

ANALISIS SOAL TOAFL IAIN CURUP

Partomuan Harahap

Institut Agama Islam Negeri Curup

Email: Partomuan73@gmail.com

Abstrak: Mengingat urgennya TOAFL dalam mengukur keterampilan bahasa Arab mahasiswa di IAIN Curup, peneliti merasa perlu untuk melakukan evaluasi terhadap soal TOAFL. Agar tes ini tidak kehilangan fungsi yang seharusnya yaitu menjadi alat pengukur pemahaman mahasiswa terhadap bahasa Arab. Tujuan penelitian ini adalah untuk mengetahui tingkat kesulitan butir soal, daya pembeda butir soal, validitas butir soal, dan reliabilitas butir soal TOAFL IAIN Curup. Metode penelitian yang digunakan adalah penelitian deskriptif kuantitatif. Teknik pengumpulan data menggunakan tes, wawancara, dokumentasi dan observasi. Hasil penelitian ini menunjukkan bahwa tingkat kesukaran soal TOAFL, terdapat 15 soal kategori mudah, 71 soal kategori sedang, dan 54 soal kategori sukar. Persentase tingkat kesukaran soal, soal masuk dalam kategori baik, karena lebih 50% indeks kesukarannya memenuhi standar. Untuk daya pembeda, tidak terdapat soal kategori baik, 48 soal kategori sedang, 56 soal kategori kurang, 4 soal tidak memiliki daya pembeda, dan 32 soal yang negatif. Melihat persentase daya pembeda soal, soal pilihan ganda kategori tidak baik, karena 0% soal kategori baik. Untuk validitas adalah 78 soal kategori valid dan 62 kategori tidak valid. Persentase validitas soal TOAFL IAIN Curup tersebut menunjukkan lebih 50% soal kualitas valid. Sedangkan reliabilitas menunjukkan bahwa soal TOAFL memiliki kategori reliabilitas yang tinggi yaitu 0,87.

Kata Kunci: Analisis, Butir Soal, TOAFL

PENDAHULUAN

Salah satu tugas penting yang sering dilupakan oleh staf pengajar adalah tugas melakukan evaluasi terhadap alat pengukur yang telah digunakan untuk mengukur keberhasilan belajar dari para peserta didiknya. Alat pengukur dimaksud adalah tes hasil belajar, yang batang tubuhnya terdiri dari kumpulan butir-butir soal.¹

Pembelajaran tanpa kegiatan evaluasi akan kehilangan makna. Sebab tenaga pengajar tidak akan memperoleh informasi penting tentang tingkat pencapaian tujuan, tingkat penguasaan materi belajar, kekuatan, kelemahan siswa dalam belajar, serta kekuatan-kelemahan mereka dalam proses pembelajaran yang dikembangkan. Walaupun evaluasi dianggap penting dan sudah merupakan pekerjaan rutin tenaga pengajar, namun dalam kenyataan sehari-hari di lapangan sistem evaluasi dalam pembelajaran bukan berarti tanpa persoalan. Berdasarkan pengamatan sepintas di lapangan, beberapa persoalan tersebut paling tidak berkaitan dengan pemahaman konsep dasar evaluasi, pelaksanaan dan pemanfaatannya, serta evaluasi program pengajaran.

¹ Anas Sudijono, Pengantar Evaluasi Pendidikan, (Jakarta: PT. Raja Grafindo Persada, 1996), h. 367-368

Selain terhadap proses pembelajaran di kelas, evaluasi juga dilakukan terhadap sebuah tes yang merupakan standar suatu lembaga pendidikan. Diantara bentuk tes tersebut adalah TOAFL. TOAFL adalah singkatan dari “Test of Arabic as a Foreign Language”. Penamaan ini diilhami oleh TOEFL, yang memang telah lebih dahulu eksis. TOAFL dilaksanakan sebagai standarisasi penilaian bahasa Arab. Dengan adanya standarisasi tersebut, diharapkan mampu menekan peserta didik sehingga memiliki keterampilan bahasa Arab yang matang. Standarisasi penilaian bahasa Arab ini telah banyak dilaksanakan di Perguruan-perguruan Tinggi di Indonesia, baik negeri maupun swasta. Salah satu Perguruan Tinggi tersebut adalah IAIN Curup. IAIN Curup telah melaksanakan tes ini terhitung sejak tahun 2015 dibawah pengawasan Unit Pengembangan Bahasa (UPB). Tes ini disusun oleh tim pembuat soal yang telah ditentukan oleh pihak Unit Pengembangan Bahasa (UPB). Tim pembuat soal tersebut adalah dosen-dosen bahasa Arab IAIN Curup. TOAFL dilaksanakan oleh pihak Sekolah Tinggi Agama Islam Indonesia (IAIN) Curup sebagai salah satu syarat bagi mahasiswa yang ingin mengikuti wisuda.

Semenjak keberadaan TOAFL di IAIN Curup, belum pernah dilaksanakan evaluasi terhadap butir soal tersebut. Mengingat urgennya TOAFL dalam mengukur keterampilan bahasa Arab mahasiswa di IAIN Curup, peneliti merasa perlu untuk melakukan evaluasi terhadap soal TOAFL tersebut. Agar tes ini tidak kehilangan fungsi yang seharusnya yaitu menjadi alat pengukur pemahaman mahasiswa terhadap bahasa Arab.

Dalam penelitian ini, peneliti akan melakukan penelitian dengan cermat terhadap butir-butir soal TOAFL IAIN Curup. Identifikasi terhadap setiap butir soal tes hasil belajar itu dilakukan dengan harapan akan menghasilkan berbagai informasi berharga, yang pada dasarnya akan merupakan umpan balik guna melakukan perbaikan, pembenahan, dan penyempurnaan kembali terhadap butir-butir item yang telah dikeluarkan dalam tes TOAFL IAIN Curup, sehingga pada masa-masa yang akan datang tes TOAFL IAIN Curup yang disusun atau dirancang betul-betul dapat menjalankan fungsinya sebagai alat pengukur keterampilan bahasa Arab yang memiliki kualitas yang tinggi.

Adapun rumusan masalah dalam penelitian ini adalah berapakah tingkat kesulitan butir soal, daya pembeda butir soal, validitas butir soal dan reliabilitas butir soal TOAFL IAIN Curup.

Berdasarkan sumber data, metode penelitian yang digunakan adalah penelitian deskriptif kuantitatif yaitu adalah penelitian yang pendeskripsian analisis datanya dinyatakan dalam angka-angka. Data dari penelitian ini berupa skor mentah mahasiswa peserta tes TOAFL IAIN Curup gelombang I tahun 2016 yang kemudian datanya diolah dengan mencari tingkat kesukaran soal, daya beda soal, validitas soal, dan reliabilitas soal untuk mengetahui kualitas soal TOAFL tersebut.

Adapun teknik pengumpulan data yang digunakan dalam penelitian ini adalah tes, wawancara, dokumentasi dan observasi. Sedangkan teknik analisis data yang digunakan dalam penelitian ini adalah :

Untuk menganalisis tingkat kesukaran butir tes dimaksudkan untuk mengetahui seberapa sulit atau mudahnya tes yang telah diselenggarakan. Dengan menggunakan rumus-rumus.

I. EVALUASI

A. Pengertian Evaluasi

Secara harfiah kata evaluasi berasal dari bahasa Inggris *evaluation* yang berarti penilaian atau penaksiran. Dan ditinjau dari sudut bahasa, penilaian diartikan sebagai proses menentukan nilai suatu obyek. Menurut istilah evaluasi merupakan kegiatan yang terencana untuk mengetahui keadaan suatu obyek dengan menggunakan instrumen dan hasilnya dibandingkan dengan tolok ukur untuk memperoleh kesimpulan.²

Evaluasi adalah suatu proses yang sangat penting dalam pendidikan guru, tetapi pihak-pihak yang terkait dalam program itu seringkali melalaikan atau tak menghayati sungguh-sungguh proses evaluasi tersebut.³

Evaluasi adalah proses untuk melihat apakah perencanaan yang sedang di bangun berhasil, sesuai dengan harapan awal atau tidak.⁴

Dengan demikian, evaluasi berarti menentukan sampai seberapa jauh sesuatu itu berharga, bermutu, atau bernilai. Evaluasi terhadap hasil belajar yang dicapai siswa dan terhadap proses pembelajaran mengandung penilaian terhadap hasil belajar atau proses belajar itu, sampai seberapa jauh keduanya dapat dinilai baik.

B. Tujuan Evaluasi

Adapun tujuan evaluasi dapat dibedakan menjadi dua, yaitu tujuan umum dan tujuan khusus, yaitu sebagai berikut:

1. Tujuan Umum

Tujuan evaluasi pembelajaran adalah untuk mengetahui keefektifan dan efisiensi sistem pembelajaran, baik yang menyangkut tentang tujuan materi, metode, media sumber belajar, lingkungan maupun sistem penilaian itu sendiri.⁵

Tujuan utama melakukan evaluasi dalam proses belajar mengajar adalah untuk mendapatkan informasi yang akurat mengenai tingkat pencapaian tujuan instruksional oleh siswa sehingga dapat diupayakan tindak lanjutnya. Tindak lanjut termaksud merupakan fungsi evaluasi dan dapat berupa.⁶

- a) Penempatan pada tempat yang tepat
- b) Pemberian umpan balik
- c) Diagnosis kesulitan belajar siswa
- d) Penentuan kelulusan

2. Tujuan Khusus

2 Sulistyorini, Evaluasi Pendidikan Dalam Meningkatkan Mutu Pendidikan, (Yogyakarta: Teras, 2009), hlm. 49

3 Oemar Hamalik, Pendidikan Guru Berdasarkan Pendekatan Kompetensi, (Jakarta: PT.Bumi Aksara) hlm. 180

4Mardiah Hayati, Desain Pembelajaran, (Pekanbaru: Yayasan Pustaka Riau, 2009), hlm. 51

5 Zainal Arifin, Evaluasi Pembelajaran, (Jakarta: Direktorat Jenderal Pendidikan Islam Departemen Agama RI, 2009)

6 Daryanto, Evaluasi Pendidikan, (Jakarta: Rineka Cipta, 2010) hlm, 11

Tujuan khusus evaluasi pembelajaran disesuaikan dengan jenis evaluasi pembelajaran itu sendiri, seperti evaluasi perencanaan dan pengembangan, evaluasi monitoring, evaluasi dampak, evaluasi efisinensi-ekonomi, dan evaluasi program komprehensif.⁷

C. Fungsi Evaluasi

Secara umum, evaluasi sebagai suatu tindakan atau proses setidaknya tidaknya memiliki tiga macam fungsi pokok, yaitu:⁸

1. Mengukur kemajuan.
2. Menunjang penyusunan rencana.
3. Memperbaiki atau melakukan penyempurnaan kembali.

Adapun secara khusus, fungsi evaluasi dalam dunia pendidikan dapat ditilik dari tiga segi, yaitu:⁹

1. Segi Psikologis

Secara psikologis, kegiatan evaluasi dalam bidang pendidikan di sekolah dapat disoroti dari dua sisi, yaitu dari sisi peserta didik dan dari sisi pendidik.

Bagi peserta didik, evaluasi pendidikan secara psikologis akan memberikan pedoman atau pegangan batin kepada mereka untuk mengenal kapasitas dan status dirinya masing-masing di tengah-tengah masyarakat atau kelasnya. Dengan dilakukannya evaluasi terhadap hasil belajar siswa misalnya, maka para siswa akan mengetahui apakah dirinya termasuk siswa yang berkemampuan tinggi, berkemampuan rendah, ataukah berkemampuan rendah. Demikian pula dengan dilakukannya evaluasi tersebut maka para siswa yang bersangkutan akan menjadi tahu atau mengerti, di manakah posisi (letak) dirinya di tengah-tengah teman-temannya. Apakah ia termasuk siswa kelompok atas (pandai), kelompok tengah (Sedang/biasa-biasa saja), ataukah termasuk kelompok bawah (bodoh).

Bagi pendidik, evaluasi pendidikan akan memberikan kepastian atau ketetapan hati kepada diri pendidik tersebut, sudah sejauh manakah kiranya usaha yang telah dilakukannya selama ini telah membawa hasil, sehingga ia secara psikologis memiliki pedoman atau pegangan batin yang pasti guna menentukan langkah-langkah apa saja yang dipandang perlu dilakukan selanjutnya.¹⁰

2. Segi Didaktik

Bagi peserta didik, secara didaktik evaluasi pendidikan akan dapat memberikan dorongan kepada mereka untuk dapat memperbaiki, meningkatkan dan mempertahankan prestasinya. Evaluasi hasil belajar itu misalnya, akan menghasilkan nilai-nilai hasil belajar untuk masing-masing individu siswa, ada siswa yang nilainya jelek (prestasinya rendah), karena

⁷Ibid.,

⁸Anas Sudijono, Pengantar Evaluasi Pendidikan, (Yogyakarta: PT Raja Grafindo Persada, 2009), hlm. 8

⁹Ibid., hlm. 10-14

¹⁰Ibid.,

itu siswa tersebut terdorong untuk memperbaikinya, agar untuk waktu-waktu yang akan datang nilai hasil belajarnya tidak sejelek sekarang.

Bagi pendidik, secara didaktik evaluasi pendidikan itu setidaknya-tidaknya memiliki lima macam fungsi, yaitu:¹¹

- a) Memberikan landasan untuk menilai hasil usaha (prestasi) yang telah dicapai oleh peserta didiknya.
 - b) Memberikan informasi yang sangat berguna, guna mengetahui posisi masing-masing peserta didik di tengah-tengah kelompoknya.
 - c) Memberikan bahan yang penting untuk memilih dan kemudian menciptakan status peserta didik.
 - d) Memberikan pedoman untuk mencari dan menemukan jalan keluar bagi peserta didik yang memang memerlukannya.
 - e) Memberikan petunjuk tentang sudah sejauh manakah program pengajaran yang telah ditentukan telah dapat dicapai.
3. Segi Administratif

Adapun secara administratif, evaluasi pendidikan setidaknya-tidaknya memiliki tiga macam fungsi, yaitu:

- a) Memberikan laporan
Dengan dilakukan evaluasi, akan dapat disusun dan disajikan laporan mengenai kemajuan dan perkembangan peserta didik setelah mereka mengikuti proses pembelajaran dalam jangka waktu tertentu.
- b) Memberikan bahan-bahan keterangan (Data)
Setiap keputusan pendidikan harus didasarkan kepada data yang lengkap dan akurat. Dalam hubungannya ini, nilai-nilai hasil belajar peserta didik yang diperoleh dari kegiatan evaluasi adalah merupakan data yang sangat penting untuk keperluan pengambilan keputusan pendidikan dan lembaga pendidikan: apakah seseorang peserta didik dapat dinyatakan tamat belajar, dapat dinyatakan naik kelas, tinggal kelas, lulus atukah tidak lulus, dan sebagainya.
- c) Memberikan gambaran
Gambaran mengenai hasil-hasil yang telah dicapai dalam proses pembelajaran tercermin antara lain dari hasil-hasil belajar para peserta didik setelah dilakukannya evaluasi hasil belajar.

D. Prinsip Evaluasi

Evaluasi memiliki beberapa prinsip, diantaranya adalah sebagai berikut:

1. Prinsip Totalitas, Keseluruhan, atau Komprehensif

Evaluasi hasil belajar harus dilakukan untuk menggambarkan perkembangan atau perubahan tingkah laku peserta didik secara menyeluruh, artinya evaluasi mampu mengungkapkan aspek kognitif, afektif, dan psikomotorik.

2. Prinsip Kesenambungan

Evaluasi yang baik dilakukan secara teratur, berkesinambungan dari waktu ke waktu, terencana dan terjadwal. Evaluasi yang demikian akan menggambarkan perkembangan pesertra didik dari waktu ke waktu.

¹¹ Ibid.,

3. Prinsip Objektivitas

Evaluasi yang baik harus terlepas dari kepentingan subyek. Hasil evaluasi tersebut harus menggambarkan kondisi peserta didik secara obyektif.¹²

4. Prinsip Kesesuaian dengan tujuan

Evaluasi harus sesuai dengan tujuan pembelajaran yang digunakan. Dalam artian evaluasi tersebut harus mempunyai gambaran yang sama dengan tujuan pembelajaran yang digunakan dan berhubungan dengan target atau tujuan akhir.

5. Prinsip Ilmiah

Di dalam alat evaluasi harus terpenuhi sifat valid (jelas menguasai suatu kemampuan), reliabel dan obyektif. Yang dimaksud dengan validitas alat evaluasi adalah kita mengukur apa yang bisa kita ukur dengan alat yang tepat, ini bisa meningkat dengan tingginya kesesuaian antara alat evaluasi dengan tujuan.

6. Prinsip Kerjasama

Yang dimaksud dengan kerjasama di sini adalah bahwa evaluasi harus dilakukan dengan melibatkan berbagai individu atau kelompok yang terkait agar bisa mencapai sasaran yang ingin dicapai. Evaluasi bukan hanya tanggung jawab guru, tetapi merupakan tanggung jawab semua pihak yang terkait semisal tenaga administrasi, para siswa, bahkan penjaga sekolah.

7. Prinsip Ekonomis

Evaluasi harus dilakukan dengan cara yang efektif dan efisien, tanpa menghambur-hamburkan waktu, tenaga, dan biaya. Tetapi dengan alasan ekonomis juga tidak boleh mengorbankan aspek akademisnya. Evaluasi hendaklah bersifat ekonomis dalam segala hlm, baik dari segi kegiatan, waktu maupun biaya.¹³

II. TOAFL

A. Pengertian TOAFL

TOAFL adalah singkatan dari "*Test of Arabic as a Foreign Language*". Penamaan ini diilhami oleh TOEFL yang memang telah lebih dulu eksis. Pengambilan ini memang dimaksudkan agar TOAFL lebih mudah diucapkan dan lebih cepat dikenal oleh banyak orang, meskipun terkesan "mirip" TOEFL. Selain itu mengapa menggunakan TOAFL, adapun alasannya adalah sebagai berikut:¹⁴

1. Selama ini (UIN, IAIN, IAIN) belum mempunyai tes bahasa Arab standar seperti TOEFL.
2. Tes ini mampu mengukur tingkat kemampuan (reseptif) seseorang dalam berbahasa Arab.
3. Tes ini mudah dikerjakan dan mudah dikoreksi.

¹² Rubiyanto, Rubino dan Sri Hartini, Evaluasi Pendidikan, (Surakarta: Progeram Akta Mengajar FKIP UMS, 2005), hlm.12

¹³ Moh. Matsna, Erta Mahyudin, Pengembangan Evaluasi dan Tes Bahasa Arab.(Tangerang Selatan: Alkitabah, 2012), hlm 32

¹⁴Muhbib Abdul Wahab, slide Presentasi Pembelajaran TOAFL di MAN 4 Jakarta tahun 2010

4. Jawaban dan hasil penilaiannya bersifat objektif dan pasti.
5. Materi tes ini cukup komprehensif, dan menurut pemahaman dan penguasaan mufradat yang cukup banyak.

B. Tujuan TOAFL

TOAFL merupakan standarisasi penilaian bahasa Arab. Diantara tujuan TOAFL adalah:

1. Menetapkan norma-norma keterampilan bahasa Arab yang kelak dijadikan sebagai pedoman standarisasi kelulusan bahasa Arab.
2. Memberlakukan standar baku kelulusan bahasa Arab dalam TOAFL.
3. Meningkatkan kualitas kemampuan dan penguasaan bahasa Arab bagi lulusan program S1, S2, dan S3 seluruh lembaga pendidikan di bawah naungan Kementerian Agama strata 1 sampai doktor.¹⁵
4. Menumbuhkan kesadaran peserta studi Islam dan ilmu pengetahuan akan signifikansi bahasa Arab sebagai media utama studi Islam dan ilmu pengetahuan.
5. Memberdayakan kemampuan memahami bahasa Arab.
6. Meningkatkan penguasaan kebahasaaraban berwacana studi Islam.¹⁶

C. Sejarah singkat TOAFL

TOAFL "*Test of Arabic as a Foreign Language*" dibuat dan diterbitkan pertama kali pada tahun 1998 oleh sebuah Tim penyusun yang diprakarsai oleh Muhib Abdul Wahab dan Suwito. Tim ini beranggotakan: Chotibul Umam, HD. Hidayat, Rofi'i, Akrom Malibary, Muhammad Matsna, Satria Effendi, dan Abdul Kadir Al-Habsyi.¹⁷

Tujuan awal pembentukan tim ini adalah untuk menyiapkan bahan tes standar bagi mahasiswa S1 dan S3 yang akan menyelesaikan studinya. Pada tahun 1999/2000, TOAFL "*Test of Arabic as a Foreign Language*". mulai digunakan sebagai salah satu materi tes, ujian masuk Program S2 dan S3 IAIN (kini UIN) Jakarta. Mulai 2005, Program S1 diwajibkan mengikuti TOAFL. Pada tahun 2000/2001. TOAFL juga digunakan sebagai materi tes masuk di beberapa Program Pascasarjana diluar UIN Jakarta, Seperti: PPs. IAIN Palembang, IAIN Lampung, IAIN Mataram, IAIN Padang, PPs. Studi Islam Universitas Muhammadiyah Jakarta, dan Diklat Diknas dan Deplu.¹⁸

D. Visi dan Misi TOAFL

Adapun visi dan misi TOAFL adalah sebagai berikut:

Visi : Menjadikan bahasa Arab sebagai bahasa studi Islam dan ilmu pengetahuan

Misi :

- Standarisasi dan sertifikasi tingkat kemampuan bahasa Arab peserta studi Islam dan ilmu pengetahuan di Indonesia, khususnya untuk program S1, S2, dan S3.

15 Muhammad Barmawi, Lulus TOAFL dengan Mudah dan Memuaskan, (Yogyakarta: DIVA Press, 2011), hlm. 11

16 Muhib Abdul Wahab, Op.cit.,

17 Muhib Abdul Wahab, Ibid.,

18 Muhib Abdul Wahab, Ibid

- Mensosialisasikan model evaluasi kemampuan bahasa Arab yang memungkinkan penstudi Islam dan ilmu pengetahuan mengembangkan kemahirannya dalam bahasa Arab, terutama kemampuan memahami teks Arab.

III. ANALISIS BUTIR SOAL

Analisis butir Soal yaitu salah satu kegiatan yang penting untuk dilaksanakan dalam rangka memperbaiki mutu suatu soal, baik mutu keseluruhan soal atau mutu tiap butir soal. Soal tes sebagai alat evaluasi diharapkan dapat memberikan nilai atau skor yang objektif dan akurat. Oleh karena itu perlu memperhatikan prinsip-prinsip dasar pembuatan soal dan perlu mengetahui ciri-ciri soal yang baik. Soal tes yang baik dapat dimasukkan dalam bank soal atau kumpulan soal sedangkan yang buruk sebaiknya tidak dipakai lagi. Analisis soal tes dilakukan untuk mengetahui baik buruknya suatu soal tes. Analisis soal tes yang dilakukan dalam penelitian ini meliputi tingkat kesulitan, daya pembeda, validitas, dan reliabilitas.

A. Uji Terhadap Butir Soal

Uji terhadap butir soal meliputi beberapa hal yaitu :

1. Tingkat Kesukaran

Bermutu atau tidaknya butir-butir item tes hasil belajar pertama-tama diketahui dari tingkat kesukaran atau taraf kesulitan yang dimiliki oleh masing-masing butir item tersebut. Butir-butir item yang baik jika tingkat kesukaran itemnya sedang¹⁹

Untuk menghitung tingkat kesukaran bentuk pilihan ganda, digunakan rumus sebagai berikut :

Rumus Tingkat Kesulitan :²⁰

$$P = \frac{R}{T}$$

Keterangan :

P = Indeks kesukaran

R = Banyaknya siswa yang menjawab soal itu dengan benar (*right*)

T = Jumlah seluruh peserta tes (*testee*)

Klasifikasi Indeks Kesukaran²¹

Indeks Kesukaran	Kualitas Butir Tes
0,00 s.d 0,30	Sukar
0,31 s.d 0,70	Sedang
0,71 s.d 1,00	Mudah

2. Daya Pembeda Item

Daya pembeda item adalah kemampuan suatu butir item tes hasil belajar untuk dapat membedakan antara *testee* yang berkemampuan tinggi (pandai) dengan *testee* yang berkemampuan

19 Anas Sudijono, Op.cit, hlm. 370

20 Moh. Matsna, Erta Mahyudin, Op.cit, hlm 183

21 Ibid.,

rendah (bodoh) demikian rupa sehingga sebagian besar *testee* yang memiliki kemampuan tinggi untuk menjawab butir item tersebut lebih banyak yang menjawab betul, sementara *testee* yang memiliki kemampuan rendah untuk menjawab butir item tersebut sebagian besar tidak dapat menjawab item dengan betul.

Rumus Daya Pembeda adalah:²²

$$D = \frac{RH-RL}{1/2 T}$$

Keterangan :

D = Daya pembeda

R_H = Kelompok mampu yang menjawab benar

R_L = Kelompok kurang mampu yang menjawab benar

$1/2 T$ = Setengah dari jumlah peserta tes kedua kelompok

Tingkat Daya Pembeda²³

Tingkat Daya Pembeda	Kualitas Butir Tes
0,50 atau lebih	Baik
Antara 0,20 – 0,50	Sedang
Di bawah 0,20	Kurang
0	Tidak ada deskriminasi
- (negatif)	negatif

3. Uji Validitas

Validitas item dari suatu tes yaitu ketepatan mengukur yang dimiliki oleh sebutir item (yang merupakan bagian yang tak terpisahkan dari tes sebagai suatu totalitas), dalam mengukur apa yang seharusnya diukur lewat butir item tersebut.²⁴

Menentukan validitas sebuah tes memerlukan beberapa langkah dan proses yang harus dilakukan sebelum menemukan nilai validitas. Adapun langkah-langkah tersebut sebagai berikut:

1. Setelah berakhirnya pelaksanaan suatu tes, dilakukan koreksi dan menghitung skor yang diperoleh oleh peserta tes.
2. Menyiapkan tabel perhitungan untuk menganalisis validitas butir soal dari soal pertama hingga terakhir.
3. Mencari mean (M_t), dengan rumus:

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

4. Mencari deviasi standar total (SD_t), dengan rumus:

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

5. Menghitung M_p dari butir soal pertama hingga terakhir

22 M. Soenardi Djiwandono, Tes Bahasa dalam Pengajaran. (Bandung: ITB, 2011), hlm. 227

23 Ibid.,1996, hlm. 144

24 Ibid., 182

6. Menghitung koefisien korelasi r_{pbi} dari butir soal pertama hingga terakhir, dengan menggunakan rumus:

$$r_{pbi} = \frac{Mp - Mt}{SDt} \sqrt{\frac{P}{q}}$$

Untuk memberikan interpretasi terhadap hasil dari perhitungan koefisien korelasi ini digunakan db sebesar $(N-2)$, kemudian hasilnya dihubungkan kepada tabel r *product moment* pada taraf 5%. Apabila dari hasil tersebut didapatkan r hitung lebih besar daripada r tabel, maka butir soal tersebut adalah valid, akan tetapi apabila r hitung lebih kecil dari r tabel, maka butir soal tersebut tidak valid.

4. Reliabilitas

Dalam rangka menentukan apakah suatu tes telah memiliki daya mengukur atau reliabilitas yang tinggi atau belum, maka dilakukanlah uji reliabilitas tes. Soal yang baik jika dilakukan beberapa kali pengujian maka hasilnya relatif sama.

Adapun langkah-langkah dalam mendapatkan nilai reliabilitas sebuah tes adalah sebagai berikut:

- 1) Menjumlah skor-skor yang dimiliki oleh butir soal yang bernomor gasal.
- 2) Menjumlahkan skor-skor yang dimiliki oleh butir soal yang bernomor genap.
- 3) Menghitung angka indek korelasi r *product moment* antara X (butir soal yang bernomor gasal) dengan Y (butir soal yang bernomor Genap).
- 4) Setelah menemukan hasil, maka disubstitusikan ke dalam rumus berikut:

$$r_{12} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

- 5) Menghitung koefisien reliabilitas tes (r_{tt} atau r_{11}) dengan menggunakan rumus:

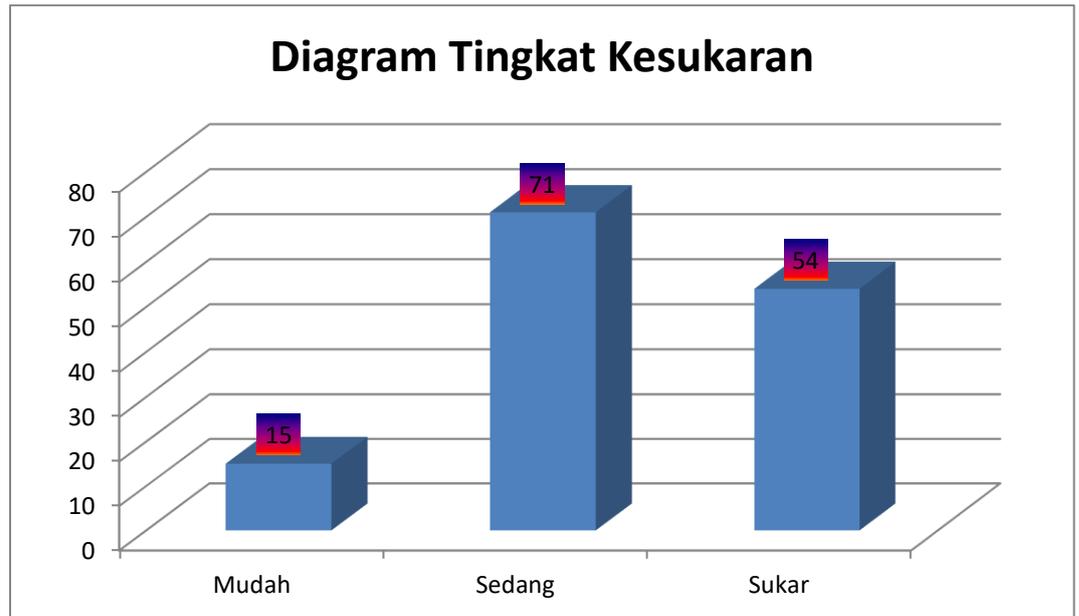
$$r_{11} = \frac{2 r^{11/22}}{1 + r^{11/22}}$$

- 6) Memberikan interpretasi terhadap r_{11}

IV. HASIL PENELITIAN DAN PEMBAHASAN

A. Analisis Tingkat Kesukaran Butir Soal TOAFL IAIN Curup

Perhitungan tingkat kesukaran soal adalah pengukuran seberapa besar derajat kesukaran suatu soal. Jika suatu soal memiliki tingkat kesukaran seimbang (proporsional), maka dapat dikatakan bahwa soal tersebut baik. Suatu soal tes hendaknya tidak terlalu sukar dan tidak pula terlalu mudah.

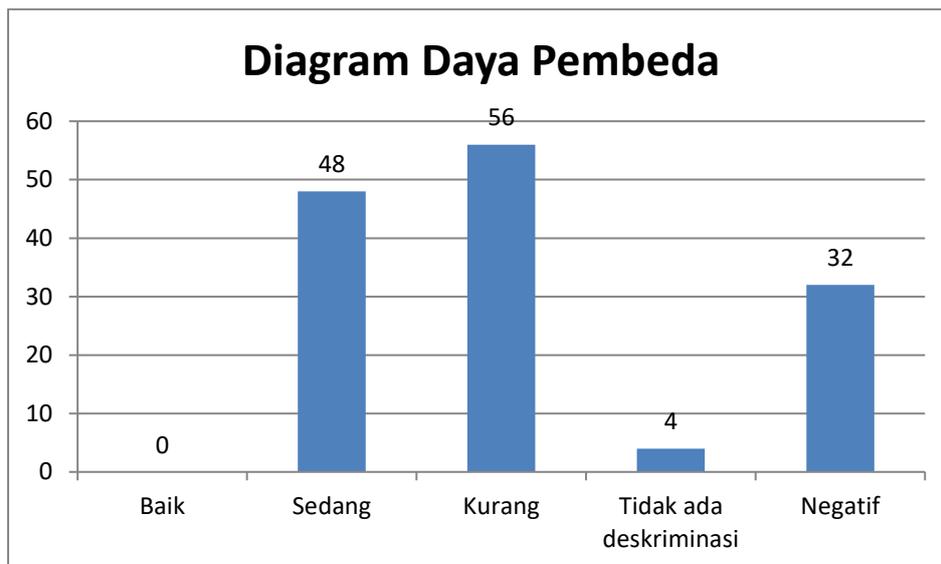


Berdasarkan pemaparan tingkat kesukaran soal, dapat diketahui soal mana yang layak diberikan pada tes berikutnya dan soal mana yang harus dibuang atau diganti. Butir tes yang memiliki tingkat kesukaran mudah, memiliki kemungkinan untuk dibuang atau direvisi agar bisa digunakan lagi dalam ujian yang akan datang, yaitu soal dengan nomor 57,68,92,93,94,95,96,97,98,99,103,106,108,115,120. Butir tes yang memiliki tingkat kesukaran sedang, dapat tetap dipertahankan dan dapat digunakan lagi untuk ujian yang akan datang, yaitu soal nomor 1,7,10,17,20,22,24,31,32,33,34,35,36,37,38,39,40,42,44,46,48,49,50,51,52,53, 55,56,58,61,62,63,64,65,67,69,71,73,74,75,76,78,79,80,83,86,87,90,91,104, 110,111,112,113,114,117,118,121,123,124,125,126,127,128,130,131,132,135, 137,138,139. Butir tes yang memiliki tingkat kesukaran sukar, dapat digunakan lagi tapi harus direvisi atau mengganti soal yang baru, yaitu soal nomor 2,3,4,5,6,,8,9,11,12,13,14,15,16,18,19,21,23,25,26,27,28,29,30,41,43,45,47,54, 59,60,66,70,72,77,81,82,84,85,88,89,100,101,102,105,107,109,116,119,122, 129,133,134,136, 140 .

Dengan melihat persentase proporsi tingkat kesukaran soal pada tabel diatas, soal TOAFL tersebut masuk dalam kategori baik, karena lebih dari 50% indeks kesukarannya memenuhi standar.

B. Analisis Daya Pembeda Butir Soal TOAFL IAIN Curup

Daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang pandai (menguasai materi) dengan siswa yang kurang pandai (kurang/tidak menguasai materi). Indeks daya pembeda biasanya dinyatakan dengan proporsi. Semakin tinggi proporsi itu, maka semakin baik soal tersebut membedakan antara siswa yang pandai dan siswa yang kurang pandai.



Berdasarkan pemaparan daya pembeda soal, dapat diketahui soal mana yang layak diberikan untuk tes berikutnya, dan soal mana yang harus dibuang atau diganti. Butir tes yang memiliki daya pembeda baik, yaitu tes yang masuk dalam kategori tes yang baik atau layak untuk siswa. Butir tes yang memiliki daya pembeda sedang, dapat digunakan lagi tapi harus direvisi lagi. Butir tes yang memiliki daya pembeda kurang, tes tersebut harus dibuang atau direvisi lagi. Selain itu juga ada butir tes yang tidak memiliki daya pembeda bahkan negatif, butir soal tersebut harus dibuang dan tidak layak digunakan kembali.

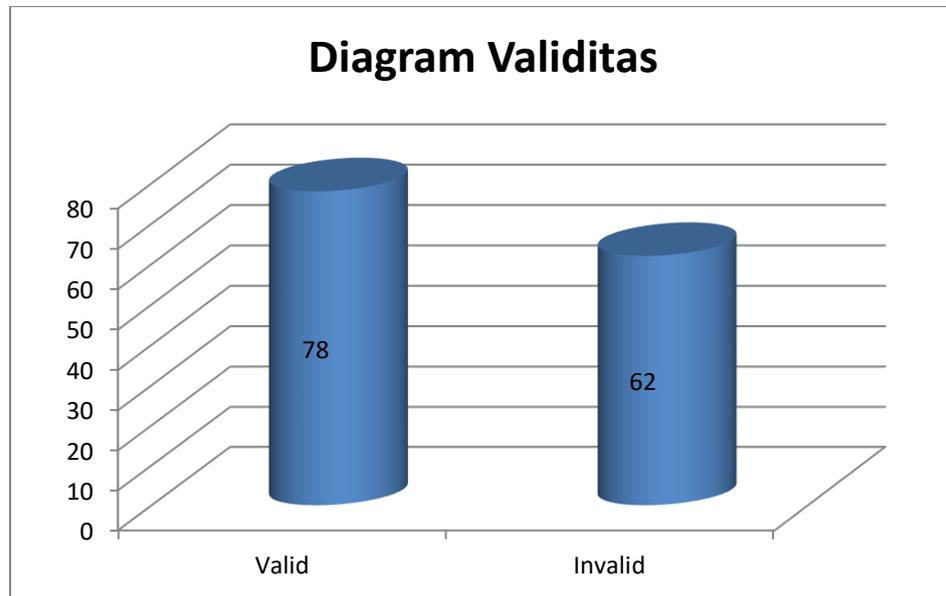
Dari tabel kualifikasi daya pembeda soal TOAFL di atas, soal yang masuk dalam kategori baik, tidak ada. Terdapat 48 soal yang masuk dalam kategori sedang, yaitu nomor 17,20,34,39,44,52,53,56,57,58,59,60,61,62, 63,64,67,69,72,73,74,77,90,91,92,93,94,95,96,98,103,104,106,108,110,111,113,114,117,118,120,123,125,126,127,131,135,139, artinya, soal tersebut dapat digunakan lagi tapi harus direvisi lagi. Terdapat 56 soal yang masuk dalam kategori kurang, yaitu soal nomor 3,9,11,12,13,19,21,22,24,26,27,28,30, 31,32,33,36,37,38,40,41,42,43,46,48,50,54,55,65,66,68,70,75,78,79,80,81,83,85,87,89,97,99,107,109,112,115,121,128,130,132,133,134,137,138,140, artinya, tes tersebut harus dibuang atau direvisi lagi. Kemudian ada 4 butir tes yang tidak memiliki daya pembeda yaitu soal nomor 7,10,51,88 dan ada 32 butir soal yang negatif yaitu soal nomor 1,2,4,5,6,8,14,15,16,18,23,25,29,35,45, 47,49,71,76,82,84,86,100, 101,102,105,116,119,122,124,129,136. Adapun soal yang tidak memiliki daya pembeda sebaiknya dibuang dan tidak digunakan lagi untuk tes berikutnya.

Dengan melihat persentase proporsi daya pembeda soal pada tabel di atas, soal TOAFL tersebut masuk dalam kategori tidak baik, karena 0% butir soal yang memiliki daya pembeda yang baik.

C. Analisis Validitas Soal TOAFL IAIN Curup

Validitas item dari suatu tes yaitu ketepatan mengukur yang dimiliki oleh sebutir item (yang merupakan bagian yang tak terpisahkan dari tes sebagai

suatu totalitas), dalam mengukur apa yang seharusnya diukur lewat butir item tersebut.²⁵



Instrumen yang valid berarti alat ukur yang digunakan untuk mendapatkan data (mengukur) itu valid. Valid berarti instrumen tersebut dapat digunakan untuk mengukur apa yang seharusnya diukur. Kriteria soal yang valid tidak perlu lagi dilakukan revisi, sedangkan soal yang masuk dalam kriteria tidak valid harus dilakukan revisi atau mengganti bentuk soalnya.

Adapun soal yang masuk kriteria valid ada 78 soal yaitu soal nomor 3,9,12,17,20,21,28,33,34,37,38,39,41,43,44,48,50,52,53,54,55,56,57,58,59,60,61,62,63,64,65,67,68,69,70,72,73,74,75,77,79,81,85,87,90,91,92,93,94,95,96,97,98,99,103,104,106,107,108,109,110,111,112,113,114,115,117,118,120,123,125,126,127,131,135,137,139,140, soal-soal tersebut tidak perlu direvisi. Sedangkan yang harus dilakukan revisi atau diganti dengan soal lain ada 62 soal, yaitu soal nomor 1,2,4,5,6,7,8,10,11,13,14,15,16,18,19,22,23,24,25,26,27,29,30,31,32,35,36,40,42,45,46,47,49,51,66,71,76,78,80,82,83,84,86,88,89,100,101,102,105,116,119,121,122,124,128,129,130,132,133,134,136,138.

D. Analisis Reliabilitas Soal TOAFL IAIN Curup

Reliabilitas adalah tingkat ketetapan suatu instrumen mengukur apa yang harus diukur. Pengukuran dilakukan dengan teknik gasal genap, yaitu membelah soal menjadi 2 bagian. Soal nomor gasal (1,3,5,7,9,11,13,dst) sebagai bagian awal, serta nomor genap (2,4,6,8,10,12, dst) sebagai bagian akhir.

Adapun langka-langkah dalam mendapatkan nilai reliabilitas sebuah tes adalah sebagai berikut:

1. Menjumlahkan skor-skor yang dimiliki oleh butir soal yang bernomor gasal.

²⁵ Ibid., 182

2. Menjumlahkan skor-skor yang dimiliki oleh butir soal yang bernomor genap.
3. Menghitung angka indeks korelasi r *product moment* antara X (butir soal yang bernomor ganjil) dengan Y (butir soal yang bernomor Genap).

Berdasarkan analisis ditemukan data sebagai berikut:

$$\begin{array}{ll} N = 93 & \sum X^2 = 76733 \\ \sum X = 2571 & \sum Y^2 = 77537 \\ \sum Y = 2607 & \sum XY = 75951 \end{array}$$

4. Setelah menemukan hasil, maka disubstitusikan ke dalam rumus berikut:

$$\begin{aligned} r_{11/22} &= \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}} \\ &= \frac{93 \times 75.951 - (2571)(2607)}{\sqrt{[93 \times 76.733 - (2571)^2][93 \times 77.537 - 2607^2]}} \\ &= \frac{7.063.443 - 6.702.597}{\sqrt{[7.136.169 - 6.610.041][7.210.941 - 6.796.449]}} \\ &= \frac{360.846}{\sqrt{526.128 \times 414.492}} \\ &= \frac{360.846}{\sqrt{218.075.846.976}} \\ &= \frac{466985,92}{360846} \\ &= 0,773 \end{aligned}$$

5. Menghitung koefisien reliabilitas tes (r_{tt} atau r_{11}) dengan menggunakan rumus:

$$\begin{aligned} r_{11} &= \frac{2 r_{11/12}}{1 + r_{11/12}} \\ &= \frac{2 \times 0,773}{1 + 0,773} \\ &= \frac{1546}{1773} \\ &= 0,8719684151 \\ &= 0,87 \text{ (dibulatkan)} \end{aligned}$$

6. Memberikan interpretasi terhadap r_{11}

Berdasarkan hasil-hasil perhitungan di atas, telah diperoleh koefisien reliabilitas tes sebesar 0,87. Koefisien reliabilitas tes sebesar 0,87 ternyata jauh lebih besar dari 0,70. Dengan demikian maka hasil tes TOAFL telah dapat dinyatakan sebagai tes yang memiliki reliabilitas tinggi.

V. SIMPULAN DAN SARAN

Berdasarkan analisis yang telah peneliti lakukan, maka dapat ditarik kesimpulan sebagai berikut:

1. Perhitungan tingkat kesukaran soal adalah pengukuran seberapa besar derajat kesukaran suatu soal. Berdasarkan analisis yang telah dilakukan terhadap 140 butir soal TOAFL IAIN Curup, terdapat 15 soal yang masuk dalam kategori mudah, 71 soal masuk dalam kategori sedang, dan 54 soal masuk dalam kategori sukar. Dengan melihat persentase proporsi tingkat kesukaran soal, soal tersebut masuk dalam kategori baik, karena lebih 50% indeks kesukarannya memenuhi standar.
2. Daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang pandai (menguasai materi) dengan siswa yang kurang

pandai (kurang/tidak menguasai materi). Berdasarkan analisis daya pembeda soal TOAFL IAIN Curup yang berjumlah 140 butir soal, tidak terdapat soal yang masuk dalam kategori baik, 48 soal yang masuk dalam kategori sedang, 56 soal yang masuk dalam kategori kurang, 4 soal yang tidak memiliki daya pembeda, dan 32 soal yang negatif. Dengan melihat persentase proporsi daya pembeda soal, soal pilihan ganda tersebut masuk dalam kategori tidak baik, karena 0% soal yang masuk kategori baik.

3. Validitas item dari suatu tes yaitu ketepatan mengukur yang dimiliki oleh sebutir item dalam mengukur apa yang seharusnya diukur lewat butir item tersebut. Dari perhitungan hasil validitas soal bentuk pilihan ganda, menunjukkan bahwa ada 78 soal yang masuk dalam kategori valid dan 62 dalam kategori tidak valid. Persentase validitas soal TOAFL IAIN Curup tersebut menunjukkan lebih 50% soal memiliki kualitas valid.
4. Reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen, Reliabilitas tes berkenaan dengan pertanyaan apakah suatu tes teliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan. Dari hasil perhitungan reliabilitas soal, menunjukkan bahwa soal TES TOAFL IAIN Curup memiliki kategori reliabilitas yang tinggi yaitu 0,87.

Berdasarkan analisis yang dilakukan oleh peneliti terhadap soal TOAFL IAIN Curup, peneliti menyarankan hal-hal sebagai berikut :

1. Analisis butir soal TOAFL IAIN Curup bertujuan untuk memperoleh gambaran kualitas soal tes yang telah dilakukan, apabila kualitas soal baik, hal ini dapat memberikan hasil tes yang baik pula mengenai gambaran tentang prestasi mahasiswa IAIN Curup yang sebenarnya. Kualitas soal yang baik, juga menentukan kualitas dosen bahasa Arab IAIN Curup yang telah memberikan ilmu pengetahuan sebagai bekal mengikuti tes TOAFL ini. Dengan masih ditemukannya beberapa soal yang tingkat kesukaran soalnya masih belum memenuhi standar, daya pembeda soalnya kurang, bahkan tidak ada daya pembeda sama sekali, validitas soalnya masuk dalam kriteria tidak valid, maka soal-soal tersebut harus dibuang atau dilakukan revisi lagi dengan mengganti soal tersebut dengan soal yang baru. Agar nantinya soal tes ini benar-benar dapat mengukur kemampuan peserta tes dengan tepat.
2. Untuk meningkatkan hasil tes TOAFL, peserta tes sebaiknya lebih memahami pentingnya diadakan tes TOAFL, sehingga peserta tes dapat benar-benar mempersiapkan diri untuk mengikuti tes TOAFL.

DAFTAR PUSTAKA

- Arifin, Zaina. 2009. *Evaluasi Pembelajaran*, (Jakarta: Direktorat Jenderal Pendidikan Islam Departemen Agama RI
- Arikunto, Suharismi. *Prosedur Penelitian*, Jakarta: PT. Rineka Cipta
- Asrori, Imam dkk. *Evaluasi Pembelajaran Bahasa Arab*. Malang: Misykat
- Barmawi, Muhammad. 2011. *Lulus TOAFL dengan Mudah dan Memuaskan*, Yogyakarta: DIVA Press
- Daryanto. 2010. *Evaluasi Pendidikan*, Jakarta: Rineka Cipta
- Djiwandono, M. Soenardi Djiwandono. 2011. *Tes Bahasa dalam Pengajaran*. Bandung: ITB
- Hamalik, Oemar. *Pendidikan Guru Berdasarkan Pendekatan Kompetensi*, Jakarta: PT.Bumi Aksara
- Hayati, Mardia. 2009. *Desain Pembelajaran*, Pekanbaru: Yayasan Pustaka Riau
- Margono. 2010. *Metode Penelitian Pendidikan*, Jakarta: Rineka Cipta
- Matsna, Moh dan Mahyudin, Erta. 2012. *Pengembangan Evaluasi dan Tes Bahasa Arab*.(Tangerang Selatan: Alkitabiah
- Rubiyanto, Rubino dan Hartini, Sri. 2005. *Evaluasi Pendidikan*, Surakarta: Progeram Akta Mengajar FKIP UMS
- Sudijono, Anas. 1996. *Pengantar Evaluasi Pendidikan*, Jakarta: PT. Raja Grafindo Persada
- Sugiyono.
2008.*Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitati f dan R&D*, Bandung : Alfabeta
- Sulistiyorini.
2009. *Evaluasi Pendidikan Dalam Meningkatkan Mutu Pendidikan*, Yogyakarta: Teras
- Surahmad, Winarno. 1992. *Pengantar Penelitian Ilmiah Dasar Metode dan Teknik*, Bandung: Tarsito
- Wahab, Muhibb Abdul. 2010. *slide Presentasi Pembelajaran TOAFL di MAN 4 Jakarta*
- Wiraatmaja, Rochiati. 2005.
Metode Penelitian Tindakan Kelas, Bandung: Remaja Rosdakarya