

BAB III

METODOLOGI PENELITIAN

A. Objek Penelitian

Objek dalam penelitian ini adalah tes ulangan umum bersama matematika di Propinsi Jawa Timur tahun pelajaran 1999/2000, tes tersebut terdiri dari seperangkat tes ulangan umum matematika kelas I, seperangkat tes ulangan umum matematika kelas II, dan seperangkat tes ulangan umum matematika kelas III IPA.

Dari tiga perangkat tes itulah yang akan dicari kualitasnya dengan cara menganalisis hasil kerja siswa dalam menyelesaikan tes tersebut sebagai data. disamping itu juga akan dianalisis tentang jenis kesalahan-kesalahan yang dilakukan siswa dalam menyelesaikannya. Mengingat banyaknya MA yaitu 82 MAN dengan 38.464 siswa, dan 585 MAS dengan 63.270 siswa (Depag, 1999, h. 223), serta luas daerah Propinsi Jawa Timur yang terdiri Jawa Timur terdiri dari 7 daerah pembantu Gubernur dengan tiap daerah pembantu Gubernur terdiri dari beberapa kabupaten dan tiap Kabupaten terdiri dari beberapa MAN dan MAS, dan agar data dari hasil kerja siswa itu merupakan data yang representatif, maka pengambilan data tersebut penulis menggunakan metode "multi strategy sampling" dengan langkah-langkah sebagai berikut:

1. Pengambilan secara acak 3 daerah dari 7 daerah Pembantu Gubernur yang ada di Jawa Timur.
2. Dari hasil nomor 1, diambil lagi secara acak satu kabupaten atau kodya tiap daerah tersebut.

3. Dari kabupaten atau kodya yang terambil tersebut diambil secara acak satu Madrasah Aliyah Negeri dan satu Madrasah Aliyah Swasta.
4. Dari perolehan langkah ketiga semua lembar jawaban siswa yang telah terkumpul kemudian diberi nomor, selanjutnya diambil secara acak 50 persen dari semua lembar jawaban tersebut dengan bantuan komputer program Microsoft Excel.

Sehingga dari langkah-langkah pengambilan data tersebut di atas, maka didapatkanlah 3 MAN dan 3 MAS sebagai tempat pengambilan data dari penelitian ini yaitu MAN 2 Madiun, MAN 3 Malang, MAN Bangkalan, MA Tribakti Madiun, MA Muhammadiyah I Malang, dan MA Manba'ul Hikam Bangkalan, hal ini dapat digambarkan dalam tabel berikut ini.

Tabel 3.1
Nama Madrasah dan Banyaknya Lembar Jawaban Siswa Terpilih

| No | Nama Madrasah | Kls I | Kls II | Kls III |
|---|--------------------------------------|-------|--------|---------|
| 1 | MAN 2 Madiun | 91 | 68 | 83 |
| 2 | MAN 3 Malang | 77 | 76 | 68 |
| 3 | MAN Bangkalan | 83 | 73 | 33 |
| 4 | MA Muhammadiyah I Malang | 21 | 28 | - |
| 5 | MA Tribakti Madiun | 34 | 30 | - |
| 6 | MA Manba'ul Hikam Ketengan Bangkalan | 44 | 37 | - |
| Jumlah | | 350 | 312 | 184 |
| Pengambilan 50% secara acak (jangka ke-4) | | 175 | 156 | 92 |

B. Rancangan Penelitian

Sebagaimana disebutkan di atas tujuan penelitian ini adalah untuk mengetahui kualitas alat evaluasi yang digunakan dalam ulangan umum bersama matematika Madrasah Aliyah di Propinsi Jawa Timur tahun pelajaran 1999/2000, oleh sebab itu dalam penelitian ini tidak ada perlakuan atau manipulasi terhadap objek penelitian. Penelitian ini hanya mengungkapkan faktor-faktor berdasarkan keadaan butir-butir soal, kisi-kisi soal, lembar jawaban siswa dan GBPP matematika 1994 beserta Suplemennya.

Adapun analisis data yang akan digunakan dalam penelitian ini adalah sebagai berikut:

I. Prosedur untuk menganalisis kualitas soal secara kualitatif adalah:

a. Content Validity

Suatu tes mempunyai validitas isi yang baik atau tidak ini tergantung pada penilaian para ahli (*expert professional judgment*), tentunya para ahli disini dalam menentukan validitas isi tes yaitu dengan cara mencocokkan setiap butir tes dengan kisi-kisi, TIK, dan GBPP beserta Suplemennya.

Untuk keperluan ini penulis dalam menentukan apakah suatu butir soal memiliki validitas yang baik atau tidak yaitu dengan langkah-langkah berikut ini:

- (1) Membuat daftar pertanyaan dan atau tabel yang di dalamnya tercantum nomor urut, jenis penilaian butir soal, dan nomor soal seperti pada lampiran E.1.
- (2) Memeriksa dan mencocokkan setiap butir soal dengan GBPP, TIK, dan Indikator.

- (3) Menjawab setiap pertanyaan yang ada pada daftar pertanyaan yang sudah dibuat pada bagian (1).
- (4) Memeriksa kembali setiap jawaban yang sudah diberikan, selanjutnya merangkum dan kemudian memberikan kesimpulan.

b. Face Validity

Sebagai mana telah dikemukakan di bagian depan bahwa Face validity adalah keabsahan susunan kalimat atau kata-kata dalam menyusun butir soal sehingga jelas pengertiannya atau tidak menimbulkan tafsiran lain, dan bahkan dikatakan jika soal sulit dipahami, kurang bersih, tulisan terlalu berdesakan, tanda baca atau notasi lain yang kurang jelas, hal ini akan mengurangi validitas muka alat evaluasi itu.

Dengan demikian dalam menentukan validitas muka, penulis lakukan dengan langkah yang sama dengan menentukan validitas isi sebagaimana uraian di atas.

2. Prosedur untuk menganalisis kualitas soal secara kuantitatif adalah:

a. Reliabilitas Tes

Dalam menentukan reliabilitas tes pilihan ganda penulis menggunakan cara Kuder-Richardson 20 sebab data berbentuk dikhotomi dalam hal ini benar atau salah (Ruseffendi, 1991, h. 191), dan rumusnya adalah:

$$r = \frac{b}{b-1} \times \frac{DB^2 - \sum t.s}{DB^2}$$

dengan: r = koefisien korelasi

b = banyak soal

DB^2 = deviasi baku

t = rasio siswa yang menjawab soal tertentu betul

s = rasio siswa yang menjawab soal tertentu salah

Sedangkan tes uraiannya menggunakan rumus Alpha (Cronbach Alpha) sebab jawaban siswa beraneka ragam (Ruseffendi, 1991, h.193) dan rumusnya

$$\text{adalah: } r = \frac{b}{b-1} \times \frac{DB^2 - \sum DB_i^2}{DB_i^2}$$

dengan: r = koefisien korelasi

b = banyak soal

DB^2 = variansi skor seluruh soal menurut skor siswa perorangan

DB_i^2 = variansi skor soal ke- i

Klasifikasi besarnya koefisien reliabilitas menurut Guilford (dalam Ruseffendi, 1991, h. 197) adalah:

Tabel. 3.2

Klasifikasi Koefisien Reliabilitas

| Besarnya r | Tingkat Reliabilitas |
|--------------|----------------------|
| 0,00 - 0,20 | Kecil |
| 0,20 - 0,40 | Rendah |
| 0,40 - 0,70 | Sedang |
| 0,70 - 0,90 | Tinggi |
| 0,90 - 1,00 | Sangat Tinggi |

b. Daya Beda

Ruseffendi (1991. h.199) memberikan definisi Daya Pembeda atau Indeks Diskriminasi adalah korelasi antara skor jawaban sebuah butir soal dengan skor jawaban seluruh soal. Sehingga dengan demikian maka rumus yang digunakan untuk menentukan daya pembeda ini adalah menggunakan Korelasi Biserial Titik. Hal ini didasarkan pada pendapat Furqon (1997. h.98) mengatakan bahwa korelasi biserial titik (point biserial) merupakan salah satu bentuk korelasi dari Pearson yang digunakan dalam situasi khusus, yaitu untuk mengkorelasikan satu peubah prediktor yang bersifat dikhotomi (biner atau binomial) dengan satu peubah kriteria yang berskala interval atau rasio.

Selanjutnya Furqon (1997. h. 98) menyebutkan formula korelasi biserial titik sebagai berikut:

$$r_{pbis} = \frac{\bar{Y}_p - \bar{Y}_t}{S_t} \sqrt{\frac{p}{q}}$$

dengan: r_{pbis} = koefisien korelasi biserial titik,

\bar{Y}_p = rata-rata kelompok p (kelompok kesatu),

\bar{Y}_t = rata-rata seluruh subjek,

S_t = simpangan baku untuk seluruh subjek,

p = proposisi subjek kelompok ke satu, dan

q = proposisi subjek kelompok kedua

Adapun klasifikasi daya pembeda butir soal didasarkan pada pendapat Ebel and Frisbie(dalam Wiersma, h. 147) sebagai berikut:

Tabel 3.3
Klasifikasi Daya Pembeda

| Daya Pembeda | Evalasi Butir Soal |
|----------------|-------------------------------------|
| 0,40 dan lebih | Sangat baik |
| 0,30 - 0,39 | Cukup baik, mungkin perlu perbaikan |
| 0,20 - 0,29 | Batas minimum, perlu perbaikan |
| 0,19 ke bawah | Buruk, dirombak atau dihilangkan |

Sedangkan untuk menentukan daya pembeda dari soal bentuk uraian penulis menggunakan pendapat Karno To (1996, h. 15) yaitu dengan

menggunakan rumus: $DP = \frac{S_A - S_B}{I_A} \times 100\%$

Keterangan: DP = Indeks daya pembeda

S_A = Jumlah skor kelompok atas

S_B = Jumlah skor kelompok bawah

I_A = Jumlah skor ideal salah satu kelompok (atas/bawah)

Serta dengan kriteria daya pembeda sebagai berikut:

Negatif – 9% = Sangat buruk, harus dibuang

10% - 19% = buruk, sebaiknya dibuang

20% - 29% = Agak baik, mungkin perlu direvisi

30% - 49% = Baik

50% ke atas = Sangat baik

c. Indek Kesukaran

Setelah perangkat tes dijudgment, kemudian hasil tes dianalisis tingkat kesukarannya apakah sudah sesuai atau tidak dengan hasil judgment itu. Untuk itu tingkat kesukaran butir bentuk pilihan ganda ditentukan oleh perbandingan antara banyaknya siswa yang menjawab soal itu benar dengan banyaknya siswa yang menjawab butir soal itu. Menurut Sudjana (1989, h. 137) rumus yang dipakai adalah:

$$I = \frac{B}{N}$$

Dengan: I = Indeks kesukaran untuk tiap butir soal.

B = banyaknya siswa yang menjawab benar setiap butir soal.

N = banyaknya siswa yang memberikan jawaban pada soal tersebut.

Adapun untuk soal bentuk uraian juga menggunakan rumus yang sama dengan menterjemahkan B sebagai jumlah skor perolehan pada tiap butir soal dan N sebagai jumlah skor ideal pada tiap butir soal yang diolah.

Selanjutnya kriteria indeks kesukaran soal itu menurut Sudjana (1998, h. 137) adalah sebagai berikut pada tabel 3.4.

Tabel 3.4
Kriteria Indeks Kesukaran

| Indeks Kesukaran | Kategori |
|------------------|----------|
| 0,00 - 0,30 | Sukar |
| 0,31 - 0,70 | Sedang |
| 0,71 - 1,00 | Mudah |

d. Analisis Distraktor

Untuk melihat apakah option berfungsi dengan baik atau tidak, dalam penelitian ini penulis mengambil gabungan dari dua pendapat yaitu:

1. Nurkencana (1986, h. 141) menyatakan “option kunci berfungsi dengan baik bila jumlah pemilih kelompok atas dan kelompok bawah tidak kurang dari 25% tetapi tidak lebih dari 75%, dan frekuensi pilihan kelompok atas harus lebih tinggi dari frekuensi pilihan kelompok bawah”. Dan dari dua keadaan atau persyaratan tersebut haruslah keduanya dipenuhi sebagai kriteria option kunci yang baik. Untuk lebih jelasnya perhatikan contoh berikut ini, dimana jumlah siswa kelompok atas maupun kelompok bawah sama yaitu 48 siswa.

Tabel 3.5
Contoh Option Kunci

| Kelompok | Nomor Soal | | |
|---------------------|------------|----|----|
| | 7 | 15 | 25 |
| Kelompok Atas (KA) | 12 | 12 | 38 |
| Kelompok Bawah (KB) | 9 | 29 | 13 |

Analisisnya yaitu:

- a. Persentase jumlah kelompok atas dan kelompok bawah untuk soal nomor (7) 21.9 %, (15) 42.7 %, dan (25) 53 % dengan demikian soal nomor 15 dan nomor 25 memenuhi syarat yang pertama.

- b. Jumlah kelompok atas dan bawah yaitu (7) $KA > KB$, (15) $KA < KB$, sedangkan (25) $KA > KB$, dengan demikian soal nomor 7 dan 25 memenuhi ketentuan yang kedua
- c. Kesimpulan yang akan diambil berdasarkan dari hasil (a) dan (b) menunjukkan bahwa soal nomor 25 mempunyai option kunci yang berfungsi dengan baik.
2. Arikunto (1999, h. 220) mengatakan bahwa suatu distraktor dapat dikatakan berfungsi dengan baik jika paling sedikit dipilih oleh 5% pengikut tes.
- e. Validitas Banding

Validitas banding adalah juga merupakan salah satu syarat yang harus dipenuhi oleh perangkat tes ulangan umum bersama MA di Propinsi Jawa Timur, untuk itu dalam menentukan validitas banding disini penulis membandingkan skor perolehan siswa (X) dengan nilai rata-rata ulangan harian (Y_1) dan juga dengan nilai rapor cawu 2 (Y_2) untuk kelas I dan II, sedangkan untuk kelas III pembandingnya adalah nilai EBTANAS murni (NEM) yang diperolehnya.

Untuk menentukan validitas banding ini penulis menggunakan korelasi produk momen dari Pearson sebagaimana dituliskan oleh Furqon (1997, h. 94) sebagi berikut:

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

Adapun interpretasi dari koefisien korelasinya sebagai mana disebutkan oleh Guilford (dalam Suherman,1990, h.147) adalah sebagai berikut:

1. Apabila $r_{xy} < 0.20$ korelasi sangat rendah.
2. Apabila $0.20 \leq r_{xy} < 0.40$ korelasi rendah.
3. Apabila $0.40 \leq r_{xy} < 0.70$ korelasi sedang.
4. Apabila $0.70 \leq r_{xy} < 0.90$ korelasi tinggi
5. Apabila $0.90 \leq r_{xy} \leq 1.00$ korelasi sangat tinggi

C. Prosedur untuk menganalisis kesalahan-kesalahan siswa

Langkah-langkah dalam menganalisis kesalahan siswa adalah sebagai berikut:

1. Mengumpulkan data atau hasil kerja siswa dari siswa MA terpilih.
2. Membuat tabel untuk mengidentifikasi jenis kesalahan siswa dari jawaban soal uraian yang telah terkumpul.
3. Merumuskan bentuk kesalahan siswa (Apakah kesalahan konsep atau kesalahan operasi?).

Untuk lebih jelasnya berikut contoh jenis kesalahan-kesalahan siswa dalam mengerjakan soal matematika. Tentukan turunan pertama dari $y = 3^{3 \cos x}$

Misalnya, beberapa jawaban siswa sebagai berikut:

$$\begin{array}{ll}
 1. \ y' = 3^{3 \cos x} (-3 \sin x) & 3. \ y' = 3 \cos x \cdot 3 \cdot (-\sin x) \\
 = -3 \cdot 3^{3 \cos x} \cdot \sin x & = -9 \cos x \cdot \sin x \\
 \\
 2. \ y' = 3^{3 \cos x} \cdot -3 \sin x \cdot \ln 3 & \\
 = -9^{3 \cos x} \sin x \cdot \ln 3 &
 \end{array}$$

Untuk jawaban yang pertama dan ketiga diduga siswa tersebut melakukan kesalahan tentang konsep, sebab siswa tersebut tidak bisa mengingat tentang konsep turunan pertama dari $y = a^u$ dengan baik. Sedangkan jawaban siswa yang ke dua diduga ia telah melakukan kesalahan operasi hitung sehingga ia mengalikan $3^{3\cos x} \cdot (-3) = -9^{3\cos x}$ yang seharusnya $(-3)3^{3\cos x}$

D. Prosedur Pengambilan Data

Untuk mendapatkan data dalam penelitian ini dilakukan langkah-langkah sebagai berikut:

1. Pengurusan surat kepada Program Pascasarjana UPI Bandung untuk mengadakan penelitian atau mendapatkan data yang ditujukan kepada MA terpilih dan Depag setempat.
2. Menyampaikan surat dari PPS UPI Bandung kepada Kepala MA terpilih dan Depag setempat, kemudian setelah mendapatkan izin maka barulah mengambil data dari hasil ulangan umum bersama matematika cawu III di MA tersebut.
3. Setelah pengambilan data, maka penulis akan meminta surat dari kepala MA terpilih, hal ini sebagai bukti peneliti telah melakukan pengambilan data di MA tersebut.

E. Metode dan Instrumen Penelitian

Metode penelitian adalah cara yang digunakan oleh peneliti dalam mengumpulkan data penelitiannya diantaranya adalah angket, wawancara, pengamatan atau tes, observasi, dan dokumentasi (Arikunto. 1997. h. 151).

Dengan demikian dalam penelitian ini metode yang digunakan peneliti untuk mendapatkan data mengenai kualitas soal ulangan umum bersama matematika Madrasah Aliyah di Jawa Timur adalah metode dokumentasi.

Dokumen yang digunakan dalam penelitian ini terdiri dari kisi-kisi soal yang disertai soal ulangan umum bersama matematika MA di Jatim untuk kelas I, II, dan III IPA tahun pelajaran 1999/2000, kunci jawaban soal pilihan ganda, penyelesaian soal uraiannya serta aturan penskorannya, nilai rapor cawu 2 dan nilai rata-rata ulangan harian untuk kelas I dan II, daftar nilai NEM untuk kelas III, GBPP matematika SMU berdasarkan kurikulum 1994 beserta suplemennya, 350 lembar jawaban siswa kelas I, 312 lembar jawaban siswa kelas II, dan 184 lembar jawaban siswa kelas III IPA.

Sedangkan instrumen penelitian adalah alat atau fasilitas yang digunakan oleh peneliti dalam mengumpulkan data agar pekerjaan lebih mudah dan hasilnya lebih baik, dalam arti cermat, lengkap, dan sistematis sehingga lebih mudah diolah, misalnya angket, ceklis, pedoman wawancara, dan pedoman pengamatan (Arikunto, 1997, h. 151). Adapun instrumen dalam penelitian ini adalah berupa angket atau daftar isian yang berisikan penilaian tentang validitas isi dan validitas muka, tingkat kesukaran dari soal ulangan umum bersama matematika MA di Jawa Timur tahun pelajaran 1999/2000 oleh penilai.