

**STUDI DESKRIPTIF ANALISIS INSTRUMEN TES PILIHAN
GANDA MATA PELAJARAN AKIDAH AKHLAK BUATAN
GURU KKM RAYON 2 SEMESTER I SISWA KELAS VII MTs
MIFTAHUSSALAM 1 WONOSALAM DEMAK
TAHUN 2008/2009**

SKRIPSI

Diajukan untuk Memenuhi Tugas dan Melengkapi Syarat
Memperoleh Gelar Sarjana Pendidikan Islam
Ilmu Pendidikan Agama Islam



Oleh:

MAKRUF

NIM : 3102310

**FAKULTAS TARBIYAH
INSTITUT AGAMA ISLAM NEGERI WALISONGO
SEMARANG
2009**

PERSETUJUAN PEMBIMBING

Tanggal

Tanda Tangan

Drs. Karnadi, M.Pd.
Pembimbing I

Dra. Muntholi'ah, M.Pd.
Pembimbing II

TANDA TERIMA NASKAH SKRIPSI

Nama : MAKRUF
NIM : 3102310
Judul Sekripsi : STUDI DESKRIPTIF ANALISIS INSTRUMEN TES PILIHAN GANDA MATA PELAJARAN AKIDAH AKHLAK BUATAN GURU KKM RAYON 2 SEMESTER I SISWA KELAS VII MTs MIFTAHUSSALAM 1 WONOSALAM DEMAK TAHUN 2008/2009.

Tanggal

Tandatangan

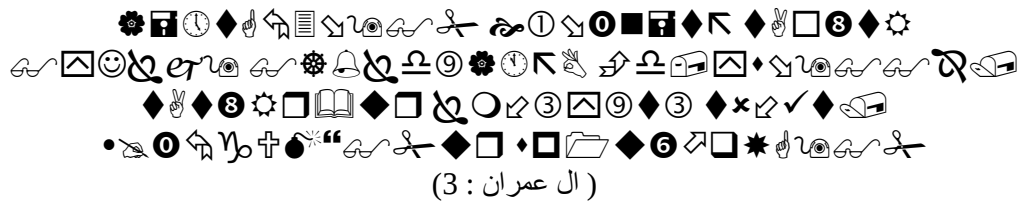
Drs.Karnadi, M.Pd.
Pembimbing I

Dra. Muntholi'ah, M.Pd.
Pembimbing II

Dr. Raharjo, M.Ed.St.
Penguji I

Dr. Hj. Sukasih, M.Pd.
Penguji II

MOTTO



Artinya:

“Dia menurunkan Al kitab (Al Quran) kepadamu dengan sebenarnya; membenarkan kitab yang telah diturunkan sebelumnya dan menurunkan Taurat dan Injil”

Kupersembahkan kepada:

Kedua Orang Tuaku (Bapak Mukarjo, dan Ibu Watriah) Saudara (mbak Mukhasanah, Kakak subakir) dan juga kekasihku (Rini), Keponakan – keponakanku, Sahabat – Sahabatku serta semua pihak yang memberikan motivasi, semoga apa yang dikerjakan selalu mendapat ridho dari allah SWT dan senantiasa menjadi menjadi manusia yang shalih dan mulia.

Amin, Ya Rabbal ‘Alamin.

DEKLARASI

Peneliti menyatakan dengan penuh kejujuran dan tanggung jawab, bahwa skripsi ini tidak berisi materi yang telah pernah ditulis oleh orang lain atau diterbitkan. Demikian juga skripsi ini tidak berisi satupun pikiran – pikiran orang lain kecuali informasi yang terdapat dalam referensi yang dijadikan bahan rujukan.

Semarang, 15 Juni 2009

Makruf
NIM : 3102310

ABSTRAK

Makruf (NIM: 3102310). *Studi deskriptif Analisis Instrumen Tes Pilihan Ganda Mata Pelajaran Akidah Akhlak Buatan Guru KKM Roayon 2 Semester 1 Siswa Kelas VII MTs Miftahussalam 1 Wonosalam Demak Tahun 2008/2009.* Skripsi. Semarang : Program Strata I Jurusan Pendidikan Agama Islam Fakultas Tarbiyah IAIN Walisongo Semarang, 2009.

Penelitian ini bertujuan untuk mengetahui : validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan fungsi distraktor butir tes mata pelajaran Akidah Akhlak buatan Guru KKM Rayon 2 semester I siswa kelas VII MTs Miftahussalam Wonosalam Demak tahun 2008/2009.

Subjek dalam penelitian ini ialah soal tes mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 kelas VII semester I yang telah dikerjakan oleh seluruh siswa kelas VII MTs Miftahussalam Wonosalam Demak berjumlah 120 siswa.

Data penelitian yang telah diperoleh, dianalisis menggunakan teknik analisis deskriptif kuantitatif. Pengujian penelitian menggunakan poin biserial (r_{pbis}) untuk pengujian validitas butir tes, dan formula Kuder-Richardson (K-R20) untuk pengujian reliabilitas tes.

Hasil analisis penelitian menunjukkan bahwa: *pertama*, untuk validitas butir tes mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 semester I siswa kelas VII, ditunjukkan dengan 25 butir soal atau sekitar 62,5% butir-butir soal pada tes tersebut yang dinyatakan memiliki validitas, sedangkan 15 butir soal lainnya dinyatakan invalid atau sekitar 37,5%, dengan demikian tes hasil belajar bentuk objektif yang menyajikan 40 butir soal memiliki validitas sedang. Pada uji reliabilitas tes koefisien reliabilitas (r_{11}) sebesar 0,717 menunjukkan bahwa tes ini memiliki reliabilitas yang tinggi. Dilihat dari segi tingkat kesukaran terdapat 15 atau 37,5% butir tes yang terlalu mudah dikerjakan, 21 atau 52,5% butir tergolong butir tes yang sedang atau cukup dan hanya 4 atau 10% butir tes yang tergolong butir tes yang terlalu sulit sulit. Dilihat dari daya pembeda butir menunjukkan presentase 10% atau 4 butir tes mempunyai daya pembeda yang lemah sekali, ada juga butir soal tes yang mempunyai daya pembeda yang lemah yaitu yaitu 12 butir atau 30%, terdapat juga soal-soal tes tersebut yang sedang daya pembedanya yaitu sebanyak 12 butir tes atau 30%, baik sekitar 10 butir atau 25%, dan baik sekali 2 butir atau 5%. Dilihat dari fungsi distraktor menunjukkan bahwa sekitar 74,166% berfungsi dengan baik.

Berdasarkan penelitian diharapkan dapat dijadikan bahan informasi dan masukan bagi para guru khususnya evaluator bidang pendidikan dan sivitas akademika, para pembuat soal dan para pakar pendidikan agar senantiasa meningkatkan profesionalismenya melaksanakan evaluasi pendidikan agar betul – betul dapat memberikan masukan yang dibutuhkan secara akurat, demi terciptanya evaluasi pendidikan pada masa yang akan datang.

KATA PENGANTAR

Bismillaahirrahmaanirrahiim

Dengan mengucap rasa syukur dan bersujud atas kehadiran Allah SWT atas rahmat, ni'mat, taufiq serta hidayah-Nya penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Analisis Instrumen Tes Mata Pelajaran Aqidah Akhlak Buatan Guru KKM Rayon 2 Semester I Siswa Kelas VII MTs Miftahussalam 1 Wonosalam Demak Tahun 2008/2009”. Skripsi ini disusun untuk memenuhi salah satu syarat mengikuti ujian Munaqosah yang selanjutnya akan memperoleh gelar Sarjana Strata-1 (S.1) Institut Agama Islam Negeri (IAIN) Walisongo Semarang.

Diwaktu menyusun skripsi ini penulis banyak mendapatkan bimbingan dan saran – saran dari berbagai pihak, sehingga dapat terealisasi. Untuk itu penulis menyampaikan terima kasih kepada:

1. Prof. DR. H. Ibnu Hajar, M.Ed., selaku Dekan Fakultas Tarbiyah IAIN Walisongo.
2. Achmad Muthohar, M.Ag., Ketua Juran Pendidikan Agama Islam (PAI) Fakultas Tarbiyah dan segenap staff yang telah memberikan arahan tentang skripsi ini. Tidak lupa kepada Amin Amin Farikh, M.Ag., sebagai Dosen Wali yang selalu membimbing mulai semester I.
3. Drs. Karnadi, M.Pd. selaku Pembimbing I dan Dra. Muntholi'ah, M.Pd. selaku pembimbing II yang telah berkenan meluangkan waktu, tenaga dan pikirannya untuk penyusunan skripsi ini ditengah kesibukan mereka masing – masing.
4. Para Dosen atau Staff Pengajar di lingkungan Fakultas Tarbiyah IAIN Walisongo Semarang yang telah membekali berbagai pengetahuan.
5. Najmudin, S. Ag. selaku Kepala Sekolah MTs Miftahussalam Wonosalam Demak beserta Staff dan Guru – Guru, yang telah berkenan memberikan tempat dan waktu guna penyusunan skripsi.

6. Petugas Perpustakaan, baik Fakultas, Institut serta Pascasarjan IAIN Walisongo yang telah memberikan izin dan layanan yang ramah.
7. Kedua Orang Tuaku (Bapak Mukarjo dan Ibu Watriah), Mbah kirah, Mbak Muhasanah, Kakak Subakir, Calon Istriku Rini, dan dan semua saudaraku yang secara langsung maupun tidak langsung telah membantu baik moril maupun materiil dalam penyusunan skripsi ini.
8. Teman – temanku semua yang telah membantu dalam penyusunan skripsi ini.

Akhirnya penulis menyadari bahwa penyusunan skripsi ini masih jauh dari sempurna. Penulis berharap semoga skripsi ini dapat bermanfaat bagi penulis sendiri khususnya dan bagi para pembaca pada umumnya.

Semarang, 15 Juni 2009
Penulis

Makruf
NIM : 3102310

DAFTAR ISI

Halaman Judul.....	i
Persetujuan Pembimbing.....	ii
Pengesahan.....	iii
Deklarasi	iv
Abstrak	v
Motto	vi
Persembahan	vii
Kata Pengantar	viii
Daftar Isi	ix
Daftar Tabel	x
Daftar Lampiran.....	xi
Bab I PENDAHULUAN.....	1
A. Latar Belakang Masalah.....	1
B. Penegasan Istilah.....	3
C. Perumusan Masalah.....	5
D. Tujuan Penelitian.....	6
E. Manfaat Penelitian.....	6
F. Sistematika Penulisan Sekripsi	8
Bab II KUALITAS INSTRUMEN TES HASIL BELAJAR.....	12
A. Instrumen Tes.....	12
1. Beberapa Istilah yang Terkait	12
2. Pengertian Tes.....	13
3. Perencanaan dalam Penyusunan Tes.....	15
4. Langkah – Langkah dalam Penyusunan Tes.....	16
5. Macam – Macam Tes.....	17
6. Bentuk – Bentuk Tes.....	19
7. Kriteria Tes yang Baik	22

8. Kegunaan Tes.....	23
B. Validitas Tes.....	25
1. Pengertian Validitas	25
2. Macam – Macam Validitas	26
a. Validitas Tes.....	26
b. Validitas Butir	29
C. Teknik Pengujian Validitas Butir	31
D. Reliabelitas Tes	32
a. Pengertian Reliabilitas.....	32
b. Faktor- faktor yang Mempengaruhi Reliabilitas	33
c. Teknik Pengujian Peliabilitas.....	34
E. Tingkat Kesukaran	37
a. Pengertian Tingkat Kesukaran	37
b. Menentukan Proporsi Jumlah soal Kategori Mudah, Sedang, dan Sukar.....	38
F. Daya Pembeda	38
a. Pengetian Daya Pembeda	38
b. Langkah – langkah Menghitung Daya Pembeda.....	39
G. Efektifitas Fungsi Distraktor	39
a. Pengertian Fungsi Distraktor.....	39
b. Tujuan Pemakaian Fungsi Distraktor	40
H. Kajian Penelitian yang Relevan	40
 Bab III	
METODE PENELITIAN	43
A. Waktu dan Tempat Penelitian	43
B. Variabel Penelitian	43
C. Metode Penelitian.....	43
D. Subyek Penelitian.....	44

	E. Teknik Pengumpulan Data.....	44
	F. Teknik Analisa Data.....	45
Bab IV	HASIL PENELITIAN DAN PEMBAHASAN	46
	A. Diskripsi Tentang Tes Aqidah Akhlak.....	46
	B. Analisis Validitas, Reliabilitas, Tingkat Kesukaran, Daya Pembeda, dan Fungsi Distraktor	47
	1. Analisis Validitas Butir	47
	2. Analisis Reliabilitas.....	52
	3. Analisis Tingkat Kesukaran	54
	4. Analisis Daya Pembeda.....	55
	5. Analisis Fungsi Distraktor.....	55
		
	C. Hasil dan Pembahasan Penelitian.....	56
	D. Keterbatasan Penelitian	58
Bab V	SIMPULAN, SARAN DAN PENUTUP	60
	A. Simpulan.....	60
	B. Saran- Saran	63
	C. Penutup.....	64

Daftar Pustaka

Lampiran – Lampiran

Daftar Riwayat Pendidikan Penulis

DAFTAR TABEL

Tabel 1	: Tabel Mencari M _p Butir
Tabel 2	: Tabel analisis Validitas Butir
Tabel 3	: Derajat Kesukaran Butir Soal Mata Pelajaran Akidah Akhlak Buatan Guru KKM Rayon 2 Semester I Kelas VII MTs Miftahussalam Wonosalam Demak
Tabel 4	: Daya Pembeda Butir Tes Objektif Mata Pelajaran Akidah Akhlak Buatan Guru KKM Rayon 2 Semester I Kelas VII MTs Miftahussalam Wonosalam Demak
Tabel 5	: Fungsi Distraktor Instrumen Tes Mata Pelajaran Akidah Akhlak Buatan Guru KKM Rayon 2 Semester I Kelas VII MTs Miftahussalam Wonosalam Demak

BAB I

PENDAHULUAN

A. LATAR BELAKANG PERMASALAHAN

Salah satu cara untuk memperbaiki proses pembelajaran yang paling efektif ialah dengan jalan mengevaluasi tes tes hasil belajar yang di peroleh dari proses belajar – mengajar itu sendiri.

Hasil dari suatu evaluasi nantinya dapat digunakan sebagai tolok ukur atau *feed back* dari peserta didik tentang apa yang harus dibenahi atau hal apa yang harus senantiasa ditambahkan. Evaluasi formal sebenarnya sangat memegang peranan penting di dunia pendidikan, antara lain memberi informasi yang dipakai sebagai dasar untuk:¹

1. Membuat kebijaksanaan dan keputusan.
2. Menilai hasil para siswa.
3. Menilai kurikulum.
4. Memberi kepercayaan kepada sekolah.
5. Memonitor dana yang telah diberikan.
6. Memperbaiki materi dan program pendidikan.

Agar tujuan evaluasi atau penilaian dapat terwujud sesuai degan prinsip-prinsip yang mendasari serta syarat - syarat yang diperlukan, pelaksanaannya perlu menyesuaikan langkah - langkah atau prosedurnya dengan menggunakan teknik yang cocok menurut jenis yang diperlukan. Jenis-jenis evaluasi dapat dilakukan dengan menggunakan beberapa teknik, yaitu teknik tes dan non tes, kemudian ditutup dengan langkah - langkah yang diambil pada pelaksanaannya.²

¹ Blaine R. Worthen and James R. Sanders, *Educational Evaluation*, (New york: Longman, 1980), hlm. 5.

² Slameto, *Evaluasi Pendidikan*, (Jakarta: Bumi Aksara, 2001), hlm. 25.

Proses evaluasi pendidikan diperlukan adanya suatu teknik, prosedur yang baik dan diharapkan akan dapat memberikan informasi yang betul-betul valid.

Pengukuran dan penilaian dalam pendidikan merupakan proses dan aktivitas yang kompleks. Disamping membutuhkan prosedur yang rumit, kemampuan guru atau evaluator yang memadai, yang tidak kalah penting yaitu penggunaan instrumen yang berkualitas. Instrumen yang berkualitas baik, memungkinkan diperolehnya data dan informasi yang akurat, sehingga keputusan-keputusan penilaian dan pendidikan dapat dibuat dengan tepat. Tanpa instrumen yang baik mustahil dapat dilakukan pengukuran dan penilaian pendidikan secara efektif.

Tes sebernarnya ialah salah satu wahana program penilaian pendidikan. Sebagai salah satu alat penilaian, tes biasanya didefinisikan sebagai kumpulan butir-butir soal yang jawabannya dapat dinyatakan dengan benar atau salah.³ Definisi ini biasanya dipakai untuk membedakan alat atau teknik lain yang mungkin dipakai dalam penilaian hasil belajar.

Pada kenyataannya, sering kali kita dengar di telinga kita siswa siswa yang mengeluhkan terhadap tes yang telah diujikan tersebut. Sebagian dari siswa menganggap bahwa bahan tes yang telah diujikan terlalu sulit dikerjakan, sebagian yang lain menganggap bahwa tes yang telah diujikan terlalu mudah dikerjakan, sebagian yang lain menganggap bahwa bahan atau instrumen tes yang diajukan mempunyai kesulitan yang cukup sedang atau biasa-biasa saja. Hal yang demikian karena pimpinan sekolah dan guru – guru kurang menyadari atas kekurangan serta kelemahan – kelemahan yang terdapat pada saat menyelenggarakan evaluasi, sehingga hal semacam itu tidak terjadi pada satu mata pelajaran semata, akan tetapi hal semacam ini terjadi hampir di semua mata pelajaran yang ada atau yang diujikan.

³ Mudjijo, *Tes Hasil Belajar*, (Jakarta: Bumi Aksara, 1995), hlm. 1.

Hal yang perlu diketahui oleh seorang evaluator untuk membuat suatu instrumen; khususnya tes, suatu instrumen tes dapat dikatakan baik bilamana instrumen tersebut memiliki ciri sebagai alat ukur yang baik. Kriteria tersebut antara lain:

1. Memiliki validitas yang cukup tinggi.
2. Memiliki reliabilitas yang baik .
3. Memiliki nilai kepraktisan.⁴

Untuk menghasilkan tes sebagai alat ukur yang baik, pastinya membutuhkan evaluator yang berkompeten, sehingga dapat mencegah terjadinya pengambilan keputusan secara sekehendak hati atau gegabah. Dengan demikian alat ukur berupa tes tersebut dapat membedakan kemampuan antar individu.

Berpangkal dari hal tersebut, penulis tertarik untuk mengkaji lebih lanjut untuk mengembangkan dengan analisa pengujian instrumen tes yang lebih mendalam dalam sebuah skripsi yang berjudul:

“STUDI DESKRIPTIF ANALISIS INSTRUMEN TES PILIHAN GANDA MATA PELAJARAN AKIDAH AKHLAK BUATAN GURU KKM RAYON 2 SEMESTER I SISWA KELAS VII MTs MIFTAHUSSALAM 1 WONOSALAM DEMAK TAHUN 2008/2009.”

B. PENEGASAN ISTILAH

1. Analisis

Analisis berarti : Penguraian suatu pokok atas berbagai bagiannya dan penelaahan bagian itu sendiri serta hubungan antar bagian untuk memperoleh pengertian yang tepat dan pemahaman arti keseluruhan⁵.

⁴ Chabib Thoha, *Teknik Evaluasi Pendidikan*, (Jakarta: Rajawali, 1991), hlm 109.

⁵Sam Chan, "Analisis Swot Kebijakan Era Pendidikan", <http://blog.unila.ac.id/redho/2009/03/26/analisa-metodologi-penelitian-sosial>, hlm. 1.

2. Tes

Tes ialah serentetan pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, intelegensi, kemasmpuan atau bakat lain yang dimiliki oleh individu atau kelompok.⁶ Tes sebenarnya sebagai salah satu wahana program penilaian pendidikan, yang menekankan pada alat pengukuran. Sedangkan Sumadi Suryabrata Tes berarti suatu cara untuk mengadakan penilaian yang berbentuk suatu tugas atau serangkaian tugas yang harus dikerjakan oleh seorang anak atau sekelompok anak sehingga menghasilkan suatu nilai tentang tingkah laku atau perilaku anak tersebut, yang dapat dibandingkan dengan nilai yang dicapai oleh anak-anak lain atau dengan nilai standar yang ditetapkan.⁷

Apabila definisi tadi kita analisa, maka akan kita temuka unsur-unsur sebagai berikut:

- a. Bahwa tes itu berbentuk suatu tugas yang terdiri dari pertanyaan-pertanyaan atau perintah-perintah.
- b. Bahwa tes itu diberikan kepada seseorang anak atau kelompok uantuk dikerjakan.
- c. Bahwa respon anak atau kelompok tersebut dinilai.

Item pilihan ganda yaitu suatu item yang terdiri dari suatu statemen yang belum lengkap. Untuk melengkapi statemen tersebut diadakan statemen sambungan. Satu diantaranya adalah sambungan yang benar, sedangkan yang lain adalah sambungan yang tidak benar. Murid-murid disuruh memilih manakah sambungan yang benar untuk statemen yang belum lengkap itu. Nomor sambungan yang benar pada lembar jawaban supaya di isi dengan tanda silang, tanda lingkaran, atau tanda yang lain sesuai dengan petunjuk tes bersangkutan. Item pilihan ganda ini dapat pula berupa suatu petunjuk yang telah disediakan beberapa jawaban yang hanya ada satu dari jawaban-jawaban

⁶ Suharsimi Arikunto, *Prosedur Penelitian*, (Jakarta: Rineka Cipta, 1993), hlm. 123

⁷ Wayan Nurkancana, *Evaluasi Pendidikan*, (Jakarta: Usah Nasional, 1986), hlm. 140.

yang benar. Pertanyaan atau statemen yang belum lengkap dalam item pilihan ganda disebut *stem*. Alternatif pilihan disebut *option*. Jawaban-jawaban yang tidak benar disebut pengecoh atau *distraktor*.⁸

3. Akidah Akhlak.

Akidah Akhlak merupakan salah satu bagian mata pelajaran Pendidikan Agama Islam yang ada di MTs Miftahussam 1 yang mengajarkan tentang keyakinan (keiman) dan budi pekerti (Akhlak).

C. RUMUSAN MASALAH

Berdasarkan latar belakang masalah tersebut, maka dapat dirumuskan masalah yang akan dibahas pada penelitian ini yaitu:

1. Bagaimanakah validitas tes soal ujian semester I mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak?
2. Bagaimanakah Reliabilitas tes soal ujian semester I mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 buatan guru KKM rayon 2 siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak?
3. Bagaimanakah tingkat kesukaran soal ujian semester I mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak?
4. Bagaimanakah dengan kualitas daya pembeda instrumen tes tersebut?
5. Bagaimanakah dengan kualitas fungsi distraktor atau fungsi pengecoh instrumen tes tersebut?

⁸ *Ibid*, hlm. 144.

D. TUJUAN PENELITIAN

Sesuai dengan perumusan masalah seperti tersebut, maka peneliti ini bertujuan untuk mengetahui:

1. Validitas tes soal ujian semester I mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak.
2. Reliabilitas tes soal ujian semester I mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak.
3. Tingkat kesukaran instrumen tes soal ujian semester I mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak.
4. Daya pembeda instrumen tes soal ujian semester I mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak.
5. Fungsi distraktor butir tes soal ujian semester I mata pelajaran Akidah Akhlak buatan guru KKM rayon 2 buatan guru KKM rayon 2 siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak.

E. MANFAAT PENELITIAN

Manfaat yang dapat diperoleh melalui penelitian ini adalah meliputi beberapa hal, diantaranya untuk:

1. Dipakai sebagai bahan pertimbangan dalam rangka merancang tes yang akan datang dengan usaha berupa perbaikan dan penyesuaian serta penyempurnaan dalam bagian tertentu,
2. Melengkapi bahan-bahan informasi terhadap lembaga-lembaga yang lain, baik yang formal maupun non formal yang membutuhkan gambaran tentang tes pada bidang studi mata pelajaran Akidah Akhlak,
3. Relevansi antara alat ukur dalam penilaian yang telah dirancang, dengan hasil yang telah dan akan dicapai untuk tahun yang akan datang.

Secara lebih rinci, manfaat penelitian ini dapat disepesifikasikan kadalam beberapa kelompok atau instansi sesuai dengan keperluan masing-masing :

a. Manfaat bagi sekolah

manfaat penelitian ini bagi sekolah diantaranya adalah membantu untuk mengetahui seberapa besar kemajuan yang telah dicapai secara keseluruhan oleh semua siswa yang telah mengikuti tes mata pelajaran akidah akhlak pada semester I tahun 2008/2009. dengan adanya penelitian terhadap instrumen tes pada mata pelajaran aqidah akhlak, sebuah lembaga pendidikan akan dapat mendapatkan informasi yang cukup berguna dalm rangka mengukur keberhasilan sekolah dalam melaksanakan proses pendidikan. Selain itu, manfaat lain yang didapat dengan adanya penelitian ini yaitu instansi atau lembaga pendidikan yang terkait (sekolah) akan mengetahui ranking atau peringkat dari sekolah tersebut bila dibandingkan dengan sekolah yang lain.

b. Kepala sekolah

Penelitian ini sangat menunjang terhadap tercapainya program pendidikan di sekolah. Sebagai langkah selanjutnya, kepala sekolah akan mengevaluasi terhadap hasil yang dicapai oleh siswanya setelah mengikuti tes. Bila hasil dari tes tersebut memuaskan, dalam arti sesuai standar yang telah ditetapkan oleh sekolah tersebut, maka pendidikan yang telah dilakukan dianggap telah berhasil dengan baik. Namun manakala hasil dari tees tersebut kurang baik, dalam arti masih jauh dari apa yang telah ditetapkan, maka hal itu menjadi tanggung jawab kepala sekolah sebagai pemimpin tertinggi di lembaga tersebut. Hal ini juga akan membuat kepala sekolah bekerja lebih ekstra keras mengadakan pengawasan terhadap guru yang mengajar pada mata pelajarran yang bersangkutan dan juga dapat meminta pertanggungjawaban terhadap guru mata pelajaran yang bersangkutan.

c. Guru

Setelah mengadakan evaluasi (evaluasi diartikan mengadakan penilaian dari jawaban siswa), seorang guru akan dapat mengetahui bahwa terdapat siswa yang mampu menguasai materi yang disampaikan dengan baik atau bahkan sebaliknya. Melihat realita semacam ini guru akan memberikan remedi (perbaikan dengan mengadakan tes ulang) terhadap siswa yang kurang dapat menjawab dengan baik terhadap soal-soal yang diujikan.

d. Pembuat soal

Manfaat yang dapat diperoleh oleh pembuat soal ujian tes mata pelajaran Akidah Akhlak kelas VII tahun 2008/2009 ini ialah mereka akan mendapat informasi tentang kualitas tes yang telah dibuatnya. Jika hasil penelitian nantinya memberikan informasi bahwa dalam instrumen tersebut terdapat butir yang mempunyai kualitas yang kurang bagus, mereka tidak akan banyank mengeluarkan tenaga untuk mencari kelemahan-kelemahan atau kekurangan terhadap tes yang mereka buat karena hal itu telah diselidiki lebih lanjut oleh penulis. Jadi dengan adanya penelitian terhadap tes mata pelajaran Akidah Akhlak mereka tinggal membenahi atau memperbaiki kekurangan-kekurangan pada butir tes itu saja.

Ke depan dapat dikatakan bahwa ketika apeneliti disini mengadakan suatu penelitian terhadap instrumen tes tersebut, maka hal positif yang akan didapat oleh para pembuat soal yaitu mereka akan mendapat informasi atau gambaran tentang instrumen yang telah dibuat oleh para pembuat soal.

F. SISTEMATIKA PENULISAN

Sistematika penulisan dalam skripsi ini dibagi lima bab, supaya lebih rinci. Penulis uraikan isi kandungan tulisan ini dengan masing-masing bab, yaitu :

Bab I : PENDAHULUAN

A. Latar Belakang Masalah

B. Penegasan Istilah

- C. Perumusan Masalah
- D. Tujuan Penelitian
- E. Manfaat Penelitian
- F. Sistematika Penulisan Sekripsi

Bab II : KUALITAS INSTRUMEN TES HASIL BELAJAR

A. Instrumen Tes.

1. Beberapa Istilah yang Terkait
2. Pengertian Tes
3. Perencanaan dalam Penyusunan Tes
4. Langkah – Langkah dalam Penyusunan Tes
5. Macam – Macam Tes
6. Bentuk – Bentuk Tes
7. Kriteria Tes yang Baik
8. Kegunaan Tes

B. Validitas Tes

1. Pengertian Validitas
2. Macam – Macam Validitas
 - a. Validitas Tes
 - b. Validitas Butir
3. Teknik Pengujian Validitas Butir

C. Reliabelitas Tes

- a. Pengertian Reliabilitas
- b. Faktor- faktor yang Mempengaruhi Reliabilitas
- c. Teknik Pengujian Reliabilitas

D. Tingkat Kesukaran

- a. Pengertian Analisis Tingkat Kesukaran
- b. Menentukan Proporsi Jumlah Soal Kategori Mudah, Sedang, dan Sukar.

- E. Daya Pembeda
 - a. Pengetian Daya Pembeda
 - b. Langkah – langkah Menghitung Daya Pembeda
- F. Efektifitas Fungsi Distraktor
 - a. Pengertian Fungsi Distraktor
 - b. Tujuan Pemakaian Fungsi Distraktor
- G. Kajian Penelitian yang Relevan

Bab III : METODE PENELITIAN

- A. Waktu dan Tempat Penelitian
- B. Variabel Penelitian
- C. Metode Penelitian
- D. Populasi Sampel dan Teknik Pengambilan Sampel
- E. Teknik Pengumpulan Data
- F. Teknik Analisa Data

Bab IV : HASIL PENELITIAN DAN PEMBAHASAN

- A. Deskripsi Tentang Tes Aqidah Akhlak
- B. Analisis Validitas, Reliabilitas, Tingkat Kesukaran, Daya Pembeda, dan Fungsi Distraktor.
 - 1. Analisis Validitas Butir
 - 2. Analisis Reliabilitas
 - 3. Analisis Tingkat Kesukaran
 - 4. Analisis Daya Pembeda
 - 5. Analisis Efektifitas Fungsi Distraktor
- C. Pembahasan dan Hasil Penelitian
- D. Keterbatasan Penelitian

Bab V: SIMPULAN, SARAN DAN PENUTUP

A. Simpulan

B. Saran- Saran

C. Penutup

BAB II

KUALITAS INSTRUMEN TES HASIL BELAJAR

A. Instrumen Tes

1. Beberapa Istilah yang Terkait.

Untuk mengetahui siswa mana yang lebih pandai dari siswa yang lain, perlu diadakan evaluasi. Salah satu alat evaluasi pendidikan ialah tes. Sebelum sampai pada uraian yang lebih jauh, maka akan diterangkan dulu arti dari beberapa istilah – istilah yang berhubungan dengan tes.

a. Tes

Tes ialah merupakan alat prosedur yang digunakan untuk mengetahui atau mengukur sesuatu dalam suasana, dengan cara dan aturan-aturan yang telah ditentukan.¹ Untuk mengerjakan tes ini tergantung dari petunjuk yang diberikan misalnya: melingkari salah satu huruf di depan pilihan jawaban, menerangkan, mencoret jawaban yang salah, menggaris bawahi jawaban yang benar, melingkari dari simbol jawaban yang benar, menjawab secara lisan, dan sebagainya.

b. Testing

Testing merupakan saat pada waktu tes itu sedang dilakdanakan atau berlangsung. Testing juga dapat dikatakan saat pengambilan tes.²

c. Testee

Testee adalah responden yang sedang mengerjakan tes. Jadi orang-orang inilah yang dinilai, diukur baik kemampuan, minat, ataupun pencapaian prestasi.³

¹ Suharsimi Arikunto, *Dasar-Dasar Evalusi Pendidikan*, (Jakarta: Bumi Aksara, 1990), hlm.

² *Ibid.*

³ *Ibid.*

d. Tester

(Dalam istilah Indonesia: pencoba), adalah orang yang disertai untuk melaksanakan pengambilan tes terhadap para responden atau testee.⁴

2. Pengertian Tes

Tes berasal dari bahasa Latin *Testum* yang berarti alat untuk mengukur tanah. Dalam bahasa Perancis kuno, kata tes berarti ukuran yang dipergunakan untuk membedakan antara emas dengan perak serta logam lainnya.⁵

Dalam kamus besar Bahasa Indonesia, tes adalah ujian secara tertulis, lisan, atau wawancara untuk mengetahui pengetahuan, kemampuan, bakat, dan kepribadian seseorang individu.⁶

Pendapat lain yang mengartikan tes, tes yaitu serentetan pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, intelegensi, kemampuan atau bakat lain yang dimiliki individu atau kelompok.⁷

Menurut Lee J. Cronbach, dalam buku berjudul "*Essentials of Psychological Testing*" mendefinisikan tes sebagai berikut:

"A test is a systematic procedure for comparing the behavior of two or more persons"

Artinya: Tes adalah prosedur sistematis untuk membedakan perilaku antara dua orang atau lebih.⁸

Menurut Douglas Brown, dalam buku kitab *Ususu At-Ta'alimu Al-Lughoti wa Ta'liimiha*, terjemah oleh Abdul Ar-Rojhi, tes berarti:

⁴ Ibid.

⁵ Chabib Thoh, *Teknik evaluasi Pendidikan*, (Jakarta: Rajawali, 1991), hlm. 43

⁶ Tim Penyusun Kamus Pembinaan dan Pengembangan Bahasa, *Kamus Besar Bahasa Indonesia*, (Jakarta: Balai Pustaka, 1997), Cet. IX, hlm 1050.

⁷ Suharsimi Arikunto, *Prosedur penelitian*, (Jakarta: Rineka Cipta, 1993) hlm. 123.

⁸ Lee J. Cronbach, Lee J. Cronbach, *Essentials of Psychological Testing*, (New York : Harper & Row Publishers, 1960), hlm. 21.

والاختبار : طريقة لقياس الافراد وادومعا رفهم في مجال معين⁹

Artinya: “Tes dalam pengertian yang singkat adalah cara untuk mengukur pengetahuan individu (*testee*) dalam bidang tertentu”.

Beberapa pengertian tes di atas dapat disimpulkan bahwa, Tes adalah merupakan suatu alat pengumpul informasi yang didapat melalui pertanyaan, petunjuk, latihan, perintah kepada *testee* untuk merespon sesuai dengan petunjuk prosedur itu, kemudian hasilnya oleh tester diolah secara sistematis menuju arah kesimpulan yang menggambarkan tingkah laku subyek tersebut.

Dari beberapa bahasan mengenai istilah tersebut diatas, dapatlah ditarik kesimpulan pengertian, antara lain:

- a. *Teste* yaitu prosedur yang sistematis, artinya (1) aitem-aitem dalam tes disusun menurut cara dan aturan tertentu, (2) prosedur administrasi dan pemberian angka (*scoring*) tes harus jelas dan dispesifikasikan secara terperinci, dan (3) setiap orang yang mengambil tes itu harus mendapat item-item ada dalam kondisi sebanding.
- b. Tes berisi sampel perilaku, artinya (1) betapapun panjangnya suatu tes isi yang tercakup di dalamnya tidak akan lebih dari seluruh item yang mungkin ada, dan (2) kelayakan tes tergantung pada sejauh mana item-item di tes itu dapat mewakili secara representatif kawasan (*domain*) perilaku yang diukur.
- c. Tes mengukur perilaku, artinya item-item dalam tes menghendaki subjek agar menunjukkan apa yang diketahui dan atau apa yang telah dipelajari subjek dengan cara menjawab item-item atau mengerjakan tugas yang dikehendaki oleh tes.

⁹ Douglas Brown, *Ususu At-Ta'alimu Al-Lughoti wa Ta'liimiha*, Terj. Abdul Ar-Rojhi dan Ali Ahmad Syu'ban, (Arab : Darun An-Nahdhoh), hlm. 266.

Adapun beberapa hal yang tidak tercakup dalam batasan tes adalah:

1. Tes tidak memberi spesifikasi formatnya, artinya tes dapat disusun dalam berbagai bentuk dan tipe sesuai dengan tujuan dan maksud diadakannya tes.
2. Tes tidak membatasi isi yang dicakupnya, artinya tes dapat melakukan fungsi ukur terhadap hasil belajar, reliabilitas, kemampuan kusus atau bakat, intelegensi, dan sebagainya sesuai dengan maksud tes itu dibuat.
3. Subjek yang dikenai tes tidak selalu perlu dan harus tahu kalau ia sedang dikenai tes. Lebih lanjut, subjek tidak selalu perlu tahu aspek psikologis apa yang sedang diukur dari dirinya. (tes prestasi Saifudin Azwar hlm. 3-4).

3. Perencanaan Dalam Penyusunan Tes

Untuk menyusun sebuah tes, ada beberapa hal yang perlu diperhatikan agar instrumen tersebut dapat menjadi instrumen tes yang baik, diantaranya:¹⁰

- a. Menentukan atau merumuskan tujuan tes.
- b. Mengidentifikasi hasil – hasil belajar (*Learning test*) yang akan diukur dengan tes itu.
- c. Menentukan atau menandai hasil – hasil yang spesifik, yang merupakan tingkah laku yang dapat diamatan sesuai dengan TIK (Tujuan Instruksional Kusus).
- d. Merinci mata pelajaran atau bahan pelajaran yang akan diukur dengan tes itu
- e. Menyiapkan tabel spesifikasi. Menggunakan tabel tersebut sebagai dasar penyusunan tes.

¹⁰ Ngalim Purwanto *Prinsip – Prinsip dan Teknik Evaluasi Pengajaran*, (Bandung : Remaja Rosdakarya, 2002), Cet. II, hlm. 17-19.

4. Langkah – Langkah Penyusunan Tes

Untuk menyusun tes diperlukan langkah – langkah sebagai berikut:¹¹

a. Menyusun *Lay Out*

Menyusun *Lay Out* harus mengacu pada tujuan, perencanaan pembelajaran, buku – buku. Menyusun *Lay Out* juga harus sesuai dengan materi yang telah disampaikan (representatif).

1. Ruang lingkup (scope)

Pengetahuan yang akan diukur sesuai rencana pelajaran yang telah ditetapkan kurikulum atau dalam program evaluasi.

2. Proporsi jumlah Item (Butir tes) disesuaikan luas masing – masing sub materi.

3. Jenis pengetahuan atau aspek proses mental yang hendak diukur sesuai dengan urgensinya.

4. Bentuk atau tipe yang akan digunakan lebih dari satu bentuk.

b. Menulis soal (*Item Writing*)

c. Menata soal

Menata soal terlebih dahulu dikumpulkan butir tes sesuai kelompoknya. Artinya bentuk soal pilihan ganda dikelompokkan dengan pilihan ganda, esai sama esai. Taraf kesulitan tes juga harus dikelompokkan antara yang ringan dengan yang ringan, sedang sama yang sedang dan sukar dengan yang sukar.

d. Menetapkan skor

e. Reproduksi tes

f. Analisa empiris terhadap suatu tes hasil belajar.

¹¹ *Ibid.*, hlm. 16.

5. Macam – Macam Tes

Tes yang merupakan salah satu teknik dalam evaluasi memiliki berbagai macam dan bentuk. Disini akan kami bahas macam-macamnya. Secara umum tes dibedakan :

- a. Berdasarkan Objek Pengukurannya
yaitu terdiri atas tes kepribadian (*personality test*) dan tes hasil belajar (*achievement test*).

- 1) Tes Kepribadian (*personality test*).

Tes kepribadian adalah tes yang diajukan untuk mengukur salah satu atau lebih aspek-aspek non intelek dari susunan mental atau psikologis individu.¹²

Yang termasuk dalam jenis tes ini dan banyak digunakan dalam pendidikan adalah : a) pengukuran sikap, b) pengukuran minat, c) pengukuran bakat, d) tes intelegensi.

- 2) Tes Hasil Belajar (*Achievement test*)

Tes prestasi atau hasil belajar adalah tes yang mengukur tingkat mana seseorang telah mencapai sesuatu dengan mempelajari informasi tertentu atau menguasai kemampuan tertentu biasanya sebagai akibat dari petunjuk / perintah khusus.¹³

Tes prestasi pada umumnya mengukur penguasaan dan kemampuan peserta didik individual dalam cakupan dan ilmu pengetahuan yang telah ditentukan oleh guru yang itu semua dilakukan oleh mereka selama waktu tertentu dan terjadi dalam proses belajar mengajar.¹⁴

¹² Chabib Thoha, *Op. Cit.*, hlm. 44.

¹³ *Ibid.*

¹⁴ Sukardi, *Metodologi Penelitian Pendidikan (Kompetensi dan Praktiknya)*, (Jakarta: Bumi Aksara, 2003), hlm. 139.

b. Berdasarkan Fungsinya

Berdasarkan fungsinya tes dibedakan menjadi empat jenis, yaitu:

1) Tes Penempatan

Tes penempatan ialah tes untuk mengukur kemampuan dasar yang dimiliki oleh anak didik; kemampuan tersebut dapat dipakai untuk meramalkan kemampuan peserta didik pada masa mendatang, sehingga kepadanya dapat dibimbing, diarahkan sesuai dengan kemampuan dasarnya.¹⁵

2) Tes Formatif

Tes formatif yaitu tes yang dilakukan pada saat proses belajar mengajar sedang berlangsung, digunakan untuk mencari umpan balik guna untuk memperbaiki proses belajar mengajar bagi guru dan murid.¹⁶

3) Tes Diagnostik

Tes diagnostik ialah tes yang digunakan untuk mengetahui atau mencari sebab-sebab kegagalan atau kesulitan belajar pada peserta didik.¹⁷

4) Tes Sumatif.

Tes sumatif yakni tes yang dilaksanakan pada akhir proses pembelajaran atau semester digunakan untuk mengukur atau menilai seberapa jauh pencapaian peserta didik terhadap bahan pelajaran yang telah diajarkan, dan selanjutnya untuk menentukan kenaikan tingkat atau kelulusan peserta didik yang bersangkutan.¹⁸

¹⁵ Chabib Thoha dan Abdul Mu'ti, PBM PAI di Sekolah (*Eksistensi dan Proses Belajar mengajar Pendidikan Agama Islam*), (Yogyakarta: Pustaka Pelajar, 1998), hlm. 289.

¹⁶ Harjanto, *Perencanaan Pengajaran*, (Jakarta: Rineka cipta, 1997), hlm. 284.

¹⁷ *Ibid.*

¹⁸ Chabib Thoha, *Op. Cit.*, hlm.48.

c. Berdasarkan Tingkatannya

Berdasarkan tingkatannya, tes dibedakan menjadi dua jenis, yaitu:

1. Tes Standar

Tes standar yaitu tes yang telah mengalami proses standarisasi, yakni proses validasi dan keandalan sehingga tes tersebut benar-benar valid dan andal untuk suatu tujuan bagi kelompok tertentu. Tes standar juga digunakan sebagai alat pengukur untuk membandingkan perorangan atau kelompok siswa yang tidak dapat dilakukan oleh tes buatan guru.¹⁹

2. Tes Non Standar

Tes non standar ialah tes yang dibuat oleh pengajar atau guru yang belum memiliki keahlian profesional dalam mpenyusunan tes, atau mereka yang memiliki keahlian tetapi tidak sempat menyusun tes secara baik, menguji cobakan, melakukan analisis sehingga validitas dan reliabilitasnya belum dapat dipertanggung jawabkan.²⁰

6. Bentuk-Bentuk Tes

Ditinjau dari bentuknya tes dibagi menjadi tiga jenis, yaitu :

a) Tes Tertulis

Tes tertulis yaitu tester mengajukan butir-butir pertanyaan atau soalnya dilakukan secara tertulis dan teste memberikan jawabannya juga secara tertulis.²¹

Tes tertulis secara umum dapat dibedakan menjadi dua bagian, yaitu :

¹⁹ Sri Esti Wuryani Djiwandono, *Psikologi Pendidikan*, (Jakarta : Grasindo, 2002), hlm. 409.

²⁰ Chabib Thoha, *Log. Cit.*, hlm. 52.

²¹ Anas Sudjiono, *Pengantar Evaluasi Pendidikan*, (Jakarta : Raja Grafindo, 2003), hlm. 75

1. Tes Obyektif

Tes obyektif ialah tes yang itemnya dapat dijawab dengan jawaban item yang sudah tersedia; sehingga peserta tes menampilkan keseragaman data, baik bagi yang menjawab benar maupun mereka yang menjawab salah.²²

Bentuk obyektif tes ada beberapa macam, antara lain :

a. *Completion Tes*, terdiri atas :

- 1) *Completion Tes* (tes melengkapi)
- 2) *Fill-in* (mengisi titik-titik dalam kalimat yang dikosongkan)

Completion tes dan *Fill-in* adalah merupakan tes yang menggunakan jawaban pendek dan bebas, butir soalnya merupakan kalimat dimana bagian-bagian tertentu harus diisi dan tidak mengenai satu hal saja melainkan bisa bermacam hal.²³

b. *Selektion Type Tes* (tes yang menjawab dengan mengadakan pilihan) yang terdiri atas :

- 1) *True – false* (benar – salah)

Tes seperti ini terdiri dari kalimat atau pertanyaan yang mengandung dua kemungkinan jawab: benar atau salah, dan teste diminta memilih apakah pernyataan – pernyataan tersebut benar atau salah dengan cara tertentu.²⁴

- 2) *Multiple Coice* (pilihan ganda)

Tes seperti ini yaitu tes obyektif yang terdiri atas pernyataan atau pertanyaan yang sifatnya belum selesai, dan untuk menyelesaikannya harus dipilih salah satu atau lebih dari

²² Chabib Thoha dan Abdul Mu'ti, *Op. Cit.*, hlm. 298.

²³ Ngalim Purwanto, *Loc. Cit.*, hlm. 40.

²⁴ Sri Esti W.D., *Op. Cit.*, hlm. 425.

beberapa kemungkinan jawaban yang telah disediakan pada tiap – tiap butir soalnya.²⁵

3) *Matching* (menjodohkan)

Matching test pada umumnya disampaikan berbentuk dua jalur, lajur kiri dan lajur kanan. Kemudian setiap lajur kiri yang merupakan tempat atau posisi dari soal atau pertanyaan, sedangkan untuk lajur kanan ialah tempat jawaban. Selanjutnya jawaban di isi pada tempat yang telah ada.²⁶

2. Tes Subyektif

Tes subyektif juga sering disebut dengan tes uraian, sehingga pada tes ini peserta didik memiliki kebebasan untuk menentukan jawaban, yang mengakibatkan data jawaban akan bervariasi dan menimbulkan subjektifitas penilaiannya.²⁷

b) Tes Lisan

Tes Lisan ialah tes dimana tester dalam mengajukan pertanyaan – pertanyaan atau soal dilakukan secara lisan, dan teste memberikan jawaban secara lisan pula.²⁸

c) Tes Tindakan

Tes tindakan yaitu tes yang persoalan atau pertanyaan disampaikan dalam bentuk tugas yang harus dikerjakan oleh peserta didik. Alat yang dapat digunakan tes ini adalah berupa observasi atau pengamatan terhadap tingkah laku tersebut, yang hasilnya kemudian diserahkan pada guru.²⁹

²⁵ Anas Sudjiono, *Op. Cit.*, hlm. 118.

²⁶ Sri Esti W.D., *Loc. Cit.*, hlm. 432.

²⁷ CHabib Thoha dan Abdul Mu'ti, *Loc. Cit.*, hlm. 298.

²⁸ Anas Sudjiono, *Loc. Cit.*, hlm. 75.

²⁹ W.S. Winkel S.J., *Psikologi Pendidikan dan Evaluasi Belajar*, (Jakarta : Gramedia, 1983), hlm. 106.

7. Kriteria Tes yang Baik

Suatu tes dapat dikatakan baik bilamana tes tersebut memiliki ciri sebagai alat ukur yang baik. Kriterianya antara lain :

a. Memiliki Validitas (*keshahihan*) yang cukup tinggi

Suatu tes dikatakan valid atau shahih jika tes tersebut mengukur tujuan atau salah satu aspek tujuan yang peneliti ukur. Salah satu metode penentuan kevalidan tes prestasi yaitu mempelajari isi tes.³⁰ Untuk penjelasan akan diterangkan penulis pada sub berikutnya.

b. Memiliki Realibitas (*keajegan / kestabilan*) yang baik

Tes dikatakan reliabel jika mengukur secara konsisten Realiabel tes tidak ditentukan dengan menguji tes itu sendiri, namun tes sebenarnya harus di uji cobakan untuk menghasilkan informasi yang diinginkan.³¹

c. Memiliki Nilai Objektivitas

Objektivitas suatu tes ditentukan oleh tingkat atau kualitas kesamaan skor-skor yang diperoleh dengan tes tersebut meskipun hasil tes itu dinilai oleh beberapa orang penilai, untuk itu diperlukan kunci jawaban tes (*scoring key*).

Kualitas Objektivitas suatu tes dapat dibedakan menjadi tiga tingkatan, yaitu:³²

- 1) Tinggi, yaitu jika hasil-hasil tes itu menunjukkan tingkat kesamaan yang tinggi.
- 2) Sedang, yaitu seperti tes yang sudah distandarisasi, tetapi pandangan subjektif skor masih mungkin muncul dalam penilaian dan interprestasinya.

³⁰ Charles E. Skinner (ed.), *Essentials of educational Psichology*, (Englewood Cliffs : Prentice-Hall, Inc, tt), hlm. 444.

³¹ *Ibid*, hlm. 445.

³² Chabib Thoha,

- 3) Fleksibel, yaitu seperti beberapa jenis tes yang digunakan oleh LBP (Lembaga Bimbingan dan Penyuluhan) untuk keperluan konseling.

d. Memiliki Nilai Kepraktisan

Kepraktisan suatu tes juga penting diperhatikan. Suatu tes dikatakan mempunyai kepraktisan yang baik jika kemungkinan untuk menggunakan tes itu besar. Kriteria untuk mengukur praktis tidaknya suatu tes dapat dilihat dari:

- 1) Biaya yang diperlukan untuk menyelenggarakan tes itu.
- 2) Waktu yang diperlukan untuk menyusun tes itu.
- 3) Sukar-mudahnya menyusun tes itu.
- 4) Sukar-mudahnya menilai tes itu.
- 5) Sulit-tidaknya menginterpretasikan (mengolah) hasil tes itu.
- 6) Lamanya waktu yang diperlukan untuk melaksanakan tes itu.³³

8. Kegunaan Tes

Penting bagi kita untuk menentukan dahulu alasan mengadakan tes. Kegunaan tes tepat memerlukan satu atau lebih tujuan yang diperlukan dan penting. Perencanaan yang cerdas merupakan dasar yang tepat dalam menggunakan hasil tes. Salah satu cara pengelompokan kegunaan tes berdasarkan fungsi ada tiga, yaitu: para pegawai sekolah atau personelnnya yang bisa kita sebut dengan administrator, supervisor, dan guru.³⁴

a. Kegunaan Bagi Administrator

Hasil tes dapat digunakan untuk menyediakan data perkembangan dan prestasi anak. Hal ini dimaksudkan kedalam kartu data kumulatif anak dan menjadi dasar data permanen evaluasi pertumbuhan dan perkembangan individu maupun kelompok kelas. Kegunaan yang lain untuk menyediakan laporan bagi orang tua. Kemudian yang lain untuk

³³ *Ibid*, hlm. 142.

³⁴ Charles E. Skinner, *Op.Cit.*, hlm. 440

menyediakan data bagi laporan periodik perkembangan sekolah untuk perlindungan dalam masyarakat. Yang terakhir adalah untuk membuat interpretasi status anak lebih baik dan memudahkan penempatan dalam ruang kelas yang cocok.³⁵

b. Kegunaan Bagi Supervisor

Demikian pula supervisor bisa menggunakan hasil tes bagi bermacam-macam tujuan. Tugas utamanya yaitu membantu guru melaksanakan tugas pengajaran yang lebih baik, itu dapat terwujud dengan baik jika antara supervisor dan guru memiliki bukti status anak. Jadi kegunaan tes bagi supervisor adalah untuk menentukan status anak ataupun kelas pada beberapa tujuan utama kurikulum. Hal ini memperbolehkannya menandai perubahan yang diperlukan untuk prosedur instruksional ataupun pembelajaran bagi siswa. Tujuan yang lain yaitu untuk mengevaluasi metode-metode pengajaran atau materi-materi instruksional.³⁶

c. Kegunaan Bagi Guru

Guru menggunakan hasil tes untuk banyak tujuan, banyak diantaranya sama dengan administrator dan supervisor. Kegunaannya hasil tes itu digunakan untuk mengukur diantaranya:

Pertama, Untuk menentukan status tiap anak dalam berbagai subyek dan tujuan kurikulum. *Kedua*, untuk mengevaluasi status dan tingkat pertumbuhan tiap anak dipandang dari segi umur dan kemampuan. *Ketiga*, untuk mengidentivikasi kebutuhan pendidikan tiap anak. *Keempat*, untuk mengidentivikasi anak berbakat, anak normal dan anak yang lamban. *Kelima*, untuk mengelompokkan anak pada kelompok kelasnya. *Keenam*, untuk menganalisa aatau mendiagnosa kesulitan anak dan tingkat

³⁵ *Ibid.*

³⁶ *Ibid*, hlm 441.

pertumbuhan secara individual. *Ketujuh*, untuk menentukan prestasi status sekolah pada awal dan akhir semester.³⁷

Jadi alat-alat evaluasi, khususnya tes hasil belajar dapat dikatakan tes yang valid apabila tes tersebut betul-betul dapat mengukur hasil belajar.³⁸ Jadi bukan sekedar mengukur daya ingatan atau kemampuan bahasa saja.

B. Validitas Tes

1. Pengertian Validitas

Suatu tes hasil belajar dapat dikatakan baik jika memiliki ciri atau mempunyai sifat valid atau shahih atau memiliki validitas. Kata valid sering diartikan dengan tepat, benar, shahih; jadi kata validitas dapat diartikan dengan ketepatan, kebenaran, keshahihian, atau keabsahan.³⁹ Apabila kata valid dikaitkan dengan fungsi tes sebagai alat pengukur, maka sebuah tes dapat dikatakan valid jika tes tersebut secara tepat, secara benar, secara shahih atau secara absah dapat mengukur apa yang seharusnya diukur.

Menurut Annen Anastasi, dalam buku: *Psychological Testing*, Yang dikutip oleh Sumadi Suryabrata mendefinisikan validitas sebagai berikut:

*“Validity, i.e., the degree to which the test actually measures what it purports to measures”*⁴⁰

Artinya: Validitas ialah tingkat dimana sesungguhnya sebuah tes dapat mengukur apa yang seharusnya diukur.

Menurut Muhammad Abdul Kholik Muhammad, dalam kitab *Ihtibarotun Al-Lughoh*, Mendefinisikan validitas:⁴¹

³⁷ *Ibid.*

³⁸ Wayan Nurkencana, *Op. Cit.*, hlm. 1280.

³⁹ Anas Sudjiono, *Pengantar Ilmu Pendidikan*, *Op. Cit.*, hlm. 93.

⁴⁰ Sumadi Suryabrata, *Pembimbing ke Psikodiagnostik*, (Yogyakarta: Andi Offset, 1984), hlm. 23.

⁴¹ Muhammad Abdul Kholik Muhammad, *Ihtibarotun Al-Lughoh*, (Jami'ah Malik Su'ud: 1989), hlm. 48.

الصدق : ان صدق الاختبار يعنى الى اي مدى يقيس الاختبار الشيء الذى
وضع من اجله

Artinya: “Validitas tes yaitu seberapa jauh tes tersebut dapat mengukur apa – apa yang hendak diukur”.

Validitas merupakan suatu ukuran yang menunjukkan tingkat – tingkat kevalidan atau keshahihan suatu instrumen. Suatu instrumen yang valid atau shahih mempunyai validitas yang tinggi. Sebaliknya instrumen yang kurang valid berarti memiliki validitas yang rendah.⁴²

Sebuah instrumen dikatakan valid apabila mampu mengukur apa yang diinginkan. Sebuah instrumen dikatakan valid apabila mampu mengungkap data dari variabel yang diteliti secara tepat. Tinggi rendahnya suatu instrumen menunjukkan seberapa jauh data yang terkumpul tidak menyimpang dari gambaran tentang variabel yang dimaksud.⁴³

Alat – alat evaluasi, khususnya tes hasil belajar dapat dikatakan tes yang valid apabila tes tersebut betul – betul dapat mengukur hasil belajar.⁴⁴ Jadi bukan sekedar mengukur daya ingat atau kemampuan bahasa saja.

2. Macam – Macam Validitas

Validitas sebuah tes dapat diketahui dari hasil pemikiran dan dari hasil pengalamannya. Dua hal yang pertama akan diperoleh Validitas logis (*logical validity*) dan hal yang kedua diperoleh Validitas empiris (*empirical validity*). Dua hal inilah yang dijadikan dasar pengelompokan Validitas.⁴⁵

⁴² Suharsimi Arikunto, *Loc. Cit.*, hlm. 136.

⁴³ *Ibid.*

⁴⁴ Wayan Nurkanca, *Loc. Cit.*, hlm. 128.

⁴⁵ Suharsimi Arikunto, *Op. Cit.*, hlm. 64.

Secara garis besar ada dua macam Validitas, yaitu Validitas tes dan Validitas butir.⁴⁶

a. Validitas Tes

Adapun jenis validitas tes secara umum dapat dikelompokkan, yaitu : Validitas Logis dan Validitas Empiris.

1) Validitas Logis

Validitas logis mengandung arti logika / penalaran. Dengan demikian maka Validitas logis untuk sebuah instrumen evaluasi menunjukkan pada kondisi bagi sebuah instrumen yang memenuhi persyaratan valid berdasarkan hasil penalaran dan sudah dirancang secara baik, sesuai dengan teori dan ketentuan yang berlaku.

Tes hasil belajar yang setelah dilakukan penganalisisan secara rasional ternyata memiliki daya ketepatan mengukur, disebut tes hasil belajar yang telah memiliki validitas logika (*logical validity*). Istilah lain untuk validitas logika adalah : Validitas rasional, Validitas ideal, atau Validitas das sollen.⁴⁷

Ada dua macam validitas logis yang dapat dicapai oleh sebuah instrumen, yaitu : Validitas Isi dan Validitas Konstruksi.

a. Validitas Isi

Validitas isi artinya kejituan suatu tes ditinjau dari isi tes tersebut. Suatu tes hasil belajar dikatakan valid, apabila materi tes tersebut betul-betul merupakan bahan-bahan yang representatif terhadap bahan-bahan pelajaran yang diberikan.⁴⁸ Dan Validitas ini mempersoalkan apakah isi butir tes yang diujikan itu

⁴⁶ Anas Sudijono, *Op. Cit.*, hlm. 65.

⁴⁷ *Ibid.* hlm. 164.

⁴⁸ Wayan Nurkencana, *Loc. Cit.*, hlm. 129.

mencerminkan isi materi kurikulum yang seharusnya diukur atau tidak.⁴⁹

b. Validitas Konstruksi

Validitas konstruksi dapat diartikan sebagai validitas yang bertitik dari segi susunan, kerangka, atau rekan. Sehingga tes hasil belajar dapat dinyatakan memiliki validitas konstruksi, apabila tes hasil belajar tersebut ditinjau dari segi susunan, kerangka atau rekaannya telah dapat secara tepat mencerminkan suatu konstruksi dalam teori psikologis.⁵⁰ Artinya dalam susunan atau kerangkanya benar-benar tepat mengukur aspek-aspek berpikir (aspek Kognitif, Afektif dan psikomotorik). Cara lain untuk menetapkan Validitas konstruksi adalah menghubungkan (korelasi) alat penilaian yang dibuat dengan alat penilaian yang sudah baku (*standardized*) seandainya telah ada yang baku. Bila menunjukkan koefisien korelasi yang tinggi, maka alat penilaian tersebut memenuhi validitas.⁵¹

2) Validitas Empiris

Dimaksudkan dengan validitas empiris adalah memiliki pengertian pengalaman, sehingga sebuah instrumen dikatakan memiliki validitas empiris apabila sudah diuji dari pengalaman. Dengan demikian Validitas empiris tidak dapat diperoleh hanya dg jalan menyusun instrument berdasarkan ketentuan seperti halnya validitas logis, tetapi harus dibuktikan dengan hasil analisis yang dilakukan terhadap data hasil pengamatan di lapangan, terbukti bahwa

⁴⁹ Chabib Thoha, *Loc. Cit.*, hlm. 111.

⁵⁰ Anas Sudjiono, *Loc. Cit.*, hlm. 166.

⁵¹ Nana sudjana, *Penilaian Hasil Proses Belajar Mengajar*, (Bandung : Remaja Rosdakarya, 1999), hlm. 15.

tes hasil belajar itu dengan secara tepat telah dapat mengukur hasil belajar yang seharusnya diukur.

Ada dua cara untuk mengetahui apakah tes hasil belajar itu sudah memiliki validitas empiris atautkah belum, yakni dari segi daya ketepatan meramalnya (*predictive validity*) dan daya ketepatan bandingannya atau “ada sekarang” (*concurrent validity*).⁵²

a) Validitas Ramalan (*predictive validity*)

Memprediksi artinya meramal, dan meramal selalu mengenai hal yang akan datang jadi sekarang belum terjadi. Sebuah tes memiliki Validitas ramalan atau prediksi apabila mempunyai kemampuan untuk meramalkan apa yang akan terjadi pada masa yang akan datang.⁵³

Jika sebuah perguruan Tinggi mampu meramalkan keberhasilan peserta tes dalam mengikuti kuliah dimasa yang akan datang. Calon yang bersaing berdasarkan hasil tes diharapkan mencerminkan tinggi rendahnya kemampuan mengikuti kuliah.

Sebagai alat pembanding validitas prediksi adalah nilai-nilai yang diperoleh setelah peserta tes mengikuti pelajaran di perguruan Tinggi. Jika ternyata siapa yang memiliki nilai tes tinggi ternyata gagal dalam ujian semester I dibandingkan dengan yang dahulu nilai tesnya lebih rendah maka tes masuk yang dimaksud tidak memiliki validitas predeksi.

b) Validitas Bandingan (*Concurrent Validity*)

Validitas bandingan suatu tes artinya membuat tes yang memiliki perbandingan atau kesamaan dengan tes yang sejenis yang telah ada atau yang telah dibakukan. Perbandingan atau kesamaan tes terlingkupnya abilitas yang diukurkannya, sasaran atau

⁵² Anas Sudjiono, *Op.Cit.*, hlm.168.

⁵³ Suharsimi Arikunto, *Loc. Cit.*, hlm. 66.

objek yang diukurnya, serta waktu yang diperlukan. Perbandingan atau kesamaan suatu tes adalah indeks korelasi berdasarkan perhitungan korelasi. Apabila menunjukkan indeks korelasi yang cukup tinggi, yakni mendekati angka satu (korelasi sempurna), berarti tes yang tersusun sudah memiliki validitas banding atau kesamaan.⁵⁴

b. Validitas Butir

Dimaksudkan dengan validitas butir dari suatu tes adalah, ketepatan mengukur yang dimiliki oleh sebutir soal (yang merupakan bagian tak terpisahkan dari tes sebagai suatu totalitas), diukur apa yg seharusnya diukur lewat butir soal tersebut.⁵⁵

Apa bila kita perhatikan secara cermat, maka tes-tes hasil belajar yang dibuat atau disusun oleh guru atau para pengajar sebenarnya adalah merupakan kumpulan dari sekian banyak butir-butir soal; dengan butir mana para penyusun tes ingin mengukur atau mengungkap hasil belajar yang telah dicapai oleh masing-masing individu peserta didik, setelah mereka proses pembelajaran dalam jangka waktu tertentu. Eratnya hubungan antara butir soal dengan tes hasil belajar sebagai suatu totalitas itu kiranya dapat dipahami dari kenyataan, bahwa semakin banyak butir-butir item yang dapat dijawab dengan betul testee, maka hasil skor-skor hasil tes – tes tersebut akan semakin tinggi. Sebaliknya semakin sedikit butir-butir item dapat dijawab betul oleh testee, maka skor-skor total hasil tes itu akan semakin rendah atau semakin menurun.

Sebutir soal dapat dikatakan telah memiliki validitas yang tinggi atau valid, jika skor-skor pada butir-butir soal yang bersangkutan memiliki kesesuaian atau kesejajaran arah dengan skor totalnya; atau dengan bahasa

⁵⁴ Nana Sudjana, *Op. Cit.*, hlm. 15-16.

⁵⁵ Anas Sudijono, *Loc. Cit.*, hlm. 182.

statistik : Ada korelasi positif yang signifikan antara skor butir dengan skor totalnya.⁵⁶

3. Teknik Pengujian Validitas Butir

Uraian yang telah dikemukakan, kiranya menjadi cukup jelas bahwa sebutir soal dapat dikatakan telah memiliki validitas yang tinggi atau valid, jika skor-skor pada butir-butir soal yang bersangkutan memiliki kesesuaian atau kesejajaran arah dengan skor totalnya; atau dengan bahasa statistik : Ada korelasi positif yang signifikan antara skor butir dengan skor totalnya. Skor total berkedudukan sebagai variabel terikat (*dependent variable*), sedangkan skor butir berkedudukan sebagai variabel bebas (*independent variable*).

Untuk sampai pada kesimpulan bahwa butir-butir yang ingin diketahui validitasnya, yaitu valid ataukah tidak, kita dapat menggunakan teknik korelasi sebagai analisisnya. Sebutir soal dapat dinyatakan valid, apabila skor butir yang bersangkutan terbukti mempunyai korelasi yang positif yang signifikan dengan skor totalnya. Seperti diketahui, pada tes obyektif maka hanya ada dua kemungkinan jawaban, yaitu betul dan salah. Setiap butir soal yang dijawab dengan betul umumnya diberi skor 1 (satu), sedangkan untuk setiap jawaban yang salah diberikan skor 0 (nol). Jenis data seperti ini dalam dunia statistik dikenal dengan nama data *diskret murni* atau data kotomik. Sedangkan skor total yang dimiliki oleh masing-masing individu testee adalah hasil penjumlahan dari setiap skor yang dimiliki oleh masing-masing butir soal adalah merupakan data kontinyu.⁵⁷

Menurut teori yang ada, apabila variabel I berupa data *kontinum* (skor hasil tes), sedangkan variabel II berupa data *diskrit murni* (betul atau salahnya

⁵⁶ *Ibid.* hlm. 184.

⁵⁷ *Ibid.* Hlm. 184 – 185.

testee dalam menjawab), maka teknik korelasi antara variabel I dengan variabel II adalah teknik korelasi point biserial (r_{pbis}).⁵⁸

C. Reliabilitas Tes

1. Pengertian Reliabilitas

Reliabilitas sering diartikan keterandalan atau keajegan. Artinya suatu tes memiliki keterandalan bilamana tes tersebut dipakai mengukur berulang-ulang hasilnya sama. Reliabilitas dapat pula diartikan keajegan atau stabilitas.⁵⁹

Menurut Anne Anastasi dalam buku *Psychological Testing*, mendefinisikan Reliabilitas tes sebagai berikut :

*“Reliability refers to the consistency of scores obtained by the same person when they are reexamined with the same test on different occasions, or with different sets of equivalent items.”*⁶⁰

Artinya : Reliabilitas adalah konsisten atau keajegan atau ketetapan dari nilai yang diperoleh dari tiap individu yang sama manakala diadakan tes ulang dengan tes yang sama pada waktu yang berbeda atau dengan butir soal yang sejenis.

Muhammad Abdul Kholik Muhammad, dalam kitab *Ihtibarotun Al-Lughoh*, Mendefinisikan reliabilitas:⁶¹

الشبّات : يقصد بالشبّات عدم التذبذب في الاختبار اذا ما قصد به ان يكون بمشابة

المقياس.

Artinya: “Reliabilitas tes ialah tidak adanya perubahan pada tes yang dilaksanakan menggunakan tes serupa”.

⁵⁸ Anas Sudijono, *Pengantar Statistik Pendidikan*, Jakarta : Rajawali, 1991), hlm. 245.

⁵⁹ Chabib Thoha, *Loc. Cit.*, hlm. 118.

⁶⁰ Anne Anastasi dan Susana Urbina, *Psychological Testing*, (New York : Prentice-Hall, 1988), hlm, 84.

⁶¹ Muhammad Abdul Kholik Muhammad, *Op. Cit.*, hlm. 39.

Reliabilitas menunjukkan pada suatu pengertian bahwa suatu instrumen cukup dapat dipercaya untuk digunakan sebagai alat pengumpul data karena instrumen tersebut sudah baik. Instrumen yang baik tidak akan bersifat tendensius mengarahkan responden untuk memilih jawaban – jawaban tertentu. Instrumen yang sudah dapat dipercaya, yang reliabel akan menghasilkan data yang dapat dipercaya juga, apabila data “memang benar sesuai dengan kenyataannya”, maka berapa kalipun diambil, tetap akan sama. Reliabilitas menunjuk kepada keterandalan sesuatu. Reliabel artinya dapat dipercaya, jadi dapat diandalkan.⁶²

Tes hasil belajar dikatakan baik apabila tes memiliki reliabilitas atau bersifat reliabel. Apabila istilah tersebut dapat dikaitkan dengan fungsi tes sebagai alat ukur keberhasilan belajar peserta didik, maka sebuah tes tersebut dapat dinyatakan reliabel apabila hasil – hasil pengukuran yang dilakukan dengan menggunakan tes tersebut secara berulang kali terhadap subyek yang sama, senantiasa menunjukkan hasil yang tetap sama atau yang sifatnya ajeg atau setabil.⁶³

2. Faktor – Faktor yang Mempengaruhi Reliabilitas

- a. Luas tidaknya sampling yang diambil

Makin luas suatu sampling, berarti tes semakin andal.

- b. Perbedaan bakat dan kemampuan murid yang dites.

Makin variatif kemampuan peserta tes, berarti makin tinggi keandalan koefisien tes. Tes yang diberikan kepada beberapa tingkat kelas yang berbeda, lebih tinggi keandalannya dari pada yang hanya diberikan kepada beberapa kelas yang sama karena tingkat kelas yang berbeda akan menghasilkan achievement yang lebih luas.

- c. Suasana dan kondisi testing.

⁶² Suharsimi Arikunto, *Loc. Cit.*, hlm. 142.

⁶³ Anas Sudjiono, *Pengantar Evaluasi Pendidikan*, *op. Cit.*, hlm. 95.

Suasana ketika sedang berlangsung testing, seperti tenang, gaduh, banyak gangguan, pengetes yang marah-marah dapat mengganggu pekerjaan tes sehingga dengan demikian mempengaruhi pula hasil dan keandalan tes.⁶⁴

3. Teknik Pengujian Reliabilitas

Berbeda dengan tes hasil belajar bentuk uraian, maka pada bentuk tes objektif penentuan reliabilitas tes dapat menggunakan tiga macam pendekatan, ketiga macam pendekatan tersebut adalah sebagai berikut:⁶⁵

a. Pendekatan *Test-retest* (Pendekatan Bentuk Ulangan)

Untuk menguji reliabilitas tes dengan jalan mengujikan tes tersebut dua kali atau lebih, kemudian hasilnya dikorelasikan. Tujuan uji reliabilitas ini yaitu mengetahui koefisien stabilitas tes. Tes tersebut memiliki keterandalan atau keajegan bilamana dipakai mengukur objek sama dalam waktu berbeda-beda hasilnya tetap sama. Ada 6 (enam) langkah untuk dapat ditempuh pada uji reliabilitas ini sebagai berikut :

- 1) Menyusun sebuah tes yang akan diukur reliabilitasnya.
- 2) Mengujikan tes yang tersusun tersebut (tahap I).
- 3) Menghitung hasil tes tersebut tahap I.
- 4) Mengujikan ulang tes yang tersusun tersebut (tahap II).
- 5) Menghitung skor hasil tes ulang (tahap II).
- 6) Menghitung reliabilitas tes tersebut dengan jalan mengkorelasikan skor tes I dengan skor tes II dengan rumus korelasi *rank – order* (teknik korelasi tat jenjang) dari *Spearman*.⁶⁶

⁶⁴ M. Ngalim Purwanto, *Op. Cit.*, hlm. 141.

⁶⁵ Anas Sudijono, *Op. Cit.*, hlm. 213.

⁶⁶ *Ibid.* hlm. 269.

b. Pendekatan *Alternate Form* (Pendekatan Bentuk Paralel)

Pengujian reliabilitas dengan jalan melakukan pengukuran dengan menggunakan dua jenis tes yang mana butir – butir soalnya sejenis tapi tidak sama, tes diujikan secara bersamaan oleh dua kelompok. Adapun untuk mencari atau menghitung reliabilitas tes, dapat dipergunakan teknik korelasi *Product Moment* dari Pearson atau teknik korelasi *rank – order* dari Spearman (khusus untuk N kurang dari 30).⁶⁷

Adapun langkah – langkah yang ditempuh adalah senbagai berikut:⁶⁸

- 1) Menyusun dua tes yang ekuivalen.
- 2) Mengujikan kedua tes tersebut (dalam kurun waktu yang beriringan).
- 3) Memberikan skor hasil tes yang sudah diujikan, disusun dengan memisahkan antara tes A dengan tes B.
- 4) Mencari Koefisien stabilitas kedua tes (A dan B) dengan jalan mencari korelasinya melalui rumus korelasi *Product Moment*.

c. Pendekatan *Single Test – Single Trial* (Pendekatan “serba satu”)

Pengujian reliabilitas dengan jalan melakukan pengukuran terhadap satu kelompok subjek, pengukuran dilakukan dengan hanya menggunakan satu jenis alat pengukur, dan pelaksanaannya dilakukan sebanyak satu kali saja. Adapun untuk mencari atau menghitung reliabilitas tes, dapat menggunakan lima jenis formula, yaitu : a) Formula Spearman-Brown, b) Formula Flanagan, c) Formula Rulon, d) Formula Kuder-Richardson, dan e) Formula C. Hoyt.⁶⁹

Pengujian reliabilitas tes, khususnya pendekatan *Single Test – Single trial* (*Pendekatan serba satu*) dapat menggunakan dua metode, yaitu :

⁶⁷ *Ibid.* hlm. 273

⁶⁸ Chabib Thoha, *Op. Cit.*, hlm. 123.

⁶⁹ Anas Sudijono, *Op. Cit.*, hlm. 214.

1) Metode Split-Half Reliability

Metode ini dipakai untuk mengetahui tingkat reliabilitas tes dengan jalan membelah tes menjadi dua bagian dan skor kedua belahan tersebut dikorelasikan pada rumus tertentu.

Cara melakukan pembelahan hasil tes tersebut dapat dilakukan dua jalan, yaitu membelah antara skor genap, atau membelah antara belahan nomor atas dan nomor bawah.

Adapun langkah secara umum yang ditempuh untuk mencari reliabilitas tes ini adalah :⁷⁰

- a) Menyusun sebuah tes sebaiknya jumlah nomornya genap, sehingga bila dibelah jumlahnya sama.
- b) Mengujikan tes tersebut pada satu sampel.
- c) Menghitung skor masing-masing peserta tes dua kelompok skor, dapat dikelompokkan skor ganjil dan genap; dapat pula dikelompokkan skor belahan atas dan bawah.
- d) Mencari realibitas setengah tes, menggunakan cara mengkorelasikan kedua skor tersebut dengan rumus Product Moment, atau mencari deviasi pada belahan ganjil genap.
- e) Mencari realibitas satu tes penuh menggunakan rumus:
 - (1) Rumus Sperman-Brown
 - (2) Rumus Flanagan
 - (3) Rumus Rulon

2) Uji Homogenitas

Di antara metode untuk mengakui koefisien konsistensi untuk mengetahui realibitas tes, dapat digunakan pendekatan yang tidak membelah tes menjadi dua. Hal ini disebabkan oleh dua kemungkinan,

- (a) Jumlah butir ganjil, sehingga tidak dapat dibelah menjadi dua, (b)

⁷⁰ Chabib Thoha, *Loc. Cit.*, hlm. 124.

Komposisi antara butir-butir ganjil dan genap tidak homogen, sehingga bila dibelah cenderung tidak memiliki korelasi positif.⁷¹

Adapun langkah-langkah pencarian realibilitasnya sebagai berikut:

- a) Membuat tabel analisis butir tanpa harus mengelompokkan nomor ganjil dan genap.
- b) Menghitung proporsi yang menjawab benar dan proporsi yang menjawab salah pada masing-masing butir dalam tabel analisis butir.
- c) Mengalikan proporsi yang menjawab benar dengan yang menjawab salah.
- d) Mencari varians (standar deviasi kuadrat) dari skor total.
- e) Menghitung reliabilitas tes dengan menggunakan rumus:
 - (1) Rumus Kuder-Richardson (K-R 20 dan K-R 21)
 - (2) Rumus C. Hoyt / Alpha

Uraian yang telah dikemukakan, kiranya menjadi cukup jelas. Kemudian langkah pengujian reliabilitas yang akan dilakukan dalam penelitian ini ialah penulis akan menggunakan rumus: K-R 20.

Alasan penulis menggunakan rumus KR-20 lebih tepat hasilnya apabila dilakukan secara langsung terhadap butir-butir soal tes yang bersangkutan.⁷²

D. Tingkat Kesukaran

a. Pengertian Analisis Tingkat kesukaran

Analisis tingkat kesukaran soal artinya: mengkaji soal – soal tes dari segi kesulitan sehingga dapat diperoleh soal – soal yang termasuk mudah, sedang, dan sukar.⁷³

⁷¹ *Ibid.* hlm. 133.

⁷² Anas Sudijono, *Loc. Cit.*, hlm. 253.

Tingkat kesukaran soal dipandang dari segi kesanggupan siswa atau kemampuan siswa waktu menjawabnya, bukan dilihat dari sudut guru sebagai pembuat soal. Persoalan yang penting melakukan analisis tingkat kesukaran soal yaitu penentuan proporsi dan kriteria soal yang termasuk mudah, sedang, dan sukar.⁷⁴

b. Menentukan Proporsi Jumlah Soal Kategori Mudah, Sedang, dan Sukar.

Ada beberapa dasar pertimbangan menentukan proporsi jumlah soal kategori mudah, sedang, dan sukar. Pertimbangan yang pertama yakni adanya keseimbangan, yakni jumlah soal sama untuk ketiga kategori tersebut. Artinya, soal mudah, sedang, dan sukar jumlahnya seimbang. Pertimbangan yang kedua proporsi jumlah soal untuk ketiga kategori tersebut didasarkan atas kurva normal. Artinya sebagian besar soal berada pada kategori sedang, sebagian lagi termasuk kategori mudah, dan sebagian lagi dalam kategori sukar dengan proporsi yang seimbang.⁷⁵

Perbandingan antara soal mudah – sedang – sukar bisa dibuat 3-4-3. artinya, 30%soal kategori mudah, 40% soal kategori sedang, dan 30% soal kategori sukar. Perbandingan lain yang sejenis dengan proporsi diatas 3-5-2. artinya, 30% soal kategori mudah, 50% soal kategori sedang, dan 20% kategori sukar.⁷⁶

E. Daya Pembeda

a. Pengertian Daya Pembeda

Yang dimaksud dengan daya pembeda suatu soal tes adalah: bagaimana kemampuan soal itu untuk membedakan siswa – siswa yang

⁷³ Nana Sudjana, *Op. Cit.*, hlm. 135.

⁷⁴ *Ibid.*

⁷⁵ *Ibid.*

⁷⁶ *Ibid.*, hlm. 136

termasuk kelompok pandai dengan siswa – siswa yang termak dalam kelompok sukurang.⁷⁷

b. Langkah – Langkah Menghitung Daya Pembeda

Untuk Menghitung daya pembeda dapat adalah dengan menempuh cara langkah sebagai berikut:⁷⁸

- 1) Memeriksa semua jawaban peserta tes.
- 2) Membuat daftar peringkat hasil tes berdasarkan sekor yang dicapainya.
- 3) Menentukan jumlah sampel sebanyak 27% dari jumlah peserta tes untuk kelompok siswa pandai (peringkat atas) dan 27% untuk kelompok siwa kurang (peringkat bawah).
- 4) Melakukan analisis butir soal, yakni menghitung jumlah siswa yang menjawab dari semua nomor soal, baik dari kelompok pandai maupun dari kelompok kurang.
- 5) Menghitung selisih jumlah siswa yang salah menjawab pada kelompok kurang dengan kelompok pandai ($SR - ST$).
- 6) Membandingkan nilai selisih yang diperoleh dengan nilai tabel.
- 7) Menentukan ada tidaknya daya pembeda pada setiap nomor soal dengan menghitung dari selisih jawabah siswa yang menjawab salah antaran kelompok kurang dan kelompok pandai ($SR - ST$) sama atau lebih besar dari nilai tabel.

F. Efektifitas Fungsi Distraktor

a. Pengertian Distraktor.

Distraktor adalah alternatif jawabah yang salah atau penyasat. Dalam setiap tes obyektif selalu digunakan alternatif jawaban yang mengandung dua unsur

⁷⁷ Daryanto, *Evaluasi Pendidikan*, (Jakarta, Rineka Cipta, 2001), hlm. 183.

⁷⁸ Nana Sudjana, *Op. Cit.*, hlm. 143 – 144.

sekaligus, yaitu jawaban yang tepat dan jawaban yang salah sebagai pengecoh.⁷⁹

b. Tujuan Pemakaian Fungsi Distraktor

Tujuan pemakaian fungsi distraktor adalah mengecoh mereka yang kurang mampu (tidak tahu) untuk dapat dibedakan dengan yang mampu. Oleh karena itu distraktor yang baik adalah yang dapat dihindari oleh anak – anak yang pandai dan terpilih oleh anak – anak yang kurang pandai.⁸⁰

G. Kajian Penelitian yang Relevan

Ada beberapa research tulisan yang membahas tentang analisis instrumen tes yang meng ilhami penulis untuk membahas tema yang cukup luas ini antara lain:

1. Skripsi dari penulis Mohamad Nurdin denngan judul “*Uji validitas dan reliabilitas tes PAI mata pelajaran Akhlak di SMP Muhammadiyah 08 Mijen Semarang Tahun 2003/2004*” yang bertujuan untuk menguji; *pertama*, validitas tes PAI mata pelajaran Akhlak kelas III yang digunakan di SMP Muhammadiyah 08 Mijen Semarang tahun 2003/2004 . penelitian ini menunjukkan hasil bahwa:⁸¹
 - a. Untuk validitas PAI mata pelajaran Akhlak kelas III, ditunjukkan dengan 28 butir soal yang bersangkutan dinyatakan valid dan 22 butir soal dinyatakan invalid.
 - b. Untuk reliabilitas tesnya ditunjukkan dengan koefesien korelasi $r_{11} = 0,640$. hal ini menunjukkan bahwa reliabilitas tes PAI mata pelajaran

⁷⁹ Chabib Thoha, *Op. Cit.*, hlm. 149.

⁸⁰ *Ibid.*

⁸¹ Mohamad Nurdin, “Uji Validitas dan Reliabilitas tes PAI Mata Pelajaran Akhlak di SMP Muhammadiyah 08 Mijen Semarang Tahun 2003/2004”, Skripsi S1 IAIN Walisongo Semarang, (Semarang: Perpustakaan Tarbiyah, IAIN Walisongo Semarang, 2004), hlm. 61, t.d.

Akhlak kelas III tahun 2003/2004 adalah memiliki reliabilitas yang sedang.

2. Skripsi dari penulis Mawardi dengan judul “Validitas dan *reliabilitas instrumen tes buatan guru pelajaran fiqh kelas VII MTs Al – ma’arif NU Boja Kendal*” yang menganalisa validitas dan reliabilitas tes buatan guru pelajaran fiqh kelas VII di MTs Al – Ma’arif Boja dengan hasil Sebagai berikut:⁸²

- a. Hasil uji validitas terhadap butir- butir soal tes objektif pada mata pelajaran fiqh kelas VII buatan guru PAI di MTs NU 02 Al – Ma’arif Boja, diketahui bahwa ada 13 butir soal atau sekitar 37% dari butir soal tersebut yang dinyatakan memiliki validitas, sedangkan 22 butir soal atau sekitar 63% yang dinyatakan drop. Dengan demikian tes buatan guru PAI mata pelajaran Fiqh kelas VII MTs NU 02 Al – Ma’arif Boja Kendal tersebut memiliki validitas butir yang sedang.
- b. Berdasarkan analisis terhadap Reliabilitas tes objektif pada mata pelajaran Fiqh kelas VII MTs NU 02 Al – Ma’arif Boja Kendal, diperoleh koefisien reliabilitas (r_{11}) sebesar 0,5. Dengan demikian, maka penulis dapat menyimpulkan bahwa butir soal pada mata pelajaran Fiqh kelas VII MTs NU 02 Al – Ma’arif Boja Kendal dapat dinyatakan memiliki reliabilitas yang sedang, hal ini didasarkan pada tabel reliabilitas tes r_{11} yang menunjukkan nilai r_{11} sebesar 0,5 terletak pada interval 0,40 – 0,60 dalam kategori sedang.

Penelitian diatas, hanya menitik beratkan pada analisis validitas dan reliabilitas. Dan tidak kalah pentingnya, dalam skripsi ini penulis akan mengkaji dan mengembangkan secara rinci tentang analisis instrumen tes yang memuat

⁸² Mawardi, “Validitas dan reliabilitas instrumen tes buatan guru pelajaran fiqh kelas VII MTs Al – ma’arif NU Boja Kendal”, Skripsi S1 IAIN Walisongo Semarang, (Semarang: Perpustakaan Tarbiyah, IAIN Walisongo Semarang, 2007), hlm. 57 – 58, t.d.

tentang: analisis validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta fungsi distraktor.

BAB III

METODE PENELITIAN

A. Waktu dan Tempat Pelaksanaan

Penelitian lapangan ini mengambil lokasi di MTs Miftahussalam 1 yang beralamat di desa Wonosalam Kec. Wonosalam Kab. Demak.

Kegiatan penelitian ini dilaksanakan pada semester I tahun 2008/2009 dengan tahapan sebagai berikut:

- a. Melakukan survei awal pada tanggal 1 Desember 2008, survei pertama ini bertujuan untuk memperoleh informasi tentang pelaksanaan ujian semester I untuk mata pelajaran Aqidah Akhlak kelas VII.
- b. Melakukan pendekatan kepada kepala sekolah MTs Miftahussalam 1 dan khususnya guru bidang studi Aqidah Akhlak untuk memberikan informasi dan memberikan ijin dalam penggandaan arsip jawaban siswa kelas VII yang mengikuti ujian Aqidah Akhlak.
- c. Pengumpulan data lapangan dan analisis data yang terkumpul akan peneliti laksanakan pada bulan April 2009.

B. Variabel Penelitian

Variabel adalah objek penelitian atau apa yang menjadi titik perhatian suatu penelitian.¹

Dalam penelitian ini terdapat dua variabel, yakni variabel bebas dan variabel terikat. Sebagai variabel bebas (*Independent Variable*) yaitu variabel yang berperan memberi pengaruh yakni skor butir soal. Untuk variabel ini menggunakan jenis data deskriptif murni (nominal); karena hanya dapat dikategorikan atas dua kutub yang berlawanan yakni "1" dan "0". Sebagai variabel terikat (*Dependent Variable*) yaitu variabel yang mendapatkan

¹ Suharsimi Arikunto, *Prosedur Penelitian (Suatu Pendekatan Praktik)*, (Jakarta: Rineka Cipta, 2002), Cet. XII, hlm. 96.

pengaruh yakni skor total instrumen tes. Sedangkan dalam variabel ini termasuk jenis data kontinum (kontinyu) : karena merupakan hasil penjumlahan dari skor masing-masing butir yakni : $0+1+0+1+0+1=3$.

C. Metode Penelitian

Metode penelitian adalah cara – carayang digunakan untuk mengumpulkan data, yang dikembangkan untuk memperoleh pengetahuan dengan mengajukan prosedur yang reliabel dan terpercaya.²

Dalam pengumpulan data peneliti menggunakan dua macam pendekatan atau penelitian, yaitu penelitian kepustakaan (*Library Research*) dan pendekatan lapangan (*Field Research*).

1. *Library Research*

Penelitian ini akan digunakan untuk memperkaya teori ilmiah dengan cara mengambil pendapat para ahli yang ada relevansinya dengan pokok permasalahan ini.

Adapun cara berpikir yang digunakan adalah: berpikir deduktif yaitu apasaja yang dipandang benar pada semua peristiwa dalam suatu kelas atau sejenis berlaku pula pada semua peristiwa (khusus maupun umum) itu menjadi peristiwa yang khusus.³

Cara ini digunakan untuk mengambil kesimpulan dari beberapa pendapat tentang definisi – definisi dari para ahli. Berpikir induktif yaitu suatu metode yang berangkat dari fakta – fakta yang khusus atau peristiwa yang kongkrit kemudian dari peristiwa yang khusus atau yang kongkrit itu ditarik generalisasi yang mempunyai sifat umum.⁴

Cara berpikir ini digunakan untuk mengambil kesimpulan apakah hasil penelitian dilappangan sesuai dengan landasan teori.

² Sutrisno Hadi, *Metodologi Research*, (Yogyakarta, Andi Offset, 1995), hlm. 42.

³ *Ibid.*, hlm. 36.

⁴ *Ibid.*, hlm. 42.

2. *Field Research*

Dalam penelitian ini, peneliti menggunakan metode survei yaitu pengamatan, penyelidikan yang kritis untuk mendapatkan keterangan yang jelas dan baik terhadap suatu persoalan tertentu dan dari dalam daerah tertentu.⁵

D. Subyek Penelitian

Salah satu langkah yang harus dilakukan dalam penelitian adalah menentukan subyek penelitian. Subyek dalam penelitian ini adalah soal tes mata pelajaran Aqidah Akhlak buatan guru KKM rayon 2 kelas VII semester I yang telah dikerjakan oleh seluruh siswa kelas VII MTs Miftahussalam 1 Wonosalam Demak yang berjumlah 120 siswa.

E. Teknik Pengumpulan Data

Untuk memperoleh data yang diperlukan dalam penelitian ini, peneliti menggunakan Metode Dokumentasi.

Dokumentasi dalam arti sempit diartikan sebagai kumpulan data verbal yang berbentuk tulisan, sedangkan dalam arti luas dokumentasi juga meliputi : monumen, foto, tape, dan lain sebagainya.⁶

Metode ini digunakan untuk menghimpun data – data dengan melihat bahan – bahan yang ada di MTs Miftahussalam 1 Wonosalam Demak, khususnya yang berkaitan dengan soal dan hasil ujian dari tes saemester I mata pelajaran Aqidah Akhlak yang telah di ujikan.

F. Teknik Analisis Data

Untuk mengolah data mentah yang telah peneliti peroleh dari tempat penelitian, peneliti menganalisisnya dengan metode deskriptif kuantitatif. Metode ini digunakan untuk memecahkan masalah penelitian dengan

⁵ Margono, *Metodologi Penelitian Pendidikan*, (jakarta, Rineka Cipta, 2000), Cet. II, hlm. 29.

⁶ Kontjoroningrat, *Metode Penelitian Masyarakat*, (Jakarta: Gramedia, tt), hlm. 63.

memaparkan keadaan sohyek yang diselidiki sebagai mana adanya, berdasarkan fakta – fakta yang aktual.⁷ Adapun tahapan-tahapan dalam pengolahan dan menganalisisnya adalah sebagai berikut:

1. Analisis Validitas Butir/ Korelasi skor Butir dengan skor Total

Analisis yang dipergunakan untuk mengolah data yang telah diperoleh dari lapangan dengan jalan mengadakan tabulasi atau tabel. Adapun analisis yang dipergunakan peneliti adalah dengan menggunakan teknik pengujian Validitas butir.

Menurut teori yang ada, apabila variable I berupa data kontinum (skor hasil tes), sedangkan variable II berupa data diskrit murni (betul atau salahnya testee dalam menjawab), maka teknik korelasi yang tepat untuk digunakan dalam mencari korelasi antara variabel I dengan variabel II adalah teknik korelasi point biserial. Adapun langkah-langkahnya adalah sebagai berikut:⁸

1. Menyiapkan tabel perhitungan dalam rangka analisis butir soal 1 sampai dengan 40.
2. Mencari mean dari skor total, yaitu M_t , dengan menggunakan rumus:

$$M_t = \frac{\sum X_t}{N}$$

3. Mencari Deviasi standar total, yaitu SD_t , dengan menggunakan rumus:

$$SD_t = \sqrt{\frac{\sum X_t^2}{N} - \left[\frac{\sum X_t}{N} \right]^2}$$

4. Mencari (menghitung) M_p untuk butir soal nomor 1 sampai dengan 40.

⁷ Hadari Nawawi dan Martini Hadari, *Instrumen Penelitian Bidang Sosial*, (Yogyakarta: Gajah Mada Universiyi Pers, 1992), hlm. 67.

⁸ *Ibid*, hlm. 187 – 189.

5. Mencari (menghitung) koefisien korelasi r_{pbis} dari nomor 1 sampai dengan nomor 40, dengan menggunakan rumus:

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

Ket:

r_{pbis} = koefisien korelasi point biserial yang melambangkan kekuatan korelasi antara variabel I dengan variabel II, (skor butir dengan skor total).

M_p = skor rata-rata hitung yang dimiliki testee, yang untuk butir soal yang bersangkutan telah dijawab dengan betul.

SD_t = deviasi standar dari skor total.

p = proporsi testee yang menjawab betul.

q = proporsi testee yang menjawab salah.

Dalam pemberian interpretasi terhadap (r_{pbis}) dengan db (derajat kebebasan) sebesar N yaitu 120. dengan memeriksa nilai “r” ternyata db sebesar 120 tidak terdapat dalam tabel, sehingga kita gunakan db yang terdekat dengan 120 yaitu db sebesar 125. Dengan demikian db sebesar 125 diperoleh harga r_{tabel} sebagai berikut:

Pada taraf signifikansi 5 % (r_{tabel}) = 0,176

Pada taraf signifikansi 1 % (r_{tabel}) = 0,230

Apabila nilai (r_{pbis}) hasil koefisien korelasi lebih besar (>) dari nilai r_{tabel} 0,176 untuk taraf 5 %, maka hasil yang diperoleh adalah signifikan, artinya butir soal tes dinyatakan valid. Apabila nilai (r_{pbis}) hasil koefisien korelasi lebih kecil (<) dari nilai r_{tabel} = 0,176, maka hasil yang diperoleh adalah non signifikan, artinya butir soal tes dinyatakan invalid.

2. Analisis Reliabilitas

Peneliti dalam menentukan reliabilitas tes adalah lebih tepat apabila dilakukan dengan secara langsung terhadap butir-butir item tes yang bersangkutan.

Adapun formula yang diajukan oleh Kuder dan Richardson ada dua buah yang masing-masing diberi kode K-R 20 dan K-R 21, tetapi yang peneliti gunakan adalah rumus K-R 20, yaitu:

$$r_{11} = \left[\frac{n}{n-1} \right] \left[- \frac{S_t^2 - p_i q_i}{S_t^2} \right]$$

Ket :

r_{11} = koefisien reliabilitas tes

n = Banyaknya butir soal

1 = Bilangan konstan

S_t = Varian total

p_i = Proporsi testee yang menjawab betul

q_i = Proporsi testee yang menjawab salah butir soal

$p_i q_i$ = Jumlah dari hasil perkalian antara $p_i q_i$.

Adapun langkah-langkah yang ditempuh dalam melakukan uji reliabilitas tes adalah sebagai berikut:

- a. Menyiapkan tabel perhitungan dalam rangka uji reliabilitas tes dengan menampilkan 40 butir soal.
- a. Mencari jumlah kuadrat total ($\sum X_t^2$),
- b. Mencari varian total (S_t),
- c. Dan yang terakhir adalah melakukan perhitungan untuk mengetahui reliabilitasnya dengan menggunakan rumus K-R 20.

Selanjutnya dalam pemberian interpretasi terhadap angka korelasi koefisien pada umumnya digunakan patokan sebagai berikut:

Interval	Interpretasi
0,90 – 1,00	Sangat tinggi
0,70 – 0,90	Tinggi
0,40 – 0,70	Cukup/ sedang
0,20 – 0,40	Rendah
0,00 – 0,20	Sangat rendah

3. Tingkat Kesukaran Butir Soal

Bermutu atau tidaknya butir-butir soal tes hasil belajar dapat diketahui dari derajat kesukaran atau taraf kesulitan yang dimiliki oleh masing-masing butir soal. Angka indeks kesukaran butir soal dapat diperoleh dengan menggunakan rumus sebagai berikut:⁹

$$p = \frac{N_p}{N}$$

Dimana:

p = Proporsi angka indeks kesukaran butir

N_p = banyaknya testee yang dapat menjawab betul

N = jumlah testee yang mengikuti tes hasil belajar

Mengenai bagaimana cara memberikan penafsiran (interpretasi) terhadap indeks angka kesukaran butir, Anas Sudjiono mengutip dari Witherington dalam bukunya berjudul *Psychological Education* adalah sebagai berikut:¹⁰

⁹ Anas Sujiono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Raja Grafindo, 2003), hlm. 372.

¹⁰ *Ibid*, hlm. 373.

Besarnya P	Interpretasi
Kurang dari 0,25	Terlalu sukar
0,25 – 0,75	Sedang atau cukup
Lebih dari 0,75	Terlalu mudah

4. Daya Pembeda

Daya pembeda butir soal adalah kemampuan suatu butir tes untuk dapat membedakan antara testee yang berkemampuan tinggi (pandai), dengan testee yang berkemampuan rendah (bodoh) sehingga sebagian besar testee yang memiliki kemampuan tinggi menjawab butir soal tersebut lebih banyak yang menjawab betul dan sebaliknya.¹¹

Teknik analisis daya pembeda butir tes ini dimana jumlah testee cukup besar yaitu sekitar 120 Orang, maka untuk analisis daya pembedanya cukup dihitung berdasar 27 % testee kelompok atas dan 27 % kelompok bawah. Hal ini disebabkan berdasarkan bukti-bukti empiris pengambilan subyek 27 % telah menunjukkan kesensitifannya atau dengan kata lain dapat diandalkan.¹²

Angka indek daya pembeda butir dapat diperoleh dengan menggunakan rumus yaitu:¹³

$$D = PA - PB \text{ atau } PH - PL$$

$$P_B = P_L = \frac{B_A}{J_B} \text{ dan } P_A = P_H = \frac{B_B}{J_A}$$

Dimana:

PA atau PH = proporsi testee kelompok atas yang dapat menjawab dengan betul butir soal yang bersangkutan.

PB atau PL = proporsi testee kelompok bawah yang dapat menjawab betul butir soal yang bersangkutan.

¹¹ *Ibid*, hlm. 386.

¹² *Ibid*, hlm. 387-389.

¹³ *Ibid*, hlm. 389.

B_A = banyaknya testee kelompok atas yang menjawab betul.

J_A = jumlah testee kelompok atas

B_B = banyaknya testee kelompok bawah yang menjawab betul

J_B = jumlah testee kelompok bawah.

Mengenai bagaimana cara memberikan penafsiran (interpretasi) terhadap indeks daya pembeda butir, Anas Sudjiono memberikan patokan sebagai berikut:¹⁴

Indeks Daya Pembeda	Interoretasi
Bertanda negatif	Lemah sekali
Kurang dari 0,20	Lemah
0,20 – 0,40	Cukup atau sedang
0,40 – 0,70	Baik
0,70 – 1,00	Baik sekali

5. Analisis Fungsi Distraktor

Option atau alternatif jawaban yang terdapat pada setiap butir soal, salah satu diantaranya adalah merupakan kunci jawaban, sedangkan yang lainnya adalah merupakan jawaban yang salah atau disebut dengan distraktor.¹⁵

Perhitungan fungsi distraktor dapat diperoleh dengan menggunakan rumus:¹⁶

$$\frac{\text{Banyaknya testee yang menjawab option}}{\text{Jumlah testee yang mengikuti tes}} \times 100 \%$$

¹⁴ *Ibid.*

¹⁵ *Ibid*, hlm. 409.

¹⁶ *Ibid*, hlm. 412.

Untuk mengetahui apakah suatu distraktor telah dapat berfungsi dengan baik, apabila distraktor tersebut telah sekurang-kurangnya dipilih oleh 5 % dari seluruh peserta tes.

BAB IV

ANALISIS DAN PEMBAHASAN HASIL PENELITIAN

Untuk mengetahui ada tidaknya kualitas atau mutu sebuah tes, dapat diketahui bagaimana validitas dan reliabilitas tesnya, tingkat kesukarannya, daya pembedanya, dan fungsi distraktor, kemudian dalam skripsi ini akan meneliti tentang kualitas tes Akidah Akhlak buatan guru KKM rayon 2 kelas VII di MTs Miftahussalam 1 Wonosalam Demak tahun 2009, dari data yang telah peneliti peroleh akan dianalisa dalam bentuk deskriptif kuantitatif. Pada umumnya penelitian deskriptif merupakan penelitian non hipotesis, sehingga dalam langkah penelitiannya tidak perlu merumuskan hipotesis.¹ Analisis deskriptif kuantitatif merupakan penelitian yang berusaha menggambarkan dan menginterpretasi objek sesuai dengan apa adanya.² Peneliti tidak melakukan manipulasi data dan tidak menetapkan peristiwa yang akan terjadi. Sedangkan kuantitatif adalah diwujudkan dalam bentuk angka - angka hasil perhitungan atau pengukuran.

A. DESKRIPSI DATA TENTANG TES AQIDAH AKHLAK

Tes Akidah Akhlak kelas VII yang diselenggarakan di MTs Miftahussalam 1 Wonosalam Demak disusun oleh Kelompok Kerja Madrasah rayon 2. Peneliti memanfaatkan subyek atau populasi yang tersedia; yaitu seluruh siswa kelas VII yang mengikuti tes semester I mata pelajaran Aqidah Akhlak tahun 2008/2009 berjumlah 120 siswa.

Tes Akidah Akhlak ini terdiri dari 45 butir soal. Jenis soal yang digunakan pada tes ini adalah berbentuk tes obyektif dan tes subyektif. Bentuk soal tes obyektif yaitu menggunakan model pilihan ganda, yang terdiri dari 40 soal dan bentuk soal tes subyektif adalah tes essay terdiri dari 5 soal. Tetapi yang menjadi fokus penelitian ini adalah jenis soal obyektif yang bentuknya *multiple choice* (pilihan ganda) yang terdiri dari 40 soal.

¹ Suharsimi Arikunto, *Prosedur Penelitian (Suatu Pendekatan Praktik)*, (Jakarta: Melton Putra, 1991), hlm. 194.

² Sukardi, *Metodologi Penelitian Pendidikan (Kompetensi dan Praktiknya)*, (Jakarta: Bumi Aksara, 2003), hlm. 157.

Selanjutnya hasil pengumpulan data yang berupa hasil jawaban tes sumatif pada mata pelajaran Aqidah Akhlak kelas tahun 2008/2009 MTs Miftahussalam 1 Wonosalam Demak dapat dilihat pada lampiran.

B. ANALISIS VALIDITAS, RELIABILITAS, TINGKAT KESUKARAN, DAYA PEMBEDA DAN FUNGSI DISTRAKTOR

1. Analisis Validitas Butir

Untuk melakukan analisis validitas butir tes yang menampilkan 40 butir soal, maka kita harus melihat pada lampiran II. Tabel analisis yang dapat digunakan untuk mencari: M_p , M_t , SD_t , p dan q .

Langkah selanjutnya adalah melakukan analisis uji validitas butir soal sebagai berikut:

- Mencari mean dari skor total, yaitu M_t , dengan menggunakan rumus:

$$M_t = \frac{\sum X_t}{N}$$

- Menghitung SD_t , dengan rumus :

$$SD_t = \sqrt{\frac{\sum X_t^2}{N} - \left[\frac{\sum X_t}{N} \right]^2}$$

- Menghitung M_p , dengan rumus:

$$M_p = \frac{\text{Jumlah skor total testee yang menjawab benar}}{\text{jumlah testee yang menjawab benar}}$$

Dari perhitungan yang telah dilakukan, dapat dilihat hasilnya pada tabel 1 sebagai berikut:

TABEL 1
TABEL UNTUK MENCARI Mp BUTIR TES
MATA PELAJARAN AQIDAH AKHLAK KELAS VII

No Btr	Total skor seluruh testee yang menjawab benar	Jumlah testee yang menjawab benar	Mp
1	2570	100	25,700
2	1847	69	26,768
3	2717	107	25,392
4	1271	46	27,630
5	1619	62	26,112
6	270	11	24,545
7	1868	71	26,309
8	417	17	24,529
9	168	6	28,000
10	2417	93	25,989
11	2498	97	25,752
12	2353	90	26,144
13	2112	79	26,734
14	2314	88	26,295
15	2392	90	26,577
16	2452	93	26,365
17	2252	83	27,132
18	1475	60	24,583
19	1964	75	26,186
20	1891	69	27,405
21	1905	70	27,214
22	399	16	24,934
23	1146	46	24,913

24	1992	73	27,287
25	2673	101	26,465
26	1670	62	26,935
27	2047	75	27,293
28	2780	107	25,981
29	3432	93	36,903
30	2986	117	25,521
31	1594	57	27,964
32	1787	65	27,492
33	1218	45	27,066
34	2583	101	25,574
35	2834	111	25,531
36	2870	113	25,398
37	1629	63	25,857
38	2782	108	25,759
38	2543	95	26,768
40	2959	116	25,508

- Langkah selanjutnya adalah memasukkan hasil tersebut untuk dicari koefisien korelasinya. Perhitungan tersebut dapat dilihat pada tabel 2. hasil perhitungannya adalah sebagai berikut:

TABEL 2
TABEL ANALISIS VALIDITAS BUTIR TES
MATA PELAJARAN AQIDAH AKHLAK KELAS VII

No Btr	Mp	Mt	SDt	$\sqrt{p : q}$	r_{pbis}
1	25,70	25,341	4,816	$\sqrt{0,833:0,167}$	0,165
2	26,768	25,341	4,816	$\sqrt{0,575:0,425}$	0,344
3	25,392	25,341	4,816	$\sqrt{0,891:0,109}$	0,028

4	27,630	25,341	4,816	$\sqrt{0,383:0,617}$	0,374
5	26,112	25,341	4,816	$\sqrt{0,516:0,484}$	0,165
6	24,545	25,341	4,816	$\sqrt{0,091:0,909}$	-0,052
7	26,309	25,341	4,816	$\sqrt{0,591:0,408}$	0,240
8	24,529	25,341	4,816	$\sqrt{0,141:0,859}$	-0,068
9	28,000	25,341	4,816	$\sqrt{0,050:0,950}$	0,126
10	25,989	25,341	4,816	$\sqrt{0,775:0,225}$	0,248
11	25,752	25,341	4,816	$\sqrt{0,808:0,192}$	0,147
12	26,144	25,341	4,816	$\sqrt{0,750:0,250}$	0,287
13	26,734	25,341	4,816	$\sqrt{0,658:0,342}$	0,400
14	26,295	25,341	4,816	$\sqrt{0,733:0,267}$	0,324
15	26,577	25,341	4,816	$\sqrt{0,750:0,250}$	0,307
16	26,365	25,341	4,816	$\sqrt{0,775:0,225}$	0,393
17	27,132	25,341	4,816	$\sqrt{0,691:0,309}$	0,552
18	24,583	25,341	4,816	$\sqrt{0,500:0,500}$	- 0,157
19	26,186	25,341	4,816	$\sqrt{0,625:0,375}$	0,225
20	27,405	25,341	4,816	$\sqrt{0,575:0,425}$	0,497
21	27,214	25,341	4,816	$\sqrt{0,583:0,417}$	0,458
22	24,934	25,341	4,816	$\sqrt{0,133:0,867}$	-0,032
23	24,913	25,341	4,816	$\sqrt{0,383:0,617}$	-0,069
24	27,287	25,341	4,816	$\sqrt{0,608:0,392}$	0,530
25	26,465	25,341	4,816	$\sqrt{0,841:0,159}$	0,535
26	26,935	25,341	4,816	$\sqrt{0,516:0,484}$	0,340
27	27,293	25,341	4,816	$\sqrt{0,625:0,375}$	0,522
28	25,981	25,341	4,816	$\sqrt{0,891:0,109}$	0,377
29	36,903	25,341	4,816	$\sqrt{0,775:0,225}$	0,445
30	25,521	25,341	4,816	$\sqrt{0,975:0,025}$	0,077
31	27,964	25,341	4,816	$\sqrt{0,475:0,525}$	0,514
32	27,492	25,341	4,816	$\sqrt{0,541:0,459}$	0,484
33	27,066	25,341	4,816	$\sqrt{0,375:0,625}$	0,277

34	25,574	25,341	4,816	$\sqrt{0,841:0,159}$	0,110
35	25,531	25,341	4,816	$\sqrt{0,925:0,075}$	0,136
36	25,398	25,341	4,816	$\sqrt{0,941:0,059}$	0,043
37	25,857	25,341	4,816	$\sqrt{0,525:0,475}$	0,112
38	25,759	25,341	4,816	$\sqrt{0,900:0,100}$	0,258
39	26,768	25,341	4,816	$\sqrt{0,791:0,209}$	0,575
40	25,508	25,341	4,816	$\sqrt{0,966:0,034}$	0,181

- Langkah yang selanjutnya adalah menginterpretasikan nilai r_{pbis} dengan r_t (r tabel) ke dalam bentuk tabel analisis, seperti tertera di bawah ini:

No. Butir	r_{pbis}	r_t (r tabel)	Interpretasi
1	0,165	0,176	<i>Invalid</i>
2	0,344	0,176	<i>Valid</i>
3	0,028	0,176	<i>Invalid</i>
4	0,374	0,176	<i>Valid</i>
5	0,165	0,176	<i>Invalid</i>
6	-0,052	0,176	<i>Invalid</i>
7	0,240	0,176	<i>Valid</i>
8	-0,068	0,176	<i>Invalid</i>
9	0,126	0,176	<i>Invalid</i>
10	0,248	0,176	<i>Valid</i>
11	0,147	0,176	<i>Invalid</i>
12	0,287	0,176	<i>Valid</i>
13	0,400	0,176	<i>Valid</i>
14	0,324	0,176	<i>Valid</i>
15	0,307	0,176	<i>Valid</i>
16	0,393	0,176	<i>Valid</i>
17	0,552	0,176	<i>Valid</i>

18	-0,157	0,176	<i>Invalid</i>
19	0,225	0,176	<i>Valid</i>
20	0,497	0,176	<i>Valid</i>
21	0,458	0,176	<i>Valid</i>
22	-0,032	0,176	<i>Invalid</i>
23	-0,069	0,176	<i>Invalid</i>
24	0,530	0,176	<i>Valid</i>
25	0,535	0,176	<i>Valid</i>
26	0,340	0,176	<i>Valid</i>
27	0,522	0,176	<i>Valid</i>
28	0,377	0,176	<i>Valid</i>
29	0,445	0,176	<i>Valid</i>
30	0,077	0,176	<i>Invalid</i>
31	0,514	0,176	<i>Valid</i>
32	0,484	0,176	<i>Valid</i>
33	0,277	0,176	<i>Valid</i>
34	0,110	0,176	<i>Invalid</i>
35	0,136	0,176	<i>Invalid</i>
36	0,043	0,176	<i>Invalid</i>
37	0,112	0,176	<i>Invalid</i>
38	0,258	0,176	<i>Valid</i>
39	0,575	0,176	<i>Valid</i>
40	0,181	0,176	<i>Valid</i>

2. Analisis Reliabilitas Tes

Untuk mencari reliabilitas instrumen tes, harus melalui beberapa langkah diantaranya:

- Menyiapkan tabel perhitungan dalam rangka uji reliabilitas tes dengan menampilkan 40 butir soal, dapat dilihat pada lampiran.
- Mencari jumlah kuadrat total $\sum X_i^2$, dengan menggunakan rumus:

$$\Sigma xt^2 = \Sigma X_t^2 - \frac{(\Sigma X_t)^2}{N}$$

$$\begin{aligned}\Sigma xt^2 &= 79844 - \frac{3041^2}{120} \\ &= 79844 - 77064,008 \\ &= 2779,992\end{aligned}$$

c. Mencari varian total (S_t^2), dengan menggunakan rumus:

$$S_t^2 = \frac{\Sigma xt^2}{N}$$

Diketahui :

$$\Sigma xt^2 = 2779,992$$

$$N = 120$$

$$\begin{aligned}S_t^2 &= \frac{2779,992}{120} \\ &= 23,166\end{aligned}$$

Dalam mencari (menghitung) reliabilitas instrumen tes ini, peneliti menggunakan rumus K-R 20, dengan rumus:

$$r_{11} = \left[\frac{n}{n-1} \right] \left[- \frac{S_t^2 - p_i q_i}{S_t^2} \right]$$

Keterangan:

r_{11} = Koefisien reliabilitas tes

n = Banyaknya butir soal

1 = Bilangan konstan

S_t^2 = Varian total

p_i = Proporsi testee yang menjawab betul

q_i = Proporsi testee yang menjawab salah butir soal.

$p_i q_i$ = Jumlah dari hasil perkalian antara $p_i q_i$

diketahui :

$$n = 40$$

$$S_t^2 = 23,166$$

$$\Sigma p_i q_i = 6,930$$

$$\begin{aligned} r_{11} &= \left[\frac{40}{40-1} \right] \left[\frac{23,166-6,930}{23,166} \right] \\ &= (1,025)(0,700) \\ &= 0,717 \end{aligned}$$

3. Analisis Tingkat Kesukaran Butir

Suatu butir soal tes hasil belajar dapat dikatakan bermutu atau tidak, dapat dilihat dari derajat kesukarannya atau taraf kesukaran yang ada pada masing-masing butir soal yang membangun tes hasil belajar tersebut.

Dan hasil uji atau analisis yang peneliti lakukan, didapatkan tabulasi data dan perhitungan angka indeks kesukaran butir soal, seperti terlihat pada lampiran IV, maka dari itu dapat kita peroleh informasi tentang derajat kesukaran butir soal tes sumatif kelas VII untuk mata pelajaran Aqidah Akhlak semester 1 tahun 2008/2009 seperti tertera pada tabel 3. berikut ini :

TABEL 3
DERAJAT KESUKARAN BUTIR SOAL MATA PELAJARAN
AQIDAH AKHLAK KELAS VII MTs MIFTAHUSSALAM 1
WONOSALAM DEMAK

No.	Derajat Kesukaran Soal	Jumlah	Prosentase
1	Terlalu sukar	4	10 %
2	Cukup (sedang)	21	52,5 %
3	Terlalu mudah	15	37,5 %

4. Analisis Daya Pembeda Butir

Suatu butir soal tes jika dilihat dari Daya Pembeda dapat dikatakan baik apabila butir soal tersebut mampu membedakan mereka (testee) yang kita anggap pandai dengan testee yang bodoh. Maksudnya jawaban pada testee pada umumnya adalah betul, sedangkan testee yang bodoh jawaban pada umumnya adalah salah. Artinya butir soal yang bersangkutan telah menunjukkan kemampuan didalam membedakan antara testee yang tergolong pandai dengan testee yang tergolong bodoh.

Berdasarkan hasil analisis soal-soal tes sumatif mata pelajaran Aqidah Akhlak semester 1 kelas VII MTs Miftahussalam 1 Wonosalam Demak tahun 2008/2009 yang dapat dilihat pada lampiran V dan hasilnya tertera lampiran VI, maka bersumber dari data yang disajikan dapat diperoleh informasi sebagaimana tertera pada tabel 4, berikut dibawah ini:

TABEL 4
DAYA PEMBEDA BUTIR SOAL TES OBJEKTIF DALAM TES
SEMESTER 1 KELAS VII MATA PELAJARAN AQIDAH AKHLAK
MTs MIFTAHUSSALAM WONOSALAM 1 DEMAK

No.	Daya Pembeda Soal	Jumlah	Prosentase
1	Lemah sekali	4	10 %
2	Lemah	12	30 %
3	Cukup/sedang	12	30 %
4	Baik	10	25 %
5	Baik sekali	2	5 %

5. Analisis Fungsi Distraktor

Menganalisis fungsi distraktor dikenal pula dengan istilah lain, yaitu: menganalisis pola penyebaran jawaban soal, yaitu suatu pola yang dapat menggambarkan bagaimana testee menentukan pilihan jawabannya

terhadap kemungkinan-kemungkinan (*option*) yang telah dipasangkan pada setiap butir soal.

Berdasarkan analisis terhadap perhitungan banyaknya testee yang memilih option/alternatif jawaban soal tes objektif bentuk pilihan ganda, model melengkapi dengan empat pilihan jawaban A, B, C, dan D pada tes objektif semester satu mata pelajaran Aqidah Akhlak MTs Miftahussalam 1 Wonosalam Demak tahun 2008/2009. Penyebaran data dapat dilihat pada lampiran VII. Maka bersumber dari data yang disajikan itu dapat diperoleh informasi sebagaimana tertera pada tabel 3, berikut di bawah ini:

TABEL 5
DATA TENTANG BERFUNGSI TIDAKNYA DISTRAKTOR
PADA TES OBJEKTIF MODEL MCI DALAM TES SUMATIF
KELAS X MATA PELAJARAN PENDIDIKAN AGAMA ISLAM
SISWA SMA NEGERI 7 SEMARANG

No.	Kondisi Distraktor	Jumlah	Prosentase
1	Telah berfungsi dengan baik	89	74,166 %
2	Tidak berfungsi dengan baik	31	25,833 %

C. HASIL DAN PEMBAHASAN PENELITIAN

1. Dari Segi Validitas Butirnya

Hasil analisis terhadap butir-bitir soal tes objektif dalam tes sumatif mata pelajaran Aqidah Akhlak semester I MTs Miftahussalam 1 Wonosalam Demak yang hasilnya dapat dilihat pada tabel 1 di atas, maka bersumber dari data itu penulis mendapatkan keterangan mengenai seberapa besar prosentase butir soal yang termasuk valid dan invalid.

Berdasarkan data pada tabel 1 di atas dan pada perhitungan yang juga tertera pada tabel 2, dapat diketahui bahwa sebanyak 25 butir soal

atau sekitar 62,5% butir-butir soal pada tes tersebut yang dinyatakan memiliki validitas, sedangkan 15 butir soal lainnya dinyatakan dalam kategori invalid atau sekitar 37,5%. Jadi dapat ditarik ketahu bahwa tes sumatif kelas VII pada mata pelajaran Aqidah Akhlak semester I tersebut memiliki validitas sedang.

2. Ditinjau Dari Segi Reliabilitas Tes

Suatu tes sebagai alat ukur baru dapat dikatakan memiliki reliabilitas sempurna apabila hasil-hasil pengukuran yang dilakukan secara berulang kali kepada subjek yang sama atau sejenis, senantiasa menunjukkan hasil (dalam hal ini adalah skor atau nilai) yang tetap sama (sifatnya ajeg atau stabil).

Berdasarkan perhitungan di atas, yaitu pada uji reliabilitas tes dimana koefisien reliabilitas (r_{11}) sebesar 0,717, jadi penulis nyatakan bahwa pada tes hasil belajar bentuk objektif yang menyajikan 40 butir soal dalam tes sumatif mata pelajaran Aqidah Akhlak semester I di MTs Miftahussalam 1 Wonosalam Demak yang diikuti 120 testee tahun 2008/2009 memiliki reliabilitas yang tinggi.

3. Dari Segi Tingkat Kesukaran

Berdasarkan tabel 3 di atas, dapat diketahui tingkat kesukaran dari butir tes yang diujikan di Mts Miftahussalam 1 Wonosalam Demak pada semester I mata pelajaran Aqidah Akhlak termasuk instrumen tes yang memiliki derajat kesukaran yang **sedang**. Hal ini dilihat dari 40 butir soal yang disajikan terdapat 15 atau 37,5% butir tes yang terlalu mudah dikerjakan, 21 atau 52,5% butir tergolong butir tes yang sedang atau cukup dan hanya 4 atau 10% butir tes yang tergolong butir tes yang terlalu sulit.

Meskipun instrumen tes tersebut telah memiliki tingkat kesukaran yang memadai, namun proporsi jumlah butir tes yang terlalu sulit, sedang dan terlalu mudah masih belum bisa dikatakan baik. Dengan melihat

tingkat kesukaran pada instrumen tersebut memiliki tingkat kesukaran yang sedang maka dapat pula diketahui bahwa siswa MTs Miftahussalam Wonosalam demak yang mengikuti ujian tes mata pelajaran Aqidah Akhlak tahun 2008/2009 memiliki tingkat kemampuan yang sedang.

4. Dari Segi Daya Pembedanya

Berdasarkan pada lampiran V dan tabel 4 di atas, dapat diketahui bahwa soal-soal tes hasil belajar tersebut, jika dilihat dari Daya Pembeda butirnya menunjukkan presentase 10% mempunyai daya pembeda yang lemahse kali, ada juga butir soal tes yang mempunyai daya pembeda yang lemah yaitu 30%, terdapat juga soal-soal tes tersebut yang sedang daya pembedanya yaitu sebanyak 30%, baik 25%, dan baik sekali 5%.

Jadi sebanyak 60% memiliki daya pembeda yang memadai sehingga dapat dikatakan bahwa untuk tes sumatif mata pelajaran Aqidah Akhlak kelas VII semester I tahun 2008/2009 ini termasuk tes yang memiliki daya pembeda cukup/sedang.

5. Dari Segi Fungsi Distraktor

Atas dasar analisis perhitungan pada lampiran VII dan tabel 5 yaitu tabel tentang kondisi distraktor dapat diketahui bahwa pemasangan distraktor pada butir soal tes tersebut sebagian telah dapat menarik atau merangsang testee yang mengikuti tes tersebut untuk tidak memilih yang bukan sebagai kunci jawaban butir soal; artinya sebagian testee terkecoh, karena 74,166% dari option distraktor yang ditampilkan dapat berfungsi dengan baik. Maka penulis dapat ditampilkan dalam tes tersebut dalam kategori memiliki fungsi distraktor yang baik.

D. KETERBATASAN PENELITIAN

Hasil penelitian yang telah peneliti lakukan secara optimal, peneliti sadari adanya beberapa keterbatasan, yaitu dilihat dari segi metode pengumpulan data, banyak sekali metode tentang pengumpulan data yang

peneliti tidak dapat terangkan dan kuasai satu persatu. Dilihat dari segi analisis data yang telah peneliti peroleh, peneliti hanya mengolah data tersebut menggunakan salah satu tekni yang peneliti kuasai dan menyajikannya dalam keterbatasan pemahaman peneliti. Dilihat dari subjek penelitian yang dalam penelitian ini peneliti tidak melakukan pemilihan subyek yang lebih banyak, tetapi peneliti menggunakan subjek atau populasi yang tersedia.

BAB V

PENUTUP

A. SIMPULAN

Setelah penulis menganalisa data yang berhasil dihimpun dalam penelitian ini, untuk mengetahui tes dalam bentuk objektif semester I pada mata pelajaran Aqidah Akhlak di MTs Miftahussalam Wonosalam Demak tahun 2008/2009 tersebut, maka penulis dapat mengetahui bahwa tes yang dianalisa pada penelitian ini adalah tes objektif yang terdiri dari 40 butir soal. Selanjutnya dapat menarik kesimpulan sebagai berikut:

1. Validitas tes pada mata pelajaran Aqidah Akhlak kelas VII MTs Miftahussalam Wonosalam Demak tahun 2008/2009 ini termasuk dalam kategori tes yang memiliki validitas yang sedang, yaitu dilihat dari hasil perhitungan pada table 2 sebanyak 25 butir soal atau sekitar 62,5% butir soal dinyatakan valid, dan 15 butir soal atau sekitar 37,5% dikategorikan invalid.
2. Reliabilitas tes mata pelajaran Aqidah Akhlak kelas VII semester I MTs Miftahussalam Wonosalam Demak tahun 2008/2009 ini termasuk dalam kategori tes yang memiliki reliabilitas yang tinggi, hal ini disebabkan karena dilihat dari koefisien reliabilitasnya sebesar:
$$r_{11} = 0,717$$
 dalam pemberian interpretasi terhadap angka korelasi koefisien, interval 0,70- 0,90 dalam interpretasi yang tinggi.
3. Tingkat kesukaran tes mata pelajaran Aqidah Akhlak semester I MTs Miftahussalam wonosalam Demak tahun 2008/2009 memiliki derajat kesukaran yang sedang. Hal ini dilihat dari 40 butir soal yang disajikan terdapat 15 atau 37,5% butir tes yang terlalu mudah dikerjakan, 21 atau 52,5% butir tergolong butir tes yang sedang atau cukup dan hanya 4 atau 10% butir tes yang tergolong butir tes yang terlalu sulit.
4. Dari segi daya pembeda tes mata pelajaran Aqidah Akhlak kelas VII semester I MTs Miftahussalam Wonosalam Demak tahun 2008/2009

ini memiliki daya pembeda yang sedang, karena sekitar 60% butir tes memiliki daya pembeda yang memadai.

5. Ditinjau dari segi fungsi distraktor tes ini memiliki termasuk tes yang memiliki fungsi distraktor yang baik, karena sebanyak 74,166% dari 120 jawaban pengecoh telah berfungsi dengan baik.
6. Dapat dilihat pada tabel 6 yang menghubungkan antara validitas butir, tingkat kesukaran butir, dan daya pembeda butir.

TABEL 6
TABEL INTERPRETASI VALIDITAS, TINGKAT KESUKARAN,
DAN DAYA PEMBEDA BUTIR

No. Btr Soal	Interpretasi Validitas Butir	Interpretasi Tingkat Kesukaran	Interpretasi Daya Pembeda
1	<i>Invalid</i>	Terlalu Mudah	Sedang
2	<i>Valid</i>	Sedang	Baik
3	<i>Invalid</i>	Terlalu Mudah	Lemah
4	<i>Valid</i>	Sedang	Sedang
5	<i>Invalid</i>	Sedang	Lemah
6	<i>Invalid</i>	Terlalu Sukar	Lemah Sekali
7	<i>Valid</i>	Sedang	Sedang
8	<i>Invalid</i>	Terlalu Sukar	Lemah Sekali
9	<i>Invalid</i>	Terlalu Sukar	Lemah
10	<i>Valid</i>	Terlalu Mudah	Sedang
11	<i>Invalid</i>	Terlalu Mudah	Lemah
12	<i>Valid</i>	Sedang	Sedang
13	<i>Valid</i>	Sedang	Baik
14	<i>Valid</i>	Sedang	Baik
15	<i>Valid</i>	Sedang	Sedang

16	<i>Valid</i>	Terlalu Mudah	Sedang
17	<i>Valid</i>	Sedang	Baik Sekali
18	<i>Invalid</i>	Sedang	Lemah
19	<i>Valid</i>	Sedang	Sedang
20	<i>Valid</i>	Sedang	Baik Sekali
21	<i>Valid</i>	Sedang	Baik
22	<i>Invalid</i>	Terlalu Sukar	Lemah Sekali
23	<i>Invalid</i>	Sedang	Lemah Sekali
24	<i>Valid</i>	Sedang	Sedang
25	<i>Valid</i>	Terlalu Mudah	Baik
26	<i>Valid</i>	Sedang	Baik
27	<i>Valid</i>	Sedang	Baik
28	<i>Valid</i>	Terlalu Mudah	Sedang
29	<i>Valid</i>	Terlalu Mudah	Sedang
30	<i>Invalid</i>	Terlalu Mudah	Lemah
31	<i>Valid</i>	Sedang	Baik
32	<i>Valid</i>	Sedang	Baik
33	<i>Valid</i>	Sedang	Sedang
34	<i>Invalid</i>	Terlalu Mudah	Lemah
35	<i>Invalid</i>	Terlalu Mudah	Lemah
36	<i>Invalid</i>	Terlalu Mudah	Lemah
37	<i>Invalid</i>	Sedang	Lemah
38	<i>Valid</i>	Terlalu Mudah	Lemah
39	<i>Valid</i>	Terlalu Mudah	Baik
40	<i>Valid</i>	Terlalu Mudah	Lemah

Dari data tabel dapat disimpulkan bahwa interpretasi butir tes yang invalid belum tentu berinterpretasi pada tingkat kesukaran yang terlalu sukar, dan berinterpretasi pada daya pembeda yang lemah.

B. SARAN – SARAN

1. Tim Penyusun Soal

- a. seyogyanya butir – butir tes hasil belajar menampilkan sample yang representatif dari populasi bahan yang telah diajarkan, sehingga dapat mewakili seluruh performance yang telah diperoleh siswa dalam mengikuti suatu unit pelajaran.
- b. Hendaknya dalam membuat soal – soal tes susunan kalimatnya diatur sedemikian rupa sehingga berpeluang menampilkan option yang satu dengan yang lain identik tapi tidak sama.
- c. Hendaknya bentuk soal yang dikeluarkan dalam tes hasil belajar dibuat bervariasi sehingga betul – betul cocok untuk mengukur hasil belajar yang diinginkan sesuai dengan tujuan tes itu sendiri.
- d. Tes hasil belajar tersebut hendaknya didesain sesuai dengan kegunaannya untuk memperoleh hasil yang diinginkan.

2. Guru Bidang Studi Aqidah Akhlak

- a. Hendaknya senantiasa mencermati butir – butir soal yang termasuk dalam kategori belum memadai untuk diteliti ulang dan dilakukan perbaikan – perbaikan, dan penyempurnaan kembali butir – butir soal yang dianggap perlu.
- b. Adalah sangat berharga sekali seandainya selalu mengusahakan evaluasi dan tindak lanjut terhadap hasil tes yang telah diujikan dengan konsentrasi harus memiliki kemampuan dan bekal pengalaman yang memadai dalam mengevaluasi tes tersebut.

C. PENUTUP

Segala puji dan syukur kepada Allah SWT, atas limpahan rahmat, taufik, dan hidayah, serta inayah-Nya, penulis dapat menyelesaikan skripsi ini yang berjudul : *“Analisis Instrumen Tes Mata Pelajaran Aqidah Akhlak Buatan Guru KKM Rayon 2 MTs Miftahussalam 1 Wonosalam Demak Tahun 2008/2009”*.

Dengan segenap kemampuan dan pengetahuan yang terbatas, penulis mencoba menyusun skripsi ini dan penulis juga menyadari sepenuhnya bahwa penelitian ini masih jauh dari kesempurnaan, baik dalam keterbatasan peneliti dalam memahami, menelaah dan mengoreksi data – data yang ada yang berkaitan dengan skripsi ini. Untuk itu kritik dan saran adalah merupakan hal yang senan tiasa dan niscaya penulis harapkan demi kesempurnaan penelitian ini.

Semoga skripsi ini dapat memberikan manfaat dan dapat memberikan kebaikan pada semua pihak; baik bagi peneliti pribadi maupun kepada para pembaca. *Amiin ya Robbal 'aalamiin.*

DAFTAR PUSTAKA

- Anastasi, Anne dan Susana, Urbina, *Psychological Testing*, New York : Prentice-Hall, 1988.
- Arikunto, Suharsimi, *Dasar-Dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 1990.
-, *Pengantar Evaluasi Pendidikan*, Jakarta : Raja Grafindo, 2003.
- Arikunto, Suharsimi, *Prosedur Penelitian (Suatu Pendekatan Praktik)*, Jakarta: Rineka Cipta, 2002, Cet. XII.
- Arikunto, Suharsimi, *Prosedur Penelitian (Suatu Pendekatan Praktik)*, Jakarta: Melton Putra, 1991.
- Azwar, Syaifuddin, *Tes Prestasi (Fungsi dan Pengembangan Pengukuran Prestasi Belajar)*, Yogyakarta: Liberty, 1987.
- Brown, Douglas, *Ususu At-Ta'alimu Al-Lughoti wa Ta'limiha*, Terj. Abdul Ar-Rojhi dan Ali Ahmad Syu'ban, Arab : Darun An-Nahdhoh, tt.
- Chan, Sam, "Analisis Swot Kebijakan Era Pendidikan",
<http://blog.unila.ac.id/redho/2009/03/26/analisa-metodologi-penelitian-sosial>.
- Cronbach, Lee j., *Essentials of Psychological testing*, new York : Harper & Row Publishers, 1960.
- Daryanto, *Evaluasi Pendidikan*, Jakarta, Rineka Cipta, 2001.
- Depdiknas, *Kamus Besar Bahasa Indonesia*, Jakarta: Balai Pustaka, 2003.
- Djiwandono, Sri Esti Wuryani, *Psikologi Pendidikan*, Jakarta : Grasindo, 2002.
- Hadi, Sutrisno, *Metodologi Research*, Yogyakarta, Andi Offset, 1995.
- Hajar, Ibnu, *Dasar-Dasar Metodologi Penelitian Kualitatif Dalam Pendidikan*, Jakarta: Raja Grafindo, 1999.
- Harjanto, *Perencanaan Pengajaran*, Jakarta: Rineka cipta, 1997.
- Kontjoroningrat, *Metode Penelitian Masyarakat*, Jakarta: Gramedia, tt.

Margono, *Metodologi Penelitian Pendidikan*, Jakarta, Rineka Cipta, 2000, Cet. II.

Mawardi, "Validitas dan reliabilitas instrumen tes buatan guru pelajaran fiqih kelas VII MTs Al – ma'arif NU Boja Kendal", Skripsi S1 IAIN Walisongo Semarang, Semarang: Perpustakaan Tarbiyah, IAIN Walisongo Semarang, 2007, t.d.

Mudjijo, *Tes Hasil Belajar*, Jakarta: Bumi Aksara, 1995.

Muhammad, Muhammad Abdul Kholik, *Ihtibarotun Al-Lughoh*, Jami'ah Malik Su'ud: 1989.

Nawawi, Hadari dan Hadari, Martini, *Instrumen Penelitian Bidang Sosial*, Yogyakarta: Gajah Mada University Press, 1992.

Nuridin, Mohamad, "Uji Validitas dan Reliabilitas tes PAI Mata Pelajaran Akhlak di SMP Muhammadiyah 08 Mijen Semarang Tahun 2003/2004", Skripsi S1 IAIN Walisongo Semarang, Semarang: Perpustakaan Tarbiyah, IAIN Walisongo Semarang, 2004, t.d.

Nurkancana, Wayan, *Evaluasi Pendidikan*, Jakarta: Usaha Nasional, 1986.

Purwanto, Ngalm, *Prinsip – Prinsip dan Teknik Evaluasi Pengajaran*, Bandung : Remaja Rosdakarya, 2002, Cet. II.

Skinner, Charles E. (ed.), *Essentials of educational Psychology*, (Englewood Cliffs : Prentice-Hall, Inc, tt.

Slameto, *Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 2001.

Sudijono, Anas, *Pengantar Statistik Pendidikan*, Jakarta : Rajawali, 1991.

sudjana, Nana ,*Penilaian Hasil Proses Belajar Mengajar*, Bandung : Remaja Rosdakarya, 1999.

Sujiono, Anas, *Pengantar Evaluasi Pendidikan*, Jakarta: Raja Grafindo, 2003.

....., *Prosedur penelitian*, Jakarta: Rineka Cipta, 1993.

Sukardi, *Metodologi Penelitian Pendidikan (Kompetensi dan Praktiknya)*, Jakarta: Bumi Aksara, 2003.

Suryobroto, Sukardi, *Pembimbing ke Psikodoagnostik*, Yogyakarta: Andi Offset, 1984.

Thoha, Chabib dan Mu'ti, Abdul PBM PAI di Sekolah (*Eksistensi dan Proses Belajar mengajar Pendidikan Agama Islam*, Yogyakarta: Pustaka Pelajar, 1998.

Thoha, Chabib, *Teknik evaluasi Pendidikan*, Jakarta: Rajawali, 1991.

Tim Penyusun Kamus Pembinaan dan Pengembangan Bahasa, *Kamus Besar Bahasa Indonesia*, Jakarta: Balai Pustaka, 1997, Cet. IX.

Winkel, W.S. S.J., *Psikologi Pendidikan dan Evaluasi Belajar*, Jakarta : Gramedia, 1983.

Worthen, Blaine R. and Sanders, James R., *Educational Evaluation*, New York: Longman, 1980.

DAFTAR RIWAYAT HIDUP

Nama : **MAKRUF**

Tempat Tanggal Lahir : Demak, 10 Nopember 1985

Alamat Asal : Trimulyo RT.01/01 Kec. Gunyur Kab. Demak

Jenjang Pendidikan :

1. SD Trimulyo 1 Guntur Demak, Lulus Tahun 1996
2. MTs Sabilurrahman Ringinharjo Gubug Grobogan, Lulus Tahun 1999.
3. MA Miftahussalam Wonosalam Demak, Lulus Tahun 2002.
4. IAIN Walisongo Semarang Angkatan Tahun 2002.

LAMPIRAN 4

HASIL PERHITUNGAN TINGKAT KESUKARAN SOAL TES SUMATIF

KELAS VII MTs MIFTAHUSSALAM 1 WONOSALAM DEMAK

N0	Jml Testee yang menjawab betul	N	TK	Interpretasi
1	100	120	0,833	Terlalu Mudah
2	69	120	0,575	Sedang
3	107	120	0,891	Terlalu Mudah
4	46	120	0,383	Sedang
5	62	120	0,516	Sedang
6	11	120	0,091	Terlalu Sukar
7	71	120	0,591	Sedang
8	17	120	0,141	Terlalu Sukar
9	6	120	0,050	Terlalu Sukar
10	93	120	0,775	Terlalu Mudah
11	97	120	0,808	Terlalu Mudah
12	90	120	0,750	Sedang
13	79	120	0,658	Sedang
14	88	120	0,733	Sedang
15	90	120	0,750	Sedang
16	93	120	0,775	Terlalu Mudah
17	83	120	0,691	Sedang
18	60	120	0,500	Sedang
19	75	120	0,625	Sedang
20	69	120	0,575	Sedang
21	70	120	0,583	Sedang
22	16	120	0,133	Terlalu Sukar
23	46	120	0,383	Sedang
24	73	120	0,608	Sedang
25	101	120	0,841	Terlalu Mudah
26	62	120	0,516	Sedang

27	75	120	0,625	Sedang
28	107	120	0,891	Terlalu Mudah
29	93	120	0,775	Terlalu Mudah
30	117	120	0,975	Terlalu Mudah
31	57	120	0,475	Sedang
32	65	120	0,541	Sedang
33	45	120	0,375	Sedang
34	101	120	0,841	Terlalu Mudah
35	111	120	0,925	Terlalu Mudah
36	113	120	0,941	Terlalu Mudah
37	63	120	0,525	Sedang
38	108	120	0,900	Terlalu Mudah
39	95	120	0,791	Terlalu Mudah
40	116	120	0,966	Terlalu Mudah

LAMPIRAN 6

HASIL PERHITUNGAN DAYA PEMBEDA SOAL

PELAJARTAN AQIDAH AKHLAK KELAS VII

MTs MIFTAHUSSALAM 1 WONOSALAM DEMAK

No Btr Soal	BA	BB	JA	JB	PA= BA/JB	PB= BB/JA	DB= PA-PB	Interpretasi
1	31	23	33	33	0,939	0,696	0,243	Sedang
2	28	7	33	33	0,848	0,212	0,636	Baik
3	31	29	33	33	0,939	0,878	0,061	Lemah
4	19	6	33	33	0,575	0,181	0,394	Sedang
5	17	14	33	33	0,515	0,424	0,091	Lemah
6	1	3	33	33	0,030	0,090	- 0,060	Lemah Sekali
7	25	14	33	33	0,757	0,424	0,333	Sedang
8	2	6	33	33	0,060	0,181	- 0,121	Lemah Sekali
9	3	0	33	33	0,909	0,000	0,090	Lemah
10	26	20	33	33	0,787	0,060	0,272	Sedang
11	28	23	33	33	0,848	0,696	0,152	Lemah
12	27	18	33	33	0,818	0,545	0,273	Sedang
13	27	13	33	33	0,818	0,393	0,425	Baik
14	29	14	33	33	0,878	0,424	0,454	Baik
15	29	16	33	33	0,878	0,484	0,394	Sedang
16	33	20	33	33	1,000	0,606	0,394	Sedang
17	33	9	33	33	1,000	0,272	0,728	Baik Sekali
18	21	15	33	33	0,636	0,454	0,182	Lemah
19	27	14	33	33	0,818	0,424	0,394	Sedang
20	32	8	33	33	0,969	0,242	0,727	Baik Sekali
21	31	10	33	33	0,939	0,303	0,636	Baik
22	1	4	33	33	0,030	0,121	- 0,091	Lemah Sekali
23	12	13	33	33	0,363	0,393	- 0,030	Lemah Sekali

24	26	14	33	33	0,787	0,424	0,363	Sedang
25	32	18	33	33	0,969	0,545	0,425	Baik
26	26	9	33	33	0,787	0,272	0,515	Baik
27	29	10	33	33	0,878	0,303	0,575	Baik
28	32	25	33	33	0,969	0,757	0,212	Sedang
29	32	22	33	33	0,969	0,666	0,303	Sedang
30	33	30	33	33	1,000	0,909	0,091	Lemah
31	25	8	33	33	0,757	0,242	0,515	Baik
32	28	9	33	33	0,848	0,272	0,576	Baik
33	18	5	33	33	0,545	0,151	0,394	Sedang
34	31	28	33	33	0,939	0,848	0,091	Lemah
35	33	30	33	33	1,000	0,909	0,091	Lemah
36	32	31	33	33	0,969	0,939	0,030	Lemah
37	19	18	33	33	0,575	0,545	0,030	Lemah
38	32	29	33	33	0,969	0,878	0,091	Lemah
39	33	18	33	33	1,000	0,545	0,454	Baik
40	33	31	33	33	1,000	0,939	0,061	Lemah

LAMPIRAN 7

**HASIL PERHITUNGAN FUNGSI DISTRAKTOR OPTION BUTIR TES
SUMATIF MATAPELAJARAN AQIDAH AKHLAK KELAS VII MTs
MIFTAHUSSALAM WONOSALAM 1 WONOSALAM DEMAK**

No. Butir	Option	Jml Testee yang Memilih Option	Interpretasi
1	A	9	Baik
	B	4	Kurang Baik
	C*		
	D	7	Baik
2	A	27	Baik
	B*		
	C	14	Baik
	D	11	Baik
3	A	5	Kurang Baik
	B	1	Kurang Baik
	C*		
	D	39	Baik
4	A	46	Baik
	B	17	Baik
	C	11	Baik
	D*		
5	A	12	Baik
	B*		
	C	41	Baik
	D	6	Baik
6	A	63	Baik

	B	26	Baik
	C*		
	D	20	Baik
7	A*		
	B	17	Baik
	C	4	Kurang Baik
	D	20	Baik
8	A	17	Baik
	B	20	Baik
	C	67	Baik
	D*		
9	A	49	Baik
	B	32	Baik
	C*		
	D	33	Baik
10	A	6	Baik
	B*		
	C	5	Kurang Baik
	D	17	Baik
11	A*		
	B	6	Baik
	C	2	Kurang Baik
	D	20	Baik
12	A	6	Baik
	B*		
	C	14	Baik
	D	12	Baik
13	A	26	Baik
	B	4	Kurang Baik
	C	10	Baik

	D*		
14	A	12	Baik
	B	9	Baik
	C*		
	D	38	Baik
15	A*		
	B	7	Baik
	C	10	Baik
	D	14	Baik
16	A	13	Baik
	B	7	Baik
	C*		
	D	10	Baik
17	A	21	Baik
	B	7	Baik
	C	11	Baik
	D*		
18	A*		
	B	7	Baik
	C	4	Kurang Baik
	D	50	Baik
19	A	11	Baik
	B*		
	C	27	Baik
	D	8	Baik
20	A	29	Baik
	B*		
	C	13	Baik
	D	9	Baik
21	A	26	Baik

	B	17	Baik
	C	11	Baik
	D*		
22	A	32	Baik
	B	40	Baik
	C*		
	D	32	Baik
23	A*		
	B	45	Baik
	C	14	Baik
	D	16	Baik
24	A	11	Baik
	B*		
	C	23	Baik
	D	13	Baik
25	A	8	Baik
	B	5	Kurang Baik
	C*		
	D	6	Baik
26	A	8	Baik
	B	26	Baik
	C*		
	D	23	Baik
27	A	10	Baik
	B	28	Baik
	C	8	Baik
	D*		
28	A	6	Baik
	B	4	Kurang Baik
	C	5	Kurang Baik

	D*		
29	A	8	Baik
	B	12	Baik
	C*		
	D	69	Baik
30	A	1	Kurang Baik
	B	2	Kurang Baik
	C	1	Kurang Baik
	D*		
31	A	43	Baik
	B*		
	C	16	Baik
	D	5	Kurang Baik
32	A*		
	B	15	Baik
	C	13	Baik
	D	25	Baik
33	A	8	Baik
	B*		
	C	31	Baik
	D	36	Baik
34	A	1	Kurang Baik
	B	9	Baik
	C*		
	D	9	Baik
35	A	4	Kurang Baik
	B	4	Kurang Baik
	C*		
	D	2	Kurang Baik

36	A	5	Kurang Baik
	B*		
	C	2	Kurang Baik
	D	1	Kurang Baik
37	A	5	Kurang Baik
	B*		
	C	2	Kurang Baik
	D	51	Baik
38	A	4	Kurang Baik
	B	7	Baik
	C	2	Kurang Baik
	D*		
39	A*		
	B	23	Baik
	C	1	Kurang Baik
	D	-	Kurang Baik
40	A*		
	B	3	Kurang Baik
	C	-	Kurang Baik
	D	2	Kurang Baik

Keterangan:

*) = Kunci jawaban.