

**ANALISIS KUALITAS TES SUMATIF CAWU I BIDANG STUDI
BAHASA ARAB KELAS DUA DI MAN YOGYAKARTA II
TAHUN PELAJARAN 1999/2000**



SKRIPSI

Diajukan Kepada Fakultas Tarbiyah
Institut Agama Islam Negeri Sunan Kalijaga
Guna Memenuhi Salah Satu Syarat Untuk
Memperoleh Gelar Sarjana Strata Satu Agama

Oleh :

STATE ISLAMIC UNIVERSITY
MUHAMMAD TAUWAB
NIM. 9642 3394
SUNAN KALIJAGA
YOGYAKARTA

**JURUSAN PENDIDIKAN BAHASA ARAB
INSTITUT AGAMA ISLAM NEGERI
SUNAN KALIJAGA
YOGYAKARTA
2000**

Drs. Tasman Hamami, M.A.
Dosen Fakultas Tarbiyah
IAIN Sunan Kalijaga Yogyakarta

NOTA DINAS

Hal : Skripsi Saudara
Muhammad Tauwab
Lamp : 8 eksemplar

Kepada
Yth. Dekan Fakultas Tarbiyah
IAIN Sunan Kalijaga
Yogyakarta

Assalamu 'alaikum warahmatullahi wabarakatuhu.

Kami selaku pembimbing, setelah membaca dan mengadakan perbaikan seperlunya, maka skripsi saudara;

Nama : Muhammad Tauwab
NIM : 96423394
Jurusan : Pendidikan Bahasa Arab
Judul : ANALISIS KUALITAS TES SUMATIF CAWU I BIDANG STUDI
BAHASA ARAB KELAS DUA DI MAN YOGYAKARTA II TAHUN
PELAJARAN 1999/2000

Sudah dapat diajukan sebagai syarat untuk memperoleh gelar Sarjana Strata Satu Agama pada Fakultas Tarbiyah IAIN Sunan Kalijaga Yogyakarta.

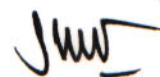
Selanjutnya kami mengharapkan dalam waktu singkat dapat segera diajukan dalam sidang munaqasyah.

Demikian harapan kami dan terima kasih atas perhatiannya.

Wassalamu 'alaikum warahmatullahi wabarakatuhu.

Yogyakarta, 22 April 2000

Dosen Pembimbing,



Drs. Tasman Hamami, M.A.

NIP. 150 226 626

Prof. Drs. Anas Sudijono
Dosen Fakultas Tarbiyah
IAIN Sunan Kalijaga Yogyakarta

NOTA DINAS KONSULTAN

Hal : Skripsi Saudara
Muhammad Tauwab
Lamp : 8 (delapan) Eksemplar

Kepada
Yth. Dekan Fakultas Tarbiyah
IAIN Sunan Kalijaga
Yogyakarta

Assalamu'alaikum warahmatullahi wabarakatuhu

Setelah membaca, meneliti dan mengadakan perbaikan seperlunya, maka kami selaku konsultan skripsi saudara :

Nama : Muhammad Tauwab
NIM : 96423394
Jurusan : Pendidikan Bahasa Arab
Judul : **ANALISIS KUALITAS TES SUMATIF CAWU I BIDANG
STUDI BAHASA ARAB KELAS DUA DI MAN
YOGYAKARTA II TAHUN PELAJARAN 1999/2000**

berpendapat bahwa skripsi tersebut sudah dapat diterima sebagai salah satu syarat guna memperoleh gelar Sarjana Strata Satu (S1) Agama pada Fakultas Tarbiyah IAIN Sunan Kalijaga Yogyakarta.

Selanjutnya kami mengharapkan agar skripsi tersebut disyahkan oleh dewan sidang monaqosah.

Demikian harapan kami dan terima kasih atas perhatiannya.

Wassalamu'alaikum warahmatullahi wabarakatuhu

Yogyakarta, 4 Mei 2000

Konsultan,


Prof. Drs. Anas Sudijono

NIP. 150 028 774

HALAMAN PENGESAHAN

Bahwa skripsi berjudul
**ANALISIS KUALITAS TES SUMATIF CAWU I BIDANG STUDI
BAHASA ARAB KELAS DUA DI MAN YOGYAKARTA II
TAHUN PELAJARAN 1999/2000**

Yang dipersiapkan oleh:

Muhammad Tauwab

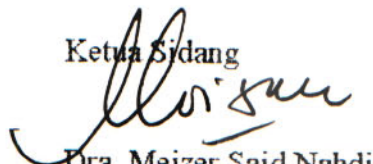
Telah dimunaqosyahkan di depan sidang munaqosyah

Pada hari Jumat, tanggal 28 April 2000

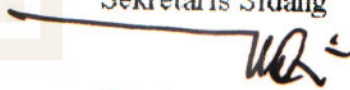
Dan dinyatakan telah memenuhi syarat untuk diterima oleh

Dewan Sidang Munaqosyah

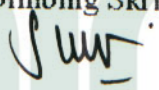
Ketua Sidang


Dra. Meizer Said Nahdi
NIP. 150 219 153

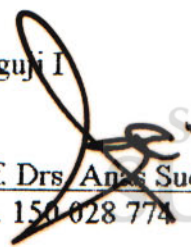
Sekretaris Sidang


Drs. Muhammad Jamroh
NIP. 150 223 031

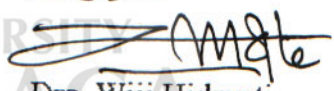
Pembimbing Skripsi


Drs. Tasman Hamami, MA.
NIP. 150 226 626

Penguji I


Prof. Drs. Anas Sudijono
NIP. 150 028 774

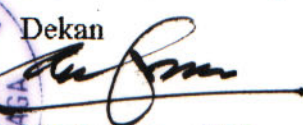
Penguji II

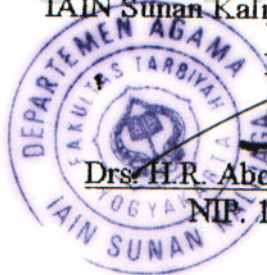

Dra. Wiji Hidayati
NIP. 150 246 924

Yogyakarta, 15 Mei 2000

IAIN Sunan Kalijaga Fakultas Tarbiyah

Dekan


Drs. H.R. Abdullah Fadjar, M.Sc.
NIP. 150 028 800



HALAMAN MOTTO

يَا أَيُّهَا الَّذِينَ ءَامَنُوا لِمَ تَقُولُونَ مَا لَا تَفْعَلُونَ
كَبُرَ مَقْتًا عِنْدَ اللَّهِ أَنْ تَقُولُوا مَا لَا تَفْعَلُونَ ^{الصف: ٢-٣}

*"Hai orang-orang yang beriman, mengapa kamu mengatakan apa yang tidak kamu perbuat. Sangat besar kebencian di sisi Allah bahwa kamu mengatakan apa-apa yang tidak kamu kerjakan." (QS. Ash Shaf ; 2-3)**

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

* Depag. RI, *Surat Ash Shaff*, Ayat 2 – 3 , Edisi Baru Revisi Terjemah 1989, Semarang, Toha Putra, 1989, hal. 928

PERSEMBAHAN



**Kupersembahkan skripsi ini kepada
Keluarga Besar Bapak H. Muhammad Idris,
Keluarga Besar Bapak Sugino, dan
Almamater Fakultas Tarbiyah IAIN Sunan Kalijaga tercinta.**



**STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA**

KATA PENGANTAR

الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ وَالصَّلَاةُ وَالسَّلَامُ عَلَى أَشْرَفِ
الْأَنْبِيَاءِ وَالْمُرْسَلِينَ وَعَلَى آلِهِ وَأَصْحَابِهِ وَمَنْ تَبِعَهُ
بِإِحْسَانٍ إِلَى يَوْمِ الدِّينِ

Segala puji syukur penulis panjatkan kepada Allah SWT yang telah memberi kekuatan dan ridloNya sehingga penulis sanggup menyelesaikan skripsi ini. Shalawat dan salam terlimpah kepada Rasulullah dan keluarganya serta seluruh ummatnya hingga akhir jaman.

Penulis sangat merasa bahagia karena telah berhasil menyelesaikan skripsi ini sebagai program akhir yang harus ditempuh guna menyelesaikan studi di fakultas Tarbiyah IAIN Sunan Kalijaga Yogyakarta. Semua ini tak mungkin tercapai jika tidak mendapatkan bantuan dari berbagai pihak, untuk itu sangat wajar kiranya penulis mengucapkan terima kasih kepada segenap pihak yang telah banyak memberikan bantuan.

Terutama ucapan terima kasih tersebut penulis haturkan kepada :

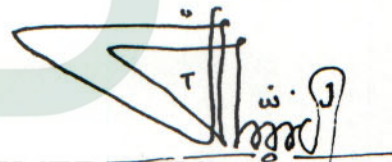
1. Bapak Dekan Fakultas Tarbiyah IAIN Sunan Kalijaga yang memiliki banyak kebijakan guna penyempurnaan fakultas tercinta.
2. Bapak Drs. Tasman Hamami, MA selaku Pembimbing skripsi yang telah mencurahkan banyak waktu dan tenaga untuk tidak bosan-bosannya membimbing penulis.
3. Bapak Drs. A. Miftah Baidlowi selaku Penasehat Akademik yang telah banyak mengarahkan penulis dalam menyelesaikan studi.
4. Bapak dan Ibu Dosen Fakultas Tarbiyah yang telah menanamkan begitu banyak ilmu pada penulis.

5. Bapak Yuwono selaku Kepala MAN Yogyakarta II yang memberikan ijin dan kesempatan pada penulis untuk melakukan reaseach di MAN Yogyakarta II.
6. Bapak Hilal selaku Guru Bidang Studi Bahasa arab kelas dua MAN Yogyakarta II yang dengan tulus ikhlas telah memberikan begitu banyak informasi dan data-data yang penulis perlukan.
7. Keluarga besar H. M. Idris terutama Ibunda Tercinta dan Almarhum Ayahanda yang senantiasa merestui dan mendoakan penulis.
8. Saudara-saudaraku mahasiswa PBA-2 angkatan 96 yang sangat banyak memberikan wawasan dan rasa persaudaraan yang begitu dalam.

Tentu masih banyak pihak yang banyak membantu penulis dalam menyelesaikan studi ini, untuk mereka semua penulis senantiasa berdoa semoga apa yang telah mereka berikan menjadi amal soleh yang berkembang biak laksana padi. Dan semoga Allah menjadikan mereka sebagai orang-orang yang senantiasa mendapatkan kesejahteraan hidup dunia dan akhirat.

Yogyakarta, 20 April 2000

Penulis,



Muhammad Tauwab
NIM. 9642 3394

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN NOTA DINAS	ii
HALAMAN PENGESAHAN.....	iv
HALAMAN MOTTO.....	v
HALAMAN PERSEMBAHAN.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI	ix
DAFTAR TABEL DAN GRAFIK.....	xi
 BAB I : PENDAHULUAN.....	 1
A. Pengertian Judul dan Penegasan Istilah.....	1
B. Latar Belakang Masalah	3
C. Rumusan Masalah.....	6
D. Hipotesa.....	6
E. Alasan Pemilihan Judul.....	7
F. Tujuan dan Kegunaan Penelitian.....	8
G. Metode Penelitian.....	9
H. Tinjauan Pustaka	16
I. Sistematika Pembahasan.....	37
 BAB II : PEMBELAJARAN BAHASA ARAB DI KELAS DUA	
MAN YOGYAKARTA II.....	38
A. Keadaan guru dan Siswa Kelas Dua.....	38
B. Proses Pembelajaran Bahasa Arab.....	40
C. Problematika Pembelajaran Bahasa arab.....	42

D. Usaha-Usaha Meningkatkan Mutu Pembelajaran	
Bahasa Arab.....	44
E. Pelaksanaan Tes Sumatif Bidang studi Bahasa Arab.....	45

BAB III	: ANALISIS KUALITAS TES SUMATIF CAWU I	
	BIDANGSTUDI BAHASA ARAB KELAS DUA	
	DI MAN YOGYAKARTA II.....	46
A.	Analisis Derajat Kesukaran Item.....	51
B.	Analisis Daya Pembeda Item.....	53
C.	Analisis Fungsi distraktor Item.....	55
D.	Analisis Validitas Item.....	57
E.	Analisis Reliabelitas Tes.....	60
BAB IV	: PENUTUP.....	67
A.	Kesimpulan.....	67
B.	Saran-Saran.....	68
C.	Kata Penutup.....	69

DAFTAR PUSTAKA

LAMPIRAN-LAMPIRAN

DAFTAR RALAT

CURRICULUM VITAE

ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR TABEL DAN GRAFIK

TABEL 1.	DISTRIBUSI FREKUENSI SKOR HASIL TES SUMATIF CAWU I BIDANG STUDI BAHASA ARAB KELAS DUA MAN YOGYAKARTA II.....	47
GRAFIK 1.	TABULASI SKOR TES SUMATIF BENTUK OBYEKTIF BIDANG STUDI BAHASA ARAB KELAS DUA.....	48
TABEL 2.	ANALISA SKOR TES SUMATIF CAWU I BIDANG STUDI BAHASA ARAB SISWA KELAS DUA MAN YOGYAKARTA II TAHUN PELAJARAN 1999/2000.....	50
TABEL 3.	PERHITUNGAN & INTERPRETASI UNTUK MENEN- TUKAN DERAJAT KESUKARAN ITEM TES SUMATIF CAWU I BIDANG STUDI BAHASA ARAB KELAS DUA MAN YOGYAKARTA II TAHUN PELAJARAN 1999/2000..	51
TABEL 4.	PERHITUNGAN-PERHITUNGAN UNTUK MENGETAHUI BESARNYA KOEFISIEN KORELASI PHI.....	54
TABEL 5.	ANALISIS FUNGSI DISTRAKTOR ITEM TES SUMATIF CAWU I BIDANG STUDI BAHASA ARAB KELAS DUA MAN YOGYAKARTA II TAHUN PELAJARAN 1999/2000.....	56

TABEL 6.	PERHITUNGAN-PERHITUNGAN UNTUK MENGETAHUI KOEFSIEN KORELASI POINT BISERIAL DALAM RANGKA UJI VALIDITAS ITEM.....	59
TABEL 7.	KESIMPULAN KUALITAS ITEM TES SUMATIF CAWU I BIDANG STUDI BAHASA ARAB.....	64



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
 YOGYAKARTA

BAB I

PENDAHULUAN

A Pengertian Judul Dan Penegasan Istilah

Skripsi ini berjudul “Analisis Kualitas Tes Sumatif Catur Wulan I Bidang Studi Bahasa Arab Kelas Dua Di MAN Yogyakarta II Tahun Pelajaran 1999/2000”. Untuk itu judul perlu dijelaskan agar terhindar dari penafsiran yang berbeda beda dari para pembaca.

Adapun mengenai istilah yang perlu mendapatkan penjelasan yaitu sebagai berikut:

1. *Analisis*, berarti penyelidikan terhadap suatu peristiwa (karangan, perbuatan dan sebagainya) untuk mengetahui keadaan yang sebenarnya. Dapat pula berarti penguraian suatu pokok atas berbagai bagiannya dan penelaahan bagian itu sendiri serta hubungan antar bagian untuk memperoleh pengertian yang tepat dan pemahaman arti keseluruhan.¹⁾ Yang dimaksud oleh penulis yaitu penyelidikan dan penelaahan terhadap tes sumatif Cawu I bidang studi bahasa arab kelas dua di MAN Yogyakarta II tahun pelajaran 1999/2000 untuk mengetahui keadaan yang sebenarnya.
2. *Kualitas*, berarti tingkat baik buruknya sesuatu; kadar; derajat atau taraf; mutu.²⁾

Yang dimaksud disini adalah kualitas atau mutu dari tes sumatif tersebut

1) *Kamus Besar Bahasa Indonesia*, Tim Penyusun Kamus Pusat Pembinaan dan Pengembangan Bahasa, cet. III, Jakarta, Balai Pustaka, 1990, hlm. 32.

2) *Ibid*, hlm. 467.

ditinjau dari aspek derajat kesukaran item, daya pembeda item, fungsi distraktor item, validitas item dan Reliabilitas tes.

3. *Tes Sumatif*, yaitu tes hasil belajar yang dilakukan setelah sekumpulan satuan program pelajaran selesai diberikan atau berakhir atau setelah seluruh unit pelajaran selesai diajarkan.³⁾ Adapun pengertian tes itu sendiri adalah cara (yang dapat digunakan) atau prosedur (yang perlu ditempuh) dalam rangka pengukuran dan penilaian di bidang pendidikan, yang berbentuk pemberian tugas atau serangkaian tugas baik berupa pertanyaan-pertanyaan yang harus dijawab atau perintah-perintah yang harus dikerjakan oleh testee, sehingga (atas dasar data yang diperoleh dari hasil pengukuran tersebut) dapat dihasilkan nilai yang melambangkan tingkah laku atau prestasi testee; nilai mana dapat dibandingkan dengan nilai-nilai yang dicapai oleh testee lainnya, atau dibandingkan dengan nilai standar tertentu.⁴⁾ Jadi tes sumatif yang dimaksud oleh penulis adalah tes sumatif cawu I bidang studi bahasa arab yang diujikan pada siswa-siswi kelas dua MAN Yogyakarta II tahun pelajaran 1999/2000 yang diadakan oleh Kanwil DEPAG DIY.
4. *Bidang Studi*, yaitu pengelompokan sejumlah mata pelajaran yang sejenis atau yang memiliki ciri-ciri yang sama; mata pelajaran yang telah berkorelasi satu

3) Anas Sudijono, *Teknik Evaluasi Pendidikan Suatu Pengantar*, Yogyakarta, U.D. Rama, 1986, hlm. 47

4) Anas Sudijono, *Pengantar Evaluasi Pendidikan*, cet. II, Jakarta, Rajawali Pers, 1998, hlm. 67.

dengan lainnya.⁵⁾ Yang dimaksud disini adalah bidang studi bahasa arab sebagai salah satu mata pelajaran wajib di MAN Yogyakarta II.

5. *Bahasa Arab*, menurut Musthofa Gholayaini adalah

اللُّغَةُ الْعَرَبِيَّةُ هِيَ الْكَلِمَاتُ الَّتِي يُجَبَّرُ بِهَا الْعَرَبُ عَنْ أَغْرَاضِهِمْ⁶⁾

yaitu kata kata yang dikemukakan oleh bangsa Arab atau penduduk Negeri Arab untuk menyatakan maksud dan tujuan mereka.

Jadi secara keseluruhan yang dimaksudkan oleh penulis mengenai judul di atas adalah penyelidikan, penguraian dan penelaahan tentang mutu tes sumatif cawu I bidang studi bahasa arab (dengan menelusuri butir-butir soalnya, apakah bermutu atau tidak) khusus pada kelas dua di MAN Yogyakarta II tahun pelajaran 1999/2000. Penelitian ini memfokuskan diri pada item-item tes bentuk obyektif dan kemampuan kerjanya sebagai alat evaluasi pendidikan.

B. LATAR BELAKANG MASALAH

Bila kita memperhatikan realita dalam dunia pendidikan, maka akan kita ketahui bahwa dalam setiap jenis pendidikan, pada waktu-waktu tertentu selama satu periode pendidikan orang selalu mengadakan evaluasi; artinya pada waktu-waktu tertentu selama satu periode pendidikan tadi orang selalu mengadakan penelitian terhadap hasil yang telah dicapai baik oleh pihak pendidik maupun pihak terdidik hingga waktu tertentu. Baik pendidikan itu terjadi di rumah tangga

5) Kamus Besar Bahasa Indonesia, *Op. cit.* hlm. 114.

6) الشيخ مصطفى غلاييني، جامع الدروس العربية، بيروت، المكتبة العصرية، ١٩٩٤، ص ٧

(pendidikan informal), di masyarakat (pendidikan nonformal) terlebih lagi di sekolah-sekolah (pendidikan formal), selalu akan kita dapati fenomena bahwa orang mengadakan evaluasi terhadap hasil usaha yang telah dilakukannya dalam jangka waktu tertentu.

Evaluasi merupakan hal yang penting dalam dunia pendidikan, sebab dengan evaluasi dapat diketahui kemajuan anak didik atau orang yang terdidik dalam jangka waktu tertentu, disamping itu diketahui pula tingkat efisiensi metode yang digunakan dalam proses belajar mengajar.⁸⁾

Proses evaluasi dalam pendidikan khususnya evaluasi terhadap prestasi belajar siswa sebagian besar bersumber dari hasil pengukuran yang berupa data yang dilambangkan dengan angka-angka. Dan atas dasar angka-angka inilah dilakukan penginterpretasian dan penafsiran. Evaluasi hasil belajar siswa tidak dapat dilakukan dengan baik manakala tidak didasarkan atas data yang bersifat kuantitatif.

Menurut Masroen, pada umumnya para pakar di bidang pendidikan sependapat, bahwa evaluasi mengenai proses pembelajaran di sekolah tidak mungkin dapat dilaksanakan secara baik apabila evaluasi itu tidak didasarkan atas data yang bersifat kuantitatif.⁹⁾ Inilah sebabnya mengapa dalam praktek, masalah pengukuran mempunyai kedudukan yang sangat penting dalam proses evaluasi. Baik buruknya evaluasi akan sangat tergantung pada hasil-hasil pengukuran yang

8) M. Buchori, *Teknik-Teknik Evaluasi dalam Pendidikan*, Bandung, Jemmars, 1980, hlm. 6-7.

9) Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta, Rajawali Pers, 1998, hlm. 6.

mendahuluinya. Hasil pengukuran yang kurang cermat akan memberikan hasil evaluasi yang kurang cermat pula; sebaliknya teknik-teknik pengukuran yang tepat dapat diharapkan memberikan landasan yang kokoh untuk mengadakan evaluasi yang tepat.¹⁰

Dalam dunia pendidikan dikenal dua jenis alat evaluasi yaitu alat yang menggunakan teknik non tes dan teknik tes. Khusus dalam evaluasi hasil belajar siswa lazim menggunakan teknik tes. Tes yang lazim digunakan dalam hal ini terbagi menjadi tiga yaitu tes diagnostik, tes formatif dan tes sumatif.¹¹⁾ Dalam pengalaman di sekolah tes formatif dapat disamakan dengan ulangan harian, sedangkan tes sumatif dapat disamakan dengan ulangan umum yang biasanya dilaksanakan pada tiap akhir catur wulan atau akhir semester.¹²⁾

Tes sumatif sebagai salah satu alat evaluasi memiliki beberapa manfaat, terutama untuk menentukan nilai guna menetapkan posisi siswa di dalam kelompok kelas. Juga untuk menentukan seseorang siswa dapat atau tidaknya mengikuti kelompok dalam menerima program berikutnya, pula untuk mengisi buku rapor.¹³⁾

Melihat urgensi dari tes sumatif tersebut maka penulis memiliki minat yang besar untuk mengetahui sejauh mana kualitas daripada tes sumatif cawu I bidang studi bahasa arab kelas dua di MAN Yogyakarta II, khususnya kemampuannya

10) *Ibid*

11) Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, cet. XI, Jakarta, Bumi Aksara, 1995, hlm. 30.

12) Anas Sudijono, *Op. cit.* hlm. 36.

13) *Ibid*, hlm. 37-39.

sebagai alat evaluasi ; apakah tes tersebut telah memberikan informasi tentang kemajuan siswa, sudah mampukah membedakan kelompok siswa yang pandai dan yang kurang pandai, apakah terdapat kelemahan-kelemahan dalam butir-bitur soalnya, apakah tes tersebut memiliki validitas dan Reliabilitas yang memadai. Untuk itu perlu dilakukan penyelidikan dan penelaahan lebih lanjut, mengingat begitu pentingnya bahasa arab dalam kancah globalisasi dunia dewasa ini dan lebih utama lagi bagi umat islam untuk memperdalam wawasan keislamannya.

C. RUMUSAN MASALAH

Beranjak dari latar belakang masalah diatas, maka penelitian tentang kualitas tes sumatif cawu I bidang studi bahasa arab kelas dua di MAN Yogyakarta II tahun pelajaran 1999/2000 penulis fokuskan pada masalah sebagai berikut:

“Apakah tes sumatif bidang studi bahasa arab (khusus pada tes bentuk obyektif) cawu I kelas dua di MAN Yogyakarta II tahun pelajaran 1999/2000 telah memiliki kualitas yang baik sebagai alat evaluasi, ditinjau dari segi derajat kesukaran item, daya pembeda item, fungsi distarktor item, validitas item dan reliabilitas tes tersebut.”

D. HIPOTESA

Dalam sebuah tes, pada umumnya suatu kelompok testee merupakan populasi yang bersifat heterogen, dimana ada sebagian kecil testee yang tergolong bodoh,

sebagian kecil pandai dan kebanyakan mereka tergolong testee yang berkemampuan sedang.

Bila hal itu terjadi maka tes yang telah dilaksanakan tersebut tidak memiliki kelainan atau wajar. Namun jika hasilnya menunjukkan bahwa sebagian besar testee mendapatkan skor rendah atau tinggi dan hanya sebagian kecil testee yang memperoleh skor sedang, berarti ada sesuatu yang tidak normal dari tes tersebut.

Dari hasil peninjauan awal terhadap skor-skor siswa kelas dua di MAN Yogyakarta II pada bidang studi bahasa arab tes sumatif cawu I tahun pelajaran 1999/2000 diperoleh informasi bahwa sebagian besar siswa memperoleh skor yang rendah.

Atas dasar itulah maka penulis memiliki hipotesa bahwa tes sumatif cawu I bidang studi bahasa arab kelas dua di MAN Yogyakarta II tahun pelajaran 1999/2000 –khusus pada tes bentuk obyektif– memiliki kualitas yang kurang baik sebagai alat evaluasi.

E. ALASAN PEMILIHAN JUDUL

Penulis tertarik pada judul tersebut dengan alasan sebagai berikut:

1. Setelah mengadakan wawancara peninjauan kepada guru bidang studi bahasa arab kelas dua di MAN Yogyakarta II berhasil didapatkan informasi bahwa para siswa kelas dua di sekolah tersebut memiliki skor yang rendah, yaitu berkisar antara 30 s.d. 60 untuk bidang studi bahasa arab, yang berarti dibawah standar yang ditetapkan oleh Kanwil DEPAG DIY yaitu 70.



2. Dalam pelaksanaan evaluasi, sebagian besar guru hanya sampai pada memberi nilai bagi siswa. Hanya sebagian kecil para guru yang mengadakan analisis terhadap alat ukur berupa tes tersebut, apakah berkualitas atau tidak. Padahal hal tersebut sangat penting. Hal ini mungkin disebabkan terbatasnya waktu dan kemampuan yang dimiliki.
3. Tes sumatif merupakan pengukuran hasil belajar siswa yang dilaksanakan sebagai puncak dari proses belajar mengajar yang telah dilalui oleh siswa setelah menyelesaikan seluruh unit pelajaran selama satu cawu.
4. Bidang studi bahasa arab perlu terus menerus mendapatkan perhatian, sebab kini bahasa arab telah menjadi salah satu bahasa resmi di PBB dan merupakan alat bantu dalam memahami kajian-kajian keislaman yang terutama bersumber dari al-quran dan al-hadits.

F. TUJUAN DAN KEGUNAAN PENELITIAN

1. Tujuan.

Mengetahui kualitas tes sumatif bidang studi bahasa arab cawu I kelas dua di MAN Yogyakarta II tahun pelajaran 1999/2000. Sehingga dapat ditentukan butir-butir soal yang telah menjalankan fungsi sebagai alat tes dengan baik, dan juga butir-butir soal yang tidak baik, guna perbaikan lebih lanjut.

2. Kegunaan.

Manfaat yang dapat diperoleh dari penelitian ini antara lain adalah dapat diperoleh informasi yang lengkap mengenai kualitas tes tersebut, baik ditinjau dari segi validitas item tes, Reliabilitas tes, derajat kesukaran item, daya pembeda item maupun fungsi distraktornya. Dimana bagi penyusun tes, hal tersebut akan menjadi informasi yang sangat berharga guna penyempurnaan dan perbaikan tes selanjutnya. Bagi guru bidang studi bahasa arab, hal itu memberikan informasi sejauh mana efektifitas dan efesinsi metode yang digunakan selama ini, sehingga untuk waktu-waktu selanjutnya diadakan penyempurnaan. Selain itu juga diketahui sejauh mana tujuan pembelajaran telah tercapai. Sedangkan bagi penulis sendiri hal ini merupakan pengalaman yang sangat berharga guna masa depan sebagai calon pendidik.

G. METODE PENELITIAN

1. Metode Penentuan Obyek Penelitian

Dalam penelitian ini yang menjadi obyek utama penelitian adalah tes sumatif cawu I bidang studi bahasa arab kelas dua di MAN Yogyakarta II tahun pelajaran 1999/2000, khususnya tes bentuk obyektif. Yang berupa alat tes itu sendiri dan lembar jawaban dari siswa . Obyek lainnya adalah kegiatan pembelajaran bahasa arab di kelas dua.

Sedangkan data penelitian diperoleh dari siswa kelas dua, guru bidang studi bahasa arab kelas dua, kepala sekolah dan dokumen hasil tes sumatif cawu I bidang studi bahasa arab kelas dua.

Berkaitan dengan besarnya populasi yang diteliti, maka penelitian ini termasuk penelitian sampel, dimana penulis mengambil sebagian saja dari seluruh populasi untuk diteliti. Besarnya sampel yang diteliti adalah 42,5% atau sebanyak 100 buah lembar jawaban testee dari sejumlah 235 buah lembar jawaban. Penetapan tersebut dilandasi oleh pendapat dari Dr. Suharsimi Arikunto, apabila populasi kurang dari 100, lebih baik diambil semua sehingga penelitiannya merupakan penelitian populasi. Selanjutnya jika berjumlah besar maka dapat diambil sampel sebesar 10-15% atau 20-25% atau lebih.¹⁴⁾

Metode yang digunakan dalam penetapan sampel ini adalah metode stratified random sampling. Dan karena seluruh testee terbagi kedalam enam kelas, dimana terdapat dua kelas yang terdiri dari 38-39 testee, maka masing-masing kelompok diambil $42.5\% \times 38 = 16$ lembar jawaban testee.

Sedangkan untuk empat kelas lainnya yang masing-masing terdiri dari 40 siswa, maka diambil sampel sebesar $42.5\% \times 40 = 17$ lembar jawaban. Sehingga seluruh lembar jawaban tersebut berjumlah $(16 \times 2) + (17 \times 4) = 100$ lembar jawaban. Seluruhnya akan diambil secara random dari masing-masing kelas/kelompok dengan menggunakan metode ordinal – kelipatan tiga -.

2. Metode Pengumpulan Data

Untuk mengumpulkan data-data, maka digunakan metode-metode sebagai berikut:

1. Dokumentasi.

Penelitian dengan menggunakan metode dokumentasi adalah penelitian yang dalam pengumpulan datanya berdasarkan pada dokumen-dokumen yang telah tersedia. Menurut Dr. Suharsimi Arikunto, metode dokumentasi adalah mencari data mengenai hal-hal atau variabel yang berupa catatan, transkrip, buku, surat kabar, majalah, notulen rapat, legger, agenda dan sebagainya.¹⁵⁾

Metode ini digunakan untuk memperoleh data tentang skor siswa dalam bidang studi bahasa arab kelas dua yang berupa lembar jawaban, alat tes sumatif berupa lembar soal bahasa arab dan kunci jawabannya, kisi-kisi soal dan keadaan guru dan siswa kelas dua.

2. Wawancara.

Interviu atau wawancara adalah metode pengumpulan dan pencatatan data, informasi dan atau pendapat yang dilakukan melalui percakapan dan tanya jawab, baik langsung maupun tidak langsung dengan sumber data. Yang dimaksud secara langsung adalah wawancara yang dilakukan secara langsung dengan sumber data. Sedangkan wawancara tidak langsung

14) Suharsimi Arikunto, *Prosedur Penelitian Suatu Pendekatan Proses*, Jakarta, Rineka Cipta, 1996, hal.120

15) *Ibid*, hal.234

artinya bahwa pewawancara menanyakan suatu data melalui perantara orang lain, tidak langsung pada sumber data.¹⁶⁾

Sehubungan dengan itu maka penulis menggunakan jenis interviu bebas terpimpin, dimana pewawancara membawa kerangka pertanyaan untuk diajukan, tetapi cara bagaimana pertanyaan tersebut akan diajukan serta irama wawancaranya sepenuhnya diserahkan pada pewawancara.¹⁷⁾ Selanjutnya metode ini dimanfaatkan untuk memperoleh data tentang keadaan guru dan siswa, gambaran umum tes, proses pembelajaran bahasa arab di kelas dua, problematika pembelajaran bahasa arab serta usaha-usaha peningkatan kualitas bahasa arab di sekolah.

3. *Observasi.*

Sebagai metode ilmiah observasi biasanya diartikan sebagai pengamatan dan pencatatan dengan sistematis fenomena-fenomena yang diselidiki, dalam arti luas sebenarnya observasi tidak terbatas hanya pada pengamatan yang dilakukan secara langsung, tapi juga tidak langsung.¹⁸⁾

Observasi ini digunakan untuk mengetahui proses pembelajaran bahasa arab di kelas dua dan problematika pembelajarannya.

16) Zainal Arifin, *Evaluasi Instruksional Prinsip-Teknik-Prosedur*, Bandung, Remaja Rosda Karya, 1991, hal.49

17) Masri Singarimbun dan Sofian Effendi, *Metode Penelitian Survei*, Jakarta, LP3ES, 1989, hal.

3. Metode Analisa Data

Dalam rangka menganalisa data yang diperoleh, maka penulis menggunakan beberapa metode, baik berupa metode non statistik maupun statistik. Metode non statistik digunakan untuk mengolah data yang bersifat kualitatif seperti proses pembelajaran bahasa arab dan problematikanya, keadaan guru dan siswa dan sebagainya. Sedangkan metode statistik digunakan untuk mengolah data yang bersifat kuantitatif, yaitu data skor-skor jawaban siswa pada tes sumatif bahasa arab.

Adapun untuk analisa statistik digunakan rumus-rumus sebagai berikut:

1. Untuk mencari tingkat indek kesukaran soal digunakan rumus.¹⁹

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

Dimana ;

P = Proporsa = angka indek kesukaran item

Np = banyaknya testee yang dapat menjawab dengan betul terhadap butir item yang bersangkutan.

N = jumlah testee yang mengikuti tes

18) Sutrisno Hadi, *Metodologi Riserch*, Yogyakarta, Yayasan Penerbit Fak. Psikologi UGM, 1986, hal. 136

19 Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta, Rajawali Pers, 1998, hal. 372

2. Untuk mencari kemampuan daya pembeda soal digunakan rumus:²⁰⁾

$$\phi = \frac{P_H - P_L}{2\sqrt{p \cdot q}}$$

Dimana ;

- ϕ = angka indek korelasi Phi, yang dalam hal ini dianggap sebagai angka diskriminasi item
- P_H = proportion of the higher group
- P_L = proportion of the lower group
- 2 = bilangan konstan
- p = proporsi seluruh testee yang jawabannya betul
- q = propori seluruh testee yang jawabannya salah

3. Untuk mengetahui berfungsi tidaknya distraktor soal digunakan prosentase dipilihnya distraktor oleh testee.
4. Untuk pengujian validitas item tes digunakan rumus teknik korelasi point biserial. Dan dalam interpretasi terhadap hasilnya digunakan db sebesar (N-nr), yaitu $100 - 2 = 98$. Derajat kebebasan sebesar 98 selanjutnya dikonsultasikan pada tabel nilai "r" product moment pada taraf signifikan 1% yakni 0,254. Adapun rumus korelasi point biserial adalah sebagai berikut;²¹⁾

20 Ibid, hal.391

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

Dimana ;

R_{pbi} = koefisien korelasi point biserial yang dipandang sebagai koefisien validitas

M_p = skor rata-rata hitung yang dimiliki oleh testee, yang untuk item yang bersangkutan telah dijawab benar

M_t = skor rata-rata dari skor total

SD_t = deviasi standar dari skor total

p = proporsi testee yang menjawab benar butir soal yang bersangkutan

q = proporsi testee yang menjawab salah butir soal yang bersangkutan

5. Untuk pengujian Reliabilitas tes bentuk obyektif digunakan formula flanagan model belahan kiri dan kanan, dengan rumus sebagai berikut,²²

$$r_{11} = 2 \left(1 - \frac{S_1^2 + S_2^2}{S_t^2} \right)$$

Dimana ;

r_{11} = koefisien reliabilitas tes secara keseluruhan

2 & 1 = bilangan konstan

21 Ibid, hal.185

22 Anas Sudijono, Ibid, Hal.231

S_1^2 = jumlah kuadrat deviasi dari skor-skor hasil tes yang termasuk pada belahan I

S_2^2 = jumlah kuadrat deviasi dari skor-skor hasil tes yang termasuk pada belahan II

S_t^2 = jumlah kuadrat total deviasi dari skor-skor belahan I dan II

H. Tinjauan Pustaka

1. Tes Sebagai Alat Evaluasi Pendidikan

Dalam dunia pendidikan evaluasi merupakan hal yang fundamental, sebab dengan evaluasi dapat diketahui kemajuan siswa dalam jangka waktu tertentu. Selain itu bagi guru dapat pula mengetahui efektifitas dan efisiensi metode yang ia gunakan selama proses pembelajaran dalam jangka waktu tersebut.

Dalam keseluruhan proses pendidikan, secara garis besar evaluasi berfungsi untuk:

1. mengetahui kemajuan kemampuan belajar murid,
2. mengetahui status akademis seorang siswa dalam kelompok/kelas,
3. mengetahui penguasaan, kekuatan dan kelemahan seorang siswa atas suatu unit pelajaran,
4. mengetahui efisiensi metode mengajar yang digunakan guru,
5. menunjang pelaksanaan BP,
6. memberikan laporan pada siswa dan orang tuanya,
7. keperluan promosi siswa,

8. keperluan pengurusan (streaming)
9. perencanaan pendidikan,
10. memberikan informasi kepada masyarakat yang memerlukan,
11. bahan feed back bagi siswa, guru dan program pengajaran'
12. alat motivasi belajar mengajar.²³

Proses evaluasi dalam pendidikan, khususnya evaluasi terhadap hasil belajar siswa, sebagian besar bersumber pada hasil pengukuran. Ini berarti bahwa kegiatan pengukuran menghasilkan data-data yang bersifat kuantitatif yang dilambangkan dengan angka-angka.

Menurut Masroen, pada umumnya para pakar di bidang pendidikan sependapat bahwa evaluasi mengenai proses pembelajaran di sekolah tidak mungkin dapat dilaksanakan secara baik apabila evaluasi itu tidak didasarkan atas data yang bersifat kuantitatif.²⁴ Oleh sebab itu diperlukan alat yang dapat digunakan untuk melakukan evaluasi.

Secara garis besar alat evaluasi pendidikan dibagi menjadi dua, yaitu yang menggunakan teknik tes dan non tes.

a. *Teknik non tes.*

Evaluasi hasil belajar siswa dengan menggunakan teknik non tes dilakukan dengan tanpa menguji siswa, melainkan dengan melakukan observasi, interviu, angket, chek list, rating scale dan biography.²⁵

23) Drs.Slameto, *Evaluasi Pendidikan*, Jakarta, Bina Aksara, 1988, hal.15-16

24) Anas Sudijono, *Pengantar Evaulasi Pendidikan*, Jakarta, Rajawali Pers, 1995, hal.6

25) *Ibid*, hal.76

b. *Teknik tes.*

Tes adalah penilaian yang komprehensif terhadap individu atau keseluruhan usaha evaluasi program.²⁶ Ada beberapa istilah yang perlu diketahui berkaitan dengan tes tersebut. Istilah-istilah tersebut adalah tes, testee, tester dan testing.

Tes adalah alat atau prosedur yang digunakan dalam rangka pengukuran dan penilaian. Testee adalah pihak yang dikenai tes atau peserta tes. Testing adalah saat dilaksanakannya tes. Tester adalah orang yang menyelenggarakan tes atau evaluator (orang yang sedang melakukan evaluasi dengan menggunakan teknik tes).²⁷

Berkaitan dengan kegiatan pembelajaran maka yang dimaksud tes hasil belajar adalah cara (yang dapat dipergunakan) atau prosedur (yang dapat ditempuh) dalam rangka pengukuran dan penilaian di bidang pendidikan, yang berbentuk pemberian tugas atau serangkaian tugas, baik berupa pertanyaan-pertanyaan (yang harus dijawab), atau perintah-perintah (yang harus dikerjakan) oleh testee, sehingga (atas dasar data yang diperoleh dari hasil pengukuran tersebut) dapat dihasilkan nilai yang melambangkan tingkah laku atau prestasi testee; nilai mana dapat dibandingkan dengan nilai-nilai yang dicapai oleh testee lainnya, atau dibandingkan dengan nilai standar tertentu.²⁸

26) Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, Jakarta, Bumi Aksara, 1995, hal.30,

27) Anas Sudijono, *Op.cit*, hal.66

28) *Ibid*, hal 67

Dalam tes prestasi hasil belajar yang tujuannya untuk mengetahui seberapa jauh penguasaan siswa terhadap bahan pelajaran dapat dikategorikan menjadi tiga macam tes yaitu tes diagnostik, tes formatif dan tes sumatif.²⁹ Tes diagnostik bertujuan untuk mengetahui kekurangan atau kelemahan siswa dalam penguasaan materi pelajaran. Tes formatif bertujuan untuk memperbaiki program pembelajaran, dan tes sumatif bertujuan untuk memberikan status kemajuan belajar siswa secara keseluruhan tentang bahan pelajaran yang telah diberikan.

2. Fungsi Tes

Secara umum ada dua macam fungsi tes, yaitu:

- a. sebagai alat ukur bagi anak didik. Dalam hal ini tes sebagai alat ukur untuk mengetahui tingkat perkembangan atau kemajuan yang telah dicapai oleh anak didik setelah menempuh proses pendidikan dalam jangka waktu tertentu.
- b. Sebagai alat ukur keberhasilan program pengajaran. Tes dikatakan berfungsi sebagai alat ukur program pengajaran sebab melalui tes tersebut akan dapat diketahui sejauh mana program pengajaran yang telah ditentukan telah tercapai.³⁰

29) Suharsimi Arikunto, *Op.cit*, hal.30

30) Anas Sudijono, *Op.cit*, hal.67

3. Analisis Item Sebagai Salah Satu Cara Untuk Mengetahui Kualitas Tes

a. Arti penting penelaahan kembali item tes.

Dalam kehidupan sehari-hari kita dapat bahwa dalam kegiatan pembelajaran guru menyelenggarakan tes, baik tes formatif (ulangan harian) maupun tes sumatif (ulangan umum). Setelah selesai memeriksa dan memberi nilai maka sebagian mereka menganggap bahwa tugas mereka telah selesai. Anggapan ini kurang tepat, sebab bagi guru yang profesional tugasnya bukan hanya membuat dan menyelenggarakan tes dan kemudian memberi nilai, melainkan pula perlu untuk mengadakan penelaahan terhadap kualitas dari tes tersebut. Sehingga semata-mata guru tidak menyalahkan siswa apabila nilai mereka rendah itu disebabkan oleh kebodohan mereka. Bisa jadi sebenarnya tes tersebut memang terlalu sulit bagi mereka.

Suatu tes yang dikatakan memiliki kualitas yang baik memiliki ciri-ciri sebagai berikut;

a. *valid (memiliki validitas)*

Tes dapat dikatakan valid apabila dengan secara tepat mengukur apa yang seharusnya diukur, atau tes yang secara benar (shahih) dapat mengungkapkan apa yang seharusnya diungkap dari anak didik.³¹

31) *Ibid*, hal.94

Oleh sebab itu dalam proses pembelajaran, tes hasil belajar dikatakan memiliki validitas apabila tes tersebut dapat mengukur secara tepat terhadap perkembangan belajar anak didik.

b. *reliabel (memiliki reliabilitas)*

Suatu tes dikatakan memiliki reliabilitas apabila tes tersebut dapat dipercaya, konsisten atau stabil dan produktif. Jadi yang dipentingkan ialah ketelitiannya; sejauh mana tes tersebut dapat dipercaya kebenarannya.³²

Dalam kaitannya dengan tes prestasi belajar, tes dipandang reliabel apabila dapat mengukur; dalam keadaan yang bagaimanapun, kapanpun dan oleh siapapun tes itu dilaksanakan akan selalu menunjukkan hasil yang sama atau relatif sama.

c. *obyektif (memiliki obyektifitas)*

Yang dimaksud disini adalah bahwa tes tersebut dapat mengukur dengan apa adanya³³ (berjalan apa adanya) yang berarti tidak ada unsur rekayasa yang dapat mempengaruhinya, terutama dalam hal pemberian skor.

d. *praktis dan ekonomis*

Tes dikatakan praktis mengandung arti bahwa tes tersebut sederhana (tidak memerlukan alat banyak dan sulit pengadaaannya)

32) Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pendidikan*, hal. 179

33) Anas Sudijono, *Op.cit*, hal.96

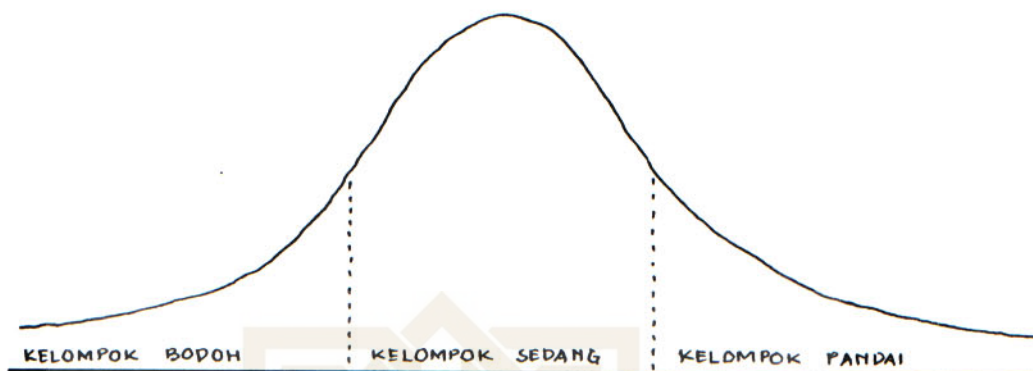
dan lengkap petunjuk pengerjaannya, kunci jawabannya dan pedoman skoringnya. Dan dikatakan ekonomis dalam arti bahwa tes tersebut tidak memerlukan biaya yang mahal, waktu yang lama dan tenaga yang besar.

Dari uraian diatas dapat difahami bahwa seorang tester perlu mempertimbangkan hal-hal tersebut guna penyusunan dan penyelenggaraan tes berikutnya. Dan untuk mengetahui ciri-ciri itu diperlukan langkah penganalisaan tes lebih lanjut.

Analisis item bukanlah ciri-ciri tes yang baik melainkan suatu kegiatan yang dapat membantu meningkatkan kebaikan tes, dengan cara melihat tingkat kebaikan setiap item tes. Dengan analisis tersebut dapat diketahui item mana yang perlu diperbaiki, dibuang atau diganti, serta item mana yang terus dapat digunakan. Dengan analisis item akan membantu meningkatkan mutu analisis validatas dan Reliabilitas tes.³⁴

Dalam suatu tes yang mengacu pada norma kelompok (norm referenced evaluation), didasarkan pada suatu asumsi bahwa pada setiap populasi yang bersifat heterogen, akan selalu dijumpai adanya kelompok baik/pandai, kelompok sedang dan kelompok rendah/bodoh. Yang distribusinya membentuk kurva normal seperti berikut :

34) Mudjijo, *Tes Hasil Belajar*, Jakarta, Bumi Aksara, 1995, hal.61-62



Asumsi ini bermakna bahwa pada setiap kegiatan pengukuran dan penilaian hasil belajar, sebagian besar testee termasuk kelompok siswa yang memiliki kemampuan sedang, dan hanya sebagian kecil yang termasuk testee dengan kemampuan rendah dan atau tinggi.

Apabila setelah dilaksanakan tes dan skor-skoranya ditabulasikan dalam sebuah kurva, ternyata terbentuk kurva sebagaimana diatas, maka itu merupakan informasi bahwa tes tersebut tidak memiliki kelainan, atau wajar. Akan tetapi apabila ditemukan bahwa kurva tersebut berbentuk miring atau juling, seperti gambar berikut;



maka berarti ada sesuatu yang tidak wajar atau tidak normal dari tes tersebut. Untuk itulah diperlukan penganalisaan terhadap item tes itu.

b. Manfaat analisa item

Dalam kegiatan evaluasi pendidikan kegiatan analisa item memberi manfaat antara lain;

1. membantu memperoleh informasi yang lengkap mengenai butir-butir soal yang telah dilaksanakan dalam tes, baik dari segi kelemahan maupun kebaikannya. Kemudian dapat dilakukan langkah lebih lanjut sehingga tes tersebut berfungsi dengan baik sebagai alat evaluasi.
2. Untuk memperoleh gambaran yang jelas tentang keadaan yang kita susun.

c. Cara menelaah tes.

Cara yang dapat ditempuh untuk menelaah kembali tes yang telah disusun dan dilaksanakan ada 4 macam cara:³⁵

1. Meneliti secara jujur soal-soal tersebut. Terkadang diperoleh ketidakjelasan perintah atau bahasa.
2. Mengadakan analisis item, yaitu suatu prosedur yang sistematis yang dapat memberikan informasi-informasi yang sangat khusus tentang item tes tersebut.
3. Mengadakan Analisis validitas.
4. Mengadakan Analisis Reliabilitas.

4. Teknik Analisis Item Tes Bentuk Obyektif

Sebagaimana telah ditegaskan pada bagian awal skripsi ini bahwa analisis kualitas tes yang dimaksud hanya dibatasi pada analisis terhadap tes

35) Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, hal.309

sumatif bentuk obyektif, maka berikut ini dipaparkan tentang teknik-teknik analisis item bentuk obyektif.

a. Analisis Derajat Kesukaran Soal.

Tinggi rendahnya atau baik buruknya suatu tes hasil belajar pertama-tama dapat ditilik dari derajat kesukaran itemnya.

Tes hasil belajar yang baik adalah yang tidak terlalu mudah dan tidak terlalu sukar. Item tes yang terlalu mudah sehingga semua testee dapat menjawab dengan benar item tersebut, bukanlah item tes yang baik. Demikian pula item tes yang terlalu sukar sehingga semua testee menjawab dengan salah, tidak dapat dikatakan sebagai item yang baik³⁶

Untuk mengetahui derajat kesukaran item itu dapat dilihat dari besar kecilnya angka indek yang melambangkannya, yaitu yang dilambangkan dengan huruf P (proporsa). Besarnya angka indek tersebut berkisar antara 0,0 sampai dengan 1,0. Semakin kecil angka indek menunjukkan semakin sukarnya item tes, dan sebaliknya semakin besar angka indek menunjukkan semakin mudahnya item tes tersebut.

Salah satu teknik yang digunakan untuk menghitung angka indek tersebut adalah dengan menggunakan rumus:³⁷

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

36 Wayan Nurkencana, *Evaluasi Pendidikan*, Surabaya, Usaha Nasional, 1986, hal.134

37. Anas Sudijono, *Pengantar Evaluasi Pendidikan*, Jakarta, Rajawali Pers, 1998, hal.372

Dimana ;

P = proportion (angka indek kesukaran item)

N_p = banyaknya testee yang dapat menjawab benar item yang bersangkutan

N = jumlah seluruh testee.

Mengenai interpretasi yang diberikan untuk indek kesukaran itu oleh Thorndike dan Elizabeth Hagen dalam buku *Measurement and Evaluation in Psychology and Education* adalah sebagai berikut;

1. apabila P kurang dari 0,3 berarti soal terlalu sukar,
2. apabila P antara 0,3 – 0,7 berarti kesukarannya sedang, dan
3. apabila besarnya P lebih dari 0,7 berarti soal terlalu mudah.³⁸

b. Analisis Daya Pembeda Soal

Daya pembeda soal yaitu kemampuan soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang bodoh (berkemampuan rendah).³⁹ Angka yang menunjukkan daya pembeda disebut *diskriminating power (D)*. Besarnya angka diskriminasi tersebut berkisar antara 0,0 hingga 1,0. Tetapi terkadang dijumpai juga angka tersebut bertanda minus. Tanda minus tersebut menunjukkan bahwa butir soal tersebut lebih banyak dijawab benar oleh testee kelompok bawah dan sedikit dijawab benar oleh testee kelompok atas.

38 .Ibid, hal.372

39) Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, hal.215

Apabila angka indeks sebesar 0,0 berarti soal tersebut tidak memiliki daya pembeda, jadi testee pandai maupun bodoh sama-sama menjawab salah soal itu. Dan bila angka indeks bertanda positif berarti telah memiliki daya pembeda.

Interpretasi yang biasa diberikan untuk indeks D adalah sebagai berikut:

1. jika besarnya angka indeks kurang dari 0,2 maka daya pembeda soal itu jelek,
2. jika besarnya angka indeks antara 0,2 s.d. 0,4 maka daya pembedanya cukup,
3. jika besarnya angka indeks antara 0,4 s.d. 0,7 maka daya pembedanya baik,
4. jika besarnya angka indeks antara 0,7 s.d. 1.0 maka daya pembedanya sangat baik.
5. Jika angka indeks bertanda minus berarti daya pembedanya jelek sekali.⁴⁰

Untuk mencari besarnya daya pembeda soal dapat dicari dengan menggunakan teknik korelasi Phi dengan rumus sebagai berikut:⁴¹

$$\phi = \frac{P_H - P_L}{2\sqrt{p.q}}$$

40) Anas Sudijono, *Op.cit*, hal.391

41) Suharsimi Arikunto, *Op.cit*, hal.223

Keterangan:

ϕ = angka indeks korelasi Phi, yang dalam hal ini dianggap sebagai angka indeks diskriminasi item.

P_H = proportion of the higher group

P_L = proportion of the lower group

2 = bilangan konstan

p = proporsi seluruh testee yang jawabannya benar

q = proporsi seluruh testee yang jawabannya salah, dimana $q = (1-p)$

P_H dapat dicari dengan rumus:

$$P_H = \frac{B_A}{J_A}$$

Dan P_L dapat dicari dengan rumus:

$$P_L = \frac{B_B}{J_B}$$

Keterangan;

P_H = proporsi testee kelompok atas yang menjawab betul item yang bersangkutan.

P_L = proporsi testee kelompok bawah yang menjawab betul item yang bersangkutan

B_A = banyaknya testee kelompok atas yang menjawab betul item yang bersangkutan

B_B = banyaknya testee kelompok bawah yang menjawab benar soal yang bersangkutan

J_A = jumlah testee kelompok atas

J_B = jumlah testee kelompok bawah

c. Analisis Fungsi Distraktor Soal

Pengujian ini dikenal dengan penyebaran jawaban soal. Yang dimaksud adalah suatu pola yang dapat menggambarkan bagaimana testee dapat menentukan pilihan jawabannya, terhadap option yang tersedia.

Pola penyebaran jawaban itu dapat diketahui dengan cara menghitung banyaknya testee yang memilih option a, b, c, d, e atau yang tidak memilih sama sekali.

Dengan itu akan diketahui apakah distraktor soal berfungsi atau tidak. Distraktor atau pengecoh yang baik adalah yang memiliki daya rangsang atau daya tarik yang besar bagi testee yang kurang memahami materi tes untuk memilihnya.

Dalam evaluasi pendidikan, suatu distraktor dikatakan berfungsi dengan baik apabila distraktor tersebut dipilih oleh minimal 5% dari seluruh testee.⁴²

5. Analisis Validitas Soal.

Suatu alat ukur evaluasi dapat dikatakan valid apabila dapat mengukur apa yang seharusnya diukur secara tepat.⁴³ Secara umum validitas tes dapat dikategorikan menjadi dua yaitu validitas rasional dan validitas empirik. Validitas rasional adalah validitas yang dapat diperoleh melalui analisa rasional. Dan validitas ini terbagi menjadi dua yaitu validitas isi dan validitas susunan. Suatu tes memiliki validitas isi jika materi tes itu benar-benar mewakili materi yang telah diberikan pada siswa. Dan memiliki validitas susunan jika tes tersebut benar-benar dapat mengukur aspek seperti kognitif, afektif dan psikomotorik sebagaimana yang dikehendaki oleh TIK.

Sedangkan validitas empirik adalah validitas yang diperoleh dari hasil analisis yang dilakukan terhadap data yang bersumber dari hasil pengamatan lapangan. Ada dua macam validitas empirik, yaitu validitas ramalan dan validitas bandingan. Suatu tes memiliki validitas ramalan jika secara tepat mampu meramalkan keberhasilan siswa untuk masa mendatang. Dan memiliki validitas bandingan jika memiliki ketepatan atau kejituan jika dibandingkan dengan tes sejenis pada kurun waktu yang sama.

Dalam penelitian ini yang menjadi fokus adalah validitas item tes itu sendiri. Karena secara teoritik suatu butir soal dikatakan memiliki validitas yang baik jika memberikan dukungan yang besar terhadap tes secara

42) *Ibid*, hal.226

43) Wayan Nurkencana, *Evaluasi Pendidikan*, hal.127

keseluruhan. Sebab tinggi rendahnya skor total akan sangat dipengaruhi oleh skor-skor tiap butirnya.⁴⁴

Untuk mencari besarnya validitas item tersebut dapat digunakan rumus korelasi point biserial sebagai berikut;

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

dimana,

r_{pbi} = koefisien korelasi point biserial yang dipandang sebagai koefisien validitas

M_p = skor rata-rata hitung yang dimiliki oleh testee, yang untuk item bersangkutan telah dijawab benar

M_t = skor rata-rata dari skor total

SD_t = deviasi standar dari skor total

P = proporsi testee yang menjawab benar butir soal yang bersangkutan

q = proporsi testee yang menjawab salah item soal yang bersangkutan

Adapun langkah-langkah yang ditempuh adalah sebagai berikut;

1. menyiapkan tabel perhitungan dalam rangka analisis validitas item.
2. mencari mean dari skor total dengan rumus :

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

44) *Ibid*, hal.48

Dimana X_t adalah jumlah skor total, dan N adalah jumlah testee.

3. mencari deviasi standar total yaitu:

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

4. mencari M_p .
5. menghitung koefisien korelasi point biserial

Koefisien korelasi selalu berada antara $-1,0$ s.d. $+1,0$. Koefisien negatif menunjukkan hubungan kebalikan, sedangkan positif menunjukkan adanya kesejajaran. Untuk menginterpretasi besarnya koefisien korelasi adalah sebagai berikut;

1. jika berada antara $0,0 - 0,2$ berarti korelasi sangat rendah,
2. jika berada antara $0,2 - 0,4$ berarti korelasi rendah,
3. jika berada antara $0,4 - 0,7$ berarti korelasi cukup,
4. jika berada antara $0,7 - 0,9$ berarti korelasi tinggi, dan
5. jika berada antara $0,9 - 1,0$ berarti korelasi sempurna.

6. Analisis Reliabilitas Tes

Suatu tes dikatakan memiliki reliabilitas yang tinggi jika tes tersebut dapat memberikan hasil tes yang konsisten (ajeg).⁴⁵ Tes yang reliabel adalah bahwa skor-skor yang dicapai oleh testee yang sejenis akan menunjukkan kestabilan, kapan saja dan dimana saja tes itu diselenggarakan.

45) Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, hal. 83

Sehubungan dengan itu Scarvia. B. Anderson cs menyatakan bahwa persyaratan bagi tes yang baik adalah validitas dan Reliabilitas, ini penting. Dalam hal ini validitas lebih penting dan reliabilitas perlu, karena menyokong terbentuknya validitas. Sebuah tes reliabel akan tetapi mungkin tidak valid, sebaliknya sebuah tes yang valid biasanya reliabel.⁴⁶

Untuk melihat Reliabilitas tes pada dasarnya dilihat dari kesejajaran hasil. Teknik yang digunakan untuk itu menggunakan rumus korelasi product moment. Ada beberapa metode untuk mengukurnya, yaitu:

1. Metode bentuk paralel

Tes paralel atau tes ekuivalen adalah dua tes yang mempunyai kesamaan tujuan, tingkat kesukaran dan susunan, tetapi butir-butir soalnya berbeda.⁴⁷ Dalam menggunakan metode paralel tester harus menyiapkan dua tes yang dicobakan sebanyak dua kali. Apabila korelasinya tinggi maka tes itu reliabel.

Penggunaan metode ini baik karena testee dihadapkan pada dua macam tes, maka tak ada faktor pengaruh sebab soalnya berbeda. Kelemahan metode ini adalah bahwa tester harus menyusun dua seri tes dan memerlukan waktu yang lama untuk mencobakan tes itu.

46) *Ibid* , hal. 83

47) *Ibid*, hal.87.

2. Metode tes ulang.

Dalam menggunakan metode ini tester hanya memiliki satu tes yang diujikan dua kali. Metode ini disebut juga single test-double trial. Selanjutnya hasil tes itu dihitung korelasinya.⁴⁸

Untuk tes yang banyak mengungkap pengetahuan dan pemahaman cara ini kurang mengena, karena testee masih ingat dengan item soalnya. Jika tenggang waktu tes pertama dan kedua terlalu sempit testee tentu masih banyak ingat soalnya. Namun jika tenggang waktunya terlalu lama maka kondisi sudah berbeda sehingga dapat mempengaruhi Reliabilitas.

Pada umumnya hasil tes yang kedua lebih baik dari hasil tes pertama. Hal ini biasa saja, yang penting adalah adanya kesejajaran hasil yang ditunjukkan oleh koefisien korelasinya

3. Metode single test-single trial

Uji reliabilitas tes dilakukan dengan mengadakan pengukuran terhadap sekelompok siswa, dimana pengukuran itu hanya dilakukan sekali dan dengan satu macam tes. Berdasarkan analisa pengukuran itu akan dapat diketahui Reliabilitas tes tersebut

Dalam aplikasinya, metode ini dapat dilakukan dengan salah satu dari formula berikut ini;

- a. formula dari Spearman dan Brown
- b. formula dari Flanagan,

48) *Ibid*, hal.88.

- c. formula dari Rulon,
- d. formula dari Kuder Richardson,
- e. formula dari C. Hoyt.⁴⁹

Walaupun kelima formula itu sama-sama termasuk dalam pendekatan single test-single trial, namun memiliki perbedaan-perbedaan. Dengan menggunakan formula Flanagan, Spearman & Brown, dan formula Rulon, maka penentuan Reliabilitas tes bentuk obyektif dilakukan dengan cara membelah dua tes (split half technique). Sedangkan dengan menggunakan formula Kuder-Richardson dan formula C. Hoyt maka penentuan Reliabilitas tes bentuk obyektif tidak menggunakan teknik belah dua.

Sekalipun menggunakan teknik belah dua, ketiga formula (Spearman, Flanagan, Rulon) berbeda dalam sasaran yang dijadikan landasan berpijak. Formula Spearman dan Brown didasarkan pada korelasi antara belahan satu dan dua. Formula Flanagan didasarkan pada deviasi. Dan formula Rulon didasarkan pada selisih skor belahan I dan II.

Dan walaupun formula Kuder-Richardson dan formula C. Hoyt sama-sama tidak menggunakan teknik belah dua, namun keduanya juga berbeda prinsip. Reliabilitas tes dengan menggunakan formula Kuder-Richardson dilakukan dengan menganalisa secara langsung terhadap skor-skor item yang bersangkutan. Sedangkan dengan formula C. Hoyt sasarannya adalah interaksi

49) Anas Sudijono, *Pengantar Evaluasi Pendidikan*, hal.214.

antara testee dengan item itu sendiri. Teknik ini dikenal dengan teknik analisis varian (ANAVA).⁵⁰

Untuk mencari Reliabilitas tes bentuk obyektif dalam skripsi ini digunakan formula Flanagan, yaitu dengan rumus;

$$r_{11} = 2 \left(1 - \frac{S_1^2 + S_2^2}{S_t^2} \right)$$

dimana,

r_{11} = koefisien Reliabilitas tes secara keseluruhan

2 & 1 = bilangan konstan

S_1^2 = jumlah kuadrat deviasi dari skor-skor hasil tes yang termasuk pada belahan I

S_2^2 = jumlah kuadrat deviasi dari skor-skor hasil tes yang termasuk belahan II

S_t^2 = jumlah kuadrat total deviasi dari skor-skor belahan I dan II

Selanjutnya untuk interpretasi hasil perhitungan maka harus berkonsultasi pada tabel nilai "r" product moment.

50) *Ibid*, hal.214-215.

I. SISTEMATIKA PEMBAHASAN

Agar pembahasan dapat runtun, logis dan sistematis, maka diperlukan adanya acuan pikir. Sehingga dapat membantu dan mempermudah dalam memahami dan menerima gagasan-gagasan yang akan dituangkan. Oleh sebab itu penulisan ini akan disusun dengan sistematika sebagai berikut:

Bab I tentang Pendahuluan yang didalamnya terkandung Pengertian Judul dan Penegasan Istilah, Latar Belakang Masalah, Perumusan Masalah, Hipotesa, Tujuan dan Kegunaan Penelitian, Tinjauan Pustaka dan Sistematika Pembahasan.

Bab II tentang Pembelajaran Bahasa Arab di Kelas Dua yang berisi keadaan guru dan siswa, proses pembelajaran bahasa arab, problematika pembelajaran bahasa arab, usaha peningkatan mutu pembelajaran bahasa arab. dan pelaksanaan tes sumatif bahasa arab.

Kemudian Bab III tentang Analisis Kualitas Tes Sumatif Bidang Studi Bahasa Arab yang terdiri dari analisis derajat kesukaran item, analisis daya pembeda item, analisis fungsi distraktor item, analisis validitas item tes, dan analisis reliabilitas tes.

Bab terakhir adalah Bab. VI tentang Penutup yang berisi kesimpulan, saran-saran dan penutup.

Perlu juga dikemukakan bahwa sebelum semua bab tersebut dipaparkan, terdapat Halaman Judul, halaman Noto Dinas, halaman Pengesahan, halaman Motto, halaman Persembahan, halaman Kata Pengantar, halaman Daftar Isi, dan halaman Daftar Tabel dan Grafik. Kemudian bagian akhir skripsi berisi Daftar Pustaka, Lampiran-Lampiran dan Curriculum Vitae.

BAB IV

PENUTUP

A. Kesimpulan

Berdasarkan seluruh analisa terhadap item tes sumatif cawu I bidang studi bahasa arab kelas dua di MAN Yogyakarta II tahun pelajaran 1999/2000, dapat diambil suatu kesimpulan bahwa tes sumatif bidang studi bahasa arab tersebut memiliki kualitas yang kurang memadai sebagai alat evaluasi.

Hal ini jelas terlihat dari jumlah item yang memiliki kualitas baik ternyata hanya dimiliki oleh dua buah item, dan empat belas item berkualitas cukup. Sedangkan dua puluh empat item lainnya memiliki kualitas yang kurang dan jelek

Dengan demikian hipotesa yang penulis ajukan sebelumnya , bahwa tes sumatif cawu I baidang studi bahasa arab kelas duadi MAN Yogyakarta II tahun pelajaran 1999/2000 – khusus pada tes bentuk obyektif- memiliki kualitas yang kurang baik sebagai alat evaluasi, telah terbukti.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA



B. Saran – Saran

Atas dasar analisa dan kesimpulan yang telah penulis paparkan, sekiranya perlu penulis sampaikan beberapa saran guna perbaikan dan penyempurnaan tes tersebut sebagai berikut :

1. Item yang telah baik hendaknya disimpan dalam bank soal agar dapat dipergunakan kembali bila diperlukan, sedangkan item yang kurang baik agar diperbaiki dan disempurnakan,
2. Dalam pembuatan dan penyusunan tes hendaknya disesuaikan dengan kemampuan siswa yang notabenenya mayoritas mereka berasal dari SLTP, dan tidak memiliki basic bahasa arab yang memadai.
3. Guru bidang studi, sekolah dan pihak-pihak yang berkompeten terhadap pendidikan bahasa arab hendaknya lebih memberikan perhatian dan keseriusan guna mencapai tujuan program pembelajaran bahasa arab itu sendiri.
4. Siswa sangat memerlukan motivasi ekstrinsik agar mereka gemar dan giat mempelajari bahasa arab.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

C. Kata Penutup

Dengan segala kesungguhan untuk segera merampungkan tugas mulia ini, akhirnya penulis dengan segenap hati kembali menghaturkan rasa terima kasih kepada seluruh pihak yang telah sangat banyak membantu studi dan kelancaran penulisan skripsi ini.

Penulis menyadari sepenuh hati bahwa pasti terdapat kekurangan dan kesalahan dalam skripsi ini, baik dalam hal redaksi bahasa ataupun dalam pengolahan data atau mungkin pula dalam hal lainnya. Karena itu penulis sangat berterimakasih jika mendapatkan kritik guna penyempurnaan penulisan selanjutnya.

Dan penulis mengharapkan agar ada pihak lain yang meneliti ulang skripsi ini ataupun melakukan penelitian sejenis, mengingat begitu urgennya evaluasi dalam dunia pendidikan.

Akhirnya penulis mengakhiri skripsi ini dengan dzikir “Alhamdu lillahi rabbi al-alamin”, sebagai rasa syukur dan pengabdian penulis padaNya. Karena sangat penulis sadari bahwa segala yang telah penulis capai selama ini takkan pernah terwujud tanpa ridlo terstuNya, terima kasih.

DAFTAR PUSTAKA

- Arifin, Zaenal, Drs., 1991, *Evaluasi Instruksional Prinsip Teknik Prosedur*, Bandung, Remaja Rosdakarya.
- Arikunto, Suharsimi, Dr., 1993, *Prosedur Penelitian Suatu Pendekatan Praktik*, Jakarta, Rineka Cipta.
- Arikunto, Suharsimi, Dr., 1995, *Dasar-Dasar Evaluasi Pendidikan*, Jakarta, Bumi Aksara.
- Buchori, Mochtar, M.ed., 1980, *Teknik-Teknik Evaluasi Dalam Pendidikan*, Bandung, Jemmars.
- Hadi, Sutrisno, Prof., Drs., MA., 1980, *Metodologi Riset*, Yogyakarta, Fak. Psikologi UGM.
- Mudjijo, Drs., M.ed., 1995, *Tes Hasil Belajar*, Jakarta, Bumi Aksara.
- Purwanto, Ngalim, Drs., 1984, *Prinsip-Prinsip Dan Teknik Evaluasi Pendidikan*, Bandung, Remaja Rosdakarya.
- Singarimbun, Masri & Sofian Effendi, 1989, *Metode Penelitian Survei*, Jakarta, LP3ES.
- Slameto, Drs., 1988, *Evaluasi Pendidikan*, Jakarta, Bina Aksara.
- Sudijono, Anas, Prof., Drs., 1986, *Teknik Evaluasi Pendidikan Suatu Pengantar*, Yogyakarta, UD.Rama.
- Sudijono, Anas, Prof., Drs., 1998, *Pengantar Evaluasi Pendidikan*, Jakarta, Rajawali Pers.
- Sumardi, Mulyanto, Dr., 1975, *Pengajaran Bahasa Asing Suatu Pendekatan Metodologi*, Jakarta, Bulan Bintang.
- Tim Penyusun Kamus P3B, 1990, *Kamus Besar Bahasa Indonesia*, Jakarta, Balai Pustaka.
- Usman, Uzer, Drs., 1996, *Menjadi Guru Profesional*, Bandung, Remaja Rosdakarya.

الشيخ مصطفى علاي، جامع الدروس العربية، بيروت، المكتبة العصرية،
١٩٩٤، ص. ٧٠