologi

Mata pelajaran Biologi bertujuan agar peserta didik memiliki kemampuan sebagai berikut:

- 1) Membentuk sikap positif terhadap biologi dengan menyadari keteraturan dan keindahan alam serta mengagungkan kebesaran Tuhan Yang Maha Esa.
- 2) Memupuk sikap ilmiah yaitu jujur, objektif, terbuka, ulet, kritis dan dapat bekerjasama dengan orang lain.
- 3) Mengembangkan pengalaman untuk dapat mengajukan dan menguji hipotesis melalui percobaan, serta mengkomunikasikan hasil percobaan secara lisan dan tertulis.
- 4) Mengembangkan kemampuan berpikir analitis, induktif, dan deduktif dengan menggunakan konsep dan prinsip biologi.
- 5) Mengembangkan penguasaan konsep dan prinsip biologi dan saling keterkaitannya dengan IPA lainnya

serta mengembangkan pengetahuan, keterampilan dan sikap percaya diri.

- 6) Menerapkan konsep dan prinsip biologi untuk menghasilkan karya teknologi sederhana yang berkaitan dengan kebutuhan manusia.
 - 7) Meningkatkan kesadaran dan berperan serta dalam menjaga kelestarian lingkungan.
- c. Ruang Lingkup Mata Pelajaran Biologi

Mata pelajaran Biologi di SMA / MA merupakan kelanjutan IPA di SMP/MTs yang menekankan pada fenomena alam dan penerapannya yang meliputi aspek-aspek sebagai berikut:

- 1) Hakikat biologi, keanekaragaman hayati dan pengelompokan makhluk hidup, hubungan antarkomponen ekosistem, perubahan materi dan energi, peranan manusia dalam keseimbangan ekosistem.
- 2) Organisasi seluler, struktur jaringan, struktur dan fungsi organ tumbuhan, hewan dan manusia serta penerapannya dalam konteks sains, lingkungan, teknologi dan masyarakat.
- 3) Proses yang terjadi pada tumbuhan, proses metabolisme, hereditas, evolusi, bioteknologi dan implikasinya pada sains, lingkungan, teknologi dan masyarakat.

- d. Standar kompetensi dan kompetensi dasar mata pelajaran Biologi kelas XI SMA atau MA.⁶⁷ Selanjutnya SK dan KD dapat dilihat pada tabel 2.4.

Tabel 2.1
Standar Kompetensi dan Kompetensi Dasar Mata Pelajaran Biologi Kelas XI Semester Gasal

Standar Kompetensi	Kompetensi Dasar
1. Memahami struktur dan fungsi sel sebagai unit terkecil kehidupan	1.1 Mendeskripsikan komponen kimiawi sel, struktur dan fungsi sel sebagai unit terkecil kehidupan 1.2 Mengidentifikasi organela sel tumbuhan dan hewan 1.3 Membandingkan mekanisme transpor pada membran (difusi, osmosis, transport aktif, endositosis, eksositosis)
2. Memahami keterkaitan antara struktur dan fungsi jaringan tumbuhan dan hewan, serta penerapannya dalam konteks Salingtemas	2.1 Mengidentifikasi struktur jaringan tumbuhan dan mengaitkannya dengan fungsinya, menjelaskan sifat totipotensi sebagai dasar kultur jaringan 2.2 Mendeskripsikan struktur jaringan hewan Vertebrata dan mengaitkannya dengan fungsinya
3. Menjelaskan struktur dan fungsi organ manusia dan hewan tertentu,	3.1 Menjelaskan keterkaitan antara struktur, fungsi, dan proses serta kelainan/penyakit yang dapat terjadi pada sistem gerak pada manusia

⁶⁷ Standar Kompetensi dan Kompetensi Dasar Mata Pelajaran Biologi SMA/ MA, (Jakarta:Dirjen Dikmenum, 2007).

Standar Kompetensi	Kompetensi Dasar
kelainan/penyakit yang mungkin terjadi serta implikasinya pada Salingtemas	3.2 Menjelaskan keterkaitan antara struktur, fungsi, dan proses serta kelainan/penyakit yang dapat terjadi pada sistem peredaran darah

B. Kajian Pustaka

Kajian pustaka digunakan untuk memperoleh suatu informasi tentang teori yang ada kaitannya dengan judul penelitian dan digunakan untuk memperoleh landasan teori ilmiah. Dalam kajian pustaka ini peneliti menelaah beberapa karya ilmiah antara lain:

Penelitian Himatul Aliyah, NIM 103111112 mahasiswa UIN Walisongo Semarang Fakultas Ilmu Tarbiyah dan Keguruan tahun 2014, dengan judul “*Studi Analisis Butir-butir Soal Objektif Berbentuk Multiple Choice Buatan MGMP LP Ma’arif Kabupaten Kendal Pada Mata Pelajaran Al-Quran Hadits Semester Ganjil Siswa Kelas VIII MTS Tahun Pelajaran 2013/2014*”. Hasil penelitian ini menunjukkan tes pada mata pelajaran Al-Quran Hadits semester ganjil siswa kelas VIII termasuk dalam kategori tes yang memiliki validitas butir tes sedang karena memiliki 17 butir soal atau sekitar 42,5% butir-butir soal tes tersebut dinyatakan valid dan 23 atau sekitar 57,5% butir soal yang lain dinyatakan dalam kategori invalid. Memiliki reliabilitas tes yang

baik, tingkat kesukaran sedang, daya pembeda sedang, dan memiliki fungsi distraktor yang baik.⁶⁸

Penelitian Lilis Tri Ariyana NIM 4401406059, mahasiswa Universitas Negeri Semarang Fakultas Matematika dan Ilmu Pengetahuan Alam tahun 2011, dengan judul “Analisis Butir Soal Ulangan Akhir Semester Gasal IPA Kelas IX SMP di Kabupaten Grobogan”. Hasil penelitian ini menunjukkan validitas logis karena sudah sesuai dengan soal standar, reliabel dengan kategori tinggi, tingkat kesukaran sedang, daya beda baik, efektifitas pengecoh berfungsi.⁶⁹

Penelitian Shinta Widyarini NIM 11403241047, mahasiswa Program Studi Pendidikan Akuntansi Jurusan Pendidikan Akuntansi Fakultas Ekonomi Universitas Negeri Yogyakarta 2015 Berdasarkan analisis butir soal secara bersama-sama soal yang sangat baik berjumlah 3 dari 50 butir soal atau 6%, soal yang baik berjumlah 11 dari 50 butir atau 22%, soal yang sedang berjumlah 7 dari 50 butir atau 14%, soal yang tidak baik berjumlah 16 dari 50 butir atau 32%, dan soal yang berkualitas sangat tidak baik berjumlah 13 dari 50 butir atau 26%. soal ujian

⁶⁸ Himatul Aliyah, “Studi Analisis Butir-butir Soal Objektif Berbentuk Multiple Choice Buatan MGMP LP Ma’arif Kabupaten Kendal Pada Mata Pelajaran Al-Quran Hadits Semester Ganjil Siswa Kelas VIII MTSTahun Pelajaran 2013/2014”, Skripsi, (Semarang: Fakuultas Ilmu Tarbiyah Dan Keguruan UIN Walisongo Semarang, 2014)

⁶⁹ Lilis Tri Ariyana, “Analisis Butir Soal Ulangan Akhir Semester Gasal IPA Kelas IX SMP di Kabupaten Grobogan”, *Skripsi*, (Semarang: Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Negeri Semarang, 2011)

akhir semester gasal mata pelajaran Ekonomi Akuntansi kelas XII IPS SMA N Kalasan tahun ajaran 2014/2015 merupakan soal yang jumlah soal yang valid 28 soal atau 56% sedangkan soal yang tidak valid berjumlah 22 soal atau 44%, memiliki tingkat reliabilitas sebesar 0,7129, memiliki daya pembeda jelek berjumlah 27 butir atau sebesar 54%, butir soal yang memiliki daya pembeda cukup berjumlah 15 butir atau sebanyak 30%, butir soal yang memiliki daya pembeda baik berjumlah 1 butir atau sebesar 5% dan butir soal yang memiliki daya pembeda negatif berjumlah 7 butir soal atau sebesar 14%. memiliki tingkat kesukaran baik berjumlah 11 butir (22%) dengan rincian butir yang tergolong sukar berjumlah 9 butir atau 18%, butir soal yang tergolong sedang berjumlah 11 butir atau 22%, dan butir soal yang tergolong mudah berjumlah 30 butir atau 60%, memiliki efektivitas pengecoh termasuk tidak baik karena soal yang efektifitas pengecohnya baik berjumlah 21 butir (42%), dengan rincian terdapat 6 (12%) butir soal baik, 15 (30%) butir soal cukup baik, 11 (22%) butir soal kurang baik, dan 18 (36%) butir soal tidak baik.⁷⁰

Meskipun hampir memiliki kesamaan dengan karya peneliti sebelumnya, yakni membahas tentang validitas, reliabilitas tes, tingkat kesulitan, daya pembeda dan fungsi

⁷⁰ Shinta Widayarni, “ Analisis Kualitas Butir Soal Ujian Akhir Semester Gasal Mata Pelajaran Ekonomi Akuntansi Kelas XII IPS SMA N 1 Kalasan Tahun 2014/2015”, *Skripsi*, (Fakultas Ekonomi Universitas Negeri Yogyakarta, 2015)

pengecoh, namun secara prinsip memiliki perbedaan yaitu pada fokus atau obyek penelitian. Fokus pada penelitian ini adalah validitas, reliabilitas, tingkat kesulitan, daya pembeda dan fungsi pengecoh pada tes buatan MGMP Biologi Kabupaten Kendal pada mata pelajaran Biologi SMA Negeri 1 Pegndon Tahun Pelajaran 2014/2015. Jadi penelitian-penelitian yang ada tersebut hanya dijadikan gambaran dan referensi saja oleh peneliti.

C. Kerangka Berpikir

Soal tes Ujian Akhir Semester (UAS) di Kabupaten Kendal yang dibuat oleh tim Musyawarah Guru Mata Pelajaran (MGMP) Dinas Pendidikan Kabupaten Kendal bertujuan untuk mengetahui seberapa baik siswa menguasai seluruh materi pelajaran yang telah disampaikan. Agar suatu instrumen tes dapat dikatakan mempunyai kualitas yang baik, harus mempunyai beberapa kriteria yaitu valid, reliabel, praktis. Kualitas instrumen tes juga harus memiliki kualitas butir soal yang baik. Kualitas butir soal di pengaruhi oleh *difficulty index* (tingkat kesukaran), *discriminating power* (daya pembeda), analisis pengecoh, homogenitas soal, dan eektivitas fungsi opsi. Untuk mengetahui kualitas soal tes sangat perlu diadakannya analisis terhadap butir soal setelah tes tersebut dikerjakan oleh siswa (*testee*). Hasil analisis terhadap butir soal tes akan memberikan informasi tentang kualitas validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh dari tiap butir soal yang dibuat oleh Dinas Pendidikan Kabupaten Kendal. Setelah diketahui kualitas

setiap butir soal maka selanjutnya akan diketahui pula pada butir mana yang termasuk dalam soal baik, soal kurang baik dan soal yang tidak baik. Hasil analisis tersebut kemudian diinformasikan kepada guru mata pelajaran terkait untuk dijadikan pertimbangan dalam pembuatan soal UAS pada semester selanjutnya. Selanjutnya kerangka berfikir divisualisasikan dengan bagan pada gambar 2.1.

Gambar 2.1
Bagan alur analisis butir soal tes mata pelajaran biologi hasil
MGMP Biologi Kabupaten Kendal dalam ujian akhir semester
gasal kelas XI tahun pelajaran 2015/2016

Soal tes Ujian Akhir Semester (UAS) yang dibuat oleh tim Musyawarah Guru Mata Pelajaran (MGMP) Dinas Pendidikan Kabupaten Kendal Semester Gasal

