

BAB III

METODE PENELITIAN

A. Desain Penelitian

Penelitian ini merupakan suatu eksperimen terhadap siswa kelas VIII MTs. Karakteristik yang akan diukur dalam penelitian ini adalah pemecahan masalah matematik. Pengukuran kemampuan pemecahan masalah matematik ini dilakukan terhadap kelompok siswa yang diberi perlakuan (eksperimen) dan kelompok siswa sebagai pembanding atau kontrol.

Desain yang digunakan dalam penelitian ini adalah desain kelompok kontrol pretes – postes (*Pretest – Posttest Control Group Design*). Desain ini melibatkan dua kelompok yaitu kelompok eksperimen yang memperoleh perlakuan (*treatment*) pembelajaran dengan strategi heuristik dalam kelompok kecil (*X*), dan kelompok kontrol yang mendapatkan pembelajaran secara konvensional. Pada desain ini terjadi pengelompokan subjek secara acak (*R*), dan adanya pretes dan postes (*O*). Dengan demikian, model desain grup kontrol pretes – postes, rumus yang digunakan sebagaimana dalam Fraenkel (1990) yaitu;

$$\begin{array}{cccc} R & O & X & O \\ R & O & & O \end{array}$$

R : Pengelompokan subjek dilakukan secara acak (*random*)

O : Observasi yang terdiri dari pretes dan postes

X : Model pembelajaran dengan strategi heuristik dalam kelompok kecil

B. Populasi dan Sampel Penelitian

Populasi dalam penelitian ini adalah seluruh siswa kelas VIII MTs Negeri Karangampel, Kabupaten Indramayu. Alasan dipilihnya kelas VIII adalah mengingat: a) siswa kelas VIII merupakan siswa menengah pada satuan pendidikan tersebut yang diperkirakan telah dapat menyesuaikan diri dengan kondisi lingkungannya; b) terdapat topik matematika yang dianggap tepat yaitu materi pokok bahasan lingkaran mengingat masih banyak siswa yang kurang memahami cara melukis garis singgung lingkaran dan aplikasinya; dan c) siswa kelas VIII telah menerima cukup banyak materi prasyarat untuk mengikuti topik matematika yang akan diteliti.

Adapun sampel penelitian diambil dua kelas dari sembilan kelas secara random (diundi) agar setiap kelas dari seluruh populasi mendapat kesempatan yang sama untuk dipilih. Kedua kelas yang terambil adalah VIII H dan VIII I. Kemudian kedua kelas tersebut diundi lagi untuk menentukan kelas eksperimen dan kelas kontrol. Hasilnya, kelas VIII H dengan jumlah siswa 46 orang sebagai kelas eksperimen dan kelas VIII I dengan jumlah siswa 46 orang sebagai kelas kontrol.

Dengan demikian, ukuran sampel di kedua kelas telah memenuhi syarat seperti yang dikatakan Gay (dalam Ruseffendi, 1998a: 92) bahwa ukuran sampel minimum untuk riset percobaan (eksperimen) paling sedikit 30 orang per kelompok dan jika penelitian eksperimen yang dikontrol dengan ketat mungkin 15 orang cukup.

C. Variabel Penelitian

Penelitian ini melibatkan dua variabel yaitu variabel bebas dan variabel terikat. Variabel bebasnya adalah pembelajaran dengan strategi heuristik dalam kelompok kecil dan variabel terikatnya adalah kemampuan pemecahan masalah matematik.

D. Instrumen Penelitian dan Pengembangannya

Penelitian ini menggunakan dua macam instrumen, yaitu: tes uraian, untuk mengukur kemampuan pemecahan masalah matematika; dan angket dengan tujuan mengetahui sikap siswa terhadap pembelajaran dengan strategi heuristik.

1. Tes Kemampuan Pemecahan Masalah Matematik

Tes kemampuan pemecahan masalah ini disusun dalam bentuk uraian. Pokok bahasan yang dipilih sebagai materi uji adalah Lingkaran, dengan pertimbangan banyak memuat masalah dalam kehidupan sehari-hari.

Pengembangan instrumen dimulai dengan menyusun kisi-kisi, dan dilanjutkan dengan menyusun butir soal tes yang sesuai. Instrumen yang telah disusun, sebelum digunakan terlebih dahulu dianalisis isinya. Aspek yang dipertimbangkan meliputi kesesuaian kisi-kisi dengan butir soal, aspek bahasa dan materi matematika. Alasan pemilihan bentuk soal uraian, karena disesuaikan dengan maksud penelitian ini yaitu mengutamakan proses daripada hasil. Jenis tes seperti ini tidak memberikan kesempatan untuk berspekulasi, tetapi memberikan keleluasan kepada siswa untuk mengutarakan jawabannya sesuai dengan kemampuannya, sehingga peneliti dapat mengungkapkan lebih banyak variasi jawaban yang dikemukakan siswa.

Sebelum digunakan dalam penelitian, soal tersebut telah dipertimbangkan oleh dosen pembimbing. Banyaknya soal tes kemampuan pemecahan masalah matematika sepuluh butir soal uraian dengan pokok bahasan lingkaran. Setelah validitas isi terpenuhi, baru kemudian diujicobakan ke sekolah lain yang mempunyai karakteristik berbeda dengan sekolah tempat penelitian. Uji coba dilaksanakan pada tanggal 1 Agustus 2005 terhadap siswa SMA kelas X 4 di SMA Negeri 1 Krangkeng Kabupaten Indramayu. Uji coba ini dilakukan untuk menganalisis validitas butir soal, reliabilitas, daya beda dan tingkat kesukaran. Alasan uji coba di SMA karena siswa kelas X (kelas satu) sudah mendapatkan materi lingkaran, sedangkan siswa kelas III SMP/MTs belum seluruhnya mendapatkan materi lingkaran. Pokok bahasan lingkaran pada Kurikulum 1994 dibagi dua yaitu diberikan pada kelas II dan kelas III sedangkan pada Kurikulum 2004 pokok bahasan lingkaran disatukan pada kelas VIII (kelas dua).

Untuk mengukur kemampuan siswa dalam memberi skor atas jawaban siswa pada soal uraian tersebut digunakan aturan penskoran yang dikemukakan oleh Sumarmo (1994). Cara penskoran keempat aspek dalam pemecahan masalah matematika tersebut ditunjukkan pada Tabel 3.1.

Tabel 3.1

Pedoman Penskoran Kemampuan Pemecahan Masalah

Skor	Memahami Masalah	Membuat Rencana Penyelesaian	Melaksanakan Perhitungan	Memeriksa Kembali
0	Salah menginterpretasikan / salah sama sekali	Tidak ada rencana, membuat rencana yang tidak relevan	Tidak melakukan perhitungan	Tidak ada pemeriksaan atau tidak ada keterampilan lain
1	Salah menginterpretasikan sebagian soal, mengabaikan kondisi soal	Membuat rencana pemecahan yang tidak dapat dilaksanakan, sehingga tidak dapat dilaksanakan	Melaksanakan prosedur yang benar mungkin menghasilkan jawaban benar tetapi salah perhitungan	Ada pemeriksaan tetapi tidak tuntas
2	Memahami masalah soal selengkapanya	Membuat rencana yang benar tetapi salah dalam hasil / tidak ada hasil	Melakukan proses yang benar dan mendapat hasil yang benar	Pemeriksaan dilaksanakan untuk melihat kebenaran proses
3	-	Membuat rencana yang benar, tetapi belum lengkap	-	-
4	-	Membuat rencana sesuai dengan prosedur dan mengarah pada solusi yang benar	-	-
	Skor Maksimal 2	Skor Maksimal 4	Skor Maksimal 2	Skor Maksimal 2

2. Analisis Instrumen

Untuk menganalisis reliabilitas tes, validitas butir, tingkat kesukaran, dan daya pembeda soal hasil ujicoba instrumen digunakan program komputer.

Rumus-rumus yang digunakan pada program komputer untuk perhitungan

reliabilitas tes, validitas butir, tingkat kesukaran, dan daya pembeda soal didasarkan pada rumus-rumus sebagai berikut:

a. Analisis Reliabilitas

Suatu alat ukur memiliki reliabilitas yang baik bila alat ukur itu memiliki konsistensi yang handal walaupun dikerjakan oleh siapapun (dalam level yang sama), di manapun dan kapanpun dipakai. Reliabilitas soal bentuk uraian menggunakan rumus *Alpha-Cronbach*:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum s_i^2}{s_t^2} \right)$$

Keterangan:

r_{11} = reliabilitas yang dicari

n = banyak butir soal

$\sum s_i^2$ = jumlah varians skor tiap-tiap item

s_t^2 = varians skor total

Untuk interpretasi nilai r_{11} digunakan *Guilford* dalam Suherman dan Kusumah (1990: 177) adalah sebagai berikut:

$r_{11} \leq 0,20$ derajat reliabilitas sangat rendah

$0,20 < r_{11} \leq 0,40$ derajat reliabilitas rendah

$0,40 < r_{11} \leq 0,60$ derajat reliabilitas sedang

$0,60 < r_{11} \leq 0,80$ derajat reliabilitas tinggi

$0,80 < r_{11} \leq 1,00$ derajat reliabilitas sangat tinggi.

Untuk lebih menyakinkan, nilai r_{11} juga dikonsultasikan pada tabel r product moment, dengan mengambil taraf signifikan 5%.

Jika $r_{11} < r_{\text{tabel}}$, maka instrumen *tidak reliabel*

Jika $r_{11} > r_{\text{tabel}}$, maka instrumen *reliabel*

Untuk r_{11} negatif, berapapun nilainya, menunjukkan bahwa instrumen tidak reliabel (Suharsimi-Arikunto, 2003 : 86).

Analisis reliabilitas dilakukan terhadap instrumen untuk mengukur kemampuan pemecahan masalah matematik. Untuk mempermudah, proses perhitungan dilakukan menggunakan SPSS, (print out software terlampir). Dari hasil analisis reliabilitas diperoleh bahwa, nilai α sebesar 0,94, nilai koefisien reliabilitas ini termasuk kategori sangat tinggi. Hal ini menunjukkan bahwa tes yang digunakan memiliki reliabilitas yang sangat tinggi sebagai instrumen untuk mengukur kemampuan pemecahan masalah matematik siswa.

b. Analisis Validitas Tes

Dalam perhitungan tingkat validitas butir soal ini, mengacu pada Suherman dan Kusumah. (1990:154), yaitu menggunakan korelasi *product moment Pearson*, dengan mengkorelasikan skor yang didapat siswa pada suatu butir soal dengan skor total yang didapatnya. Rumus yang digunakan:

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

r_{XY} = koefisien korelasi antara variabel X dan variabel Y .

N = banyak siswa

Untuk interpretasi dari koefisien korelasi tersebut dibagi ke dalam kategori-kategori yang digunakan *Guilford, J.P.* dalam Suherman dan Kusumah.

(1990:147) seperti berikut:

$0,80 < r_{XY} \leq 1,00$ korelasi sangat tinggi,

$0,60 < r_{XY} \leq 0,80$ korelasi tinggi,

$0,40 < r_{XY} \leq 0,60$ korelasi sedang,

$0,20 < r_{XY} \leq 0,40$ korelasi rendah, dan

$r_{XY} \leq 0,20$ korelasi sangat rendah.

Untuk nilai r_{XY} diartikan sebagai koefisien validitas, sehingga kriteriumnya menjadi:

$0,80 < r_{XY} \leq 1,00$ validitas sangat tinggi (sangat baik),

$0,60 < r_{XY} \leq 0,80$ validitas tinggi (baik),

$0,40 < r_{XY} \leq 0,60$ validitas sedang (cukup),

$0,20 < r_{XY} \leq 0,40$ validitas rendah (kurang),

$0,00 < r_{XY} \leq 0,20$ validitas sangat rendah, dan

$r_{XY} \leq 0,00$ tidak valid.

Dalam hal ini, juga ditampilkan penafsiran harga koefisien korelasi r dengan membandingkan pada tabel harga kritik r product moment, dengan mengambil taraf signifikan 5%, sehingga didapat kemungkinan interpretasi :

$r_{hit} < r_{kritik}$, maka korelasi tidak signifikan

$r_{hit} > r_{kritik}$, maka korelasi signifikan

c. Analisis Daya Pembeda

Daya pembeda sebuah soal menunjukkan kemampuan soal tersebut membedakan antara siswa yang pandai dengan siswa yang kurang. Sebuah soal dikatakan memiliki daya pembeda yang baik bila memang siswa yang pandai dapat mengerjakan dengan baik. Karena jumlah responden tidak mencapai 100, maka pembagian kelompok pandai dengan kelompok kurang dilakukan dengan cara membagi dua sama banyak pada kedua kelompok. Jadi pembagiannya 50% kelompok pandai dan 50% kelompok kurang. Maka untuk perhitungan daya pembeda, dilakukan langkah-langkah ,menurut Suharsimi-Arikunto. (2003 : 212), sebagai berikut:

$$\text{Untuk soal uraian : } DP = \frac{S_A - S_B}{\frac{1}{2} n \cdot maks}$$

Keterangan:

S_A = jumlah skor kelompok atas pada butir soal yang diolah

S_B = jumlah skor kelompok bawah pada butir soal yang diolah

n = Jumlah seluruh siswa kelompok atas dan bawah

maks = skor maksimal tiap butir soal

Untuk interpretasi nilai daya pembeda menurut Suherman dan Kusumah. (1990 : 202) sebagai berikut:

$$DP \leq 0,00 \text{ sangat jelek}$$

$$0,00 < DP \leq 0,20 \text{ jelek}$$

$$0,20 < DP \leq 0,40 \text{ cukup}$$

$$0,40 < DP \leq 0,70 \text{ baik}$$

$$0,70 < DP \leq 1,00 \text{ sangat baik}$$

d. Analisis Tingkat Kesukaran

Tingkat kesukaran (TK) pada masing-masing butir soal, cukup dihitung dengan menggunakan rumus:

$$\text{Soal Uraian } TK = \frac{S_A + S_B}{n \cdot maks}$$

Keterangan:

TK = Tingkat kesukaran

S_A = jumlah skor yang didapat siswa pada butir soal kelompok atas

S_B = jumlah skor yang didapat siswa pada butir soal kelompok bawah

n = jumlah seluruh siswa kelompok atas dan bawah

maks = skor maksimal tiap butir soal

Sementara kriteria interpretasi tingkat kesukaran digunakan pendapat Suherman dan Kusumah.(1990 : 213):

$TK = 0,00$ soal terlalu sukar

$0,00 < TK \leq 0,30$ soal sukar

$0,30 < TK \leq 0,70$ soal sedang

$0,70 < TK < 1,00$ soal muda

$TK = 1,00$ soal terlalu mudah

Untuk menghitung nilai koefisien validitas, daya pembeda dan tingkat kesukaran digunakan *Microsoft Excel* (print out terlampir). Berdasarkan hasil perhitungan, nilai validitas butir tes kemampuan pemecahan masalah matematika termasuk kualifikasi tinggi dan sangat tinggi (di atas 0,60), sehingga semua item soal pada instrumen dapat dikatakan memenuhi kualifikasi valid (sahih).

Secara umum hasil analisis butir menunjukkan bahwa item soal untuk mengukur kemampuan pemecahan masalah matematik memiliki kualifikasi baik sebagai instrumen. Rekapitulasi hasil perhitungan untuk analisis butir tes kemampuan pemecahan masalah matematika disajikan berikut:

Tabel 3.2

Rekapitulasi Analisis Butir Soal Kemampuan Pemecahan Masalah Matematik

Item	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
SA	141	141	143	141	145	144	145	144	143	142
SB	92	92	84	92	93	94	98	100	99	97
I	190	190	190	190	190	190	190	190	190	190
DP	0,26	0,26	0,31	0,26	0,27	0,26	0,25	0,23	0,23	0,24
TK	0,61	0,61	0,60	0,61	0,63	0,63	0,64	0,64	0,64	0,63
r_{xy}	0,74	0,76	0,91	0,83	0,82	0,85	0,95	0,79	0,75	0,78
Kesimp	Dipakai									

Berdasarkan tabel-tabel di atas, terlihat bahwa koefisien korelasi untuk setiap item soal pada setiap aspek kemampuan matematika dalam tes bernilai $r_{xy} \geq 0,60$. Hal ini menunjukkan bahwa semua butir tes termasuk ke dalam kategori tinggi atau sangat tinggi. Dengan demikian semua butir tes valid. Seluruh indeks daya pembeda untuk setiap butir tes memiliki $DP \geq 0,2$. Hal ini menunjukkan bahwa butir-butir tes tersebut memiliki daya pembeda yang cukup. Tingkat kesukaran soal pada semua butir tes termasuk kategori sedang.

Berdasarkan analisis terhadap butir tes ini secara umum dapat disimpulkan bahwa semua butir pada tes tersebut dapat digunakan sebagai instrumen untuk mengukur kemampuan pemecahan masalah matematik siswa.

E. Prosedur Pelaksanaan Penelitian

Data dikumpulkan melalui beberapa cara, yaitu tes tertulis, dan skala sikap. Tes diberikan kepada kedua kelompok siswa (kelas eksperimen dan kelas kontrol) berupa pretes dan postes. Skala sikap diberikan kepada kelompok eksperimen mengenai sikap terhadap pembelajaran dengan strategi heuristik dalam kelompok kecil.

Penelitian diawali dengan uji coba tes matematika terhadap siswa kelas X 4 SMA Negeri 1 Krangkeng Kabupaten Indramayu. Hasil analisis terhadap tes matematika itu diperoleh kesimpulan sebagaimana yang telah diuraikan pada bagian pengembangan instrumen bahwa tes tersebut layak sebagai instrumen untuk mengukur kemampuan matematik. Oleh karena itu, untuk selanjutnya dilakukan pretes terhadap siswa kelas VIII yang terambil sebagai sampel dalam penelitian. Kemudian secara acak pula ditentukan kelas mana yang masuk dalam kelas eksperimen dan kelas mana yang masuk dalam kelas kontrol. Hasil pengacakan diperoleh bahwa, siswa yang termasuk dalam kelompok eksperimen adalah kelas VIII H, yaitu yang akan diberikan pembelajaran dengan strategi heuristik dalam kelompok kecil dan siswa yang termasuk dalam kelompok kontrol adalah siswa kelas VIII I, yaitu yang akan diberikan pembelajaran secara konvensional. Pembelajaran direncanakan sebanyak 8 pertemuan. Setelah pembelajaran berakhir, siswa akan diberikan postes dengan tes yang sama seperti pada pretes.

Tahap selanjutnya adalah tahap akhir pengumpulan data dari kelompok eksperimen, yaitu pemberian perangkat skala sikap terhadap kelas yang dijadikan eksperimen. Siswa pada kelompok ini diberikan perangkat skala sikap untuk

melihat sikap mereka terhadap pembelajaran dengan strategi heuristik dalam kelompok kecil. Analisis terhadap hasil skala sikap ini dilakukan secara deskripsi.

Adapun jadwal pelaksanaan penelitian adalah sebagaimana tampak dalam Tabel 3.3 di bawah ini:

Tabel 3.3
Jadwal Penelitian Lapangan

No	Waktu	Kegiatan
1.	1 Agustus 2005	Uji coba tes pemecahan masalah matematik
2.	6 Agustus 2005	Tes awal Pemecahan masalah matematik (kelas eksperimen dan kontrol)
3.	9 Agustus 2005	Pembelajaran I (kelas eksperimen dan kontrol)
4.	13 Agustus 2005	Pembelajaran II (kelas eksperimen dan kontrol)
5.	16 Agustus 2005	Pembelajaran III (kelas eksperimen dan kontrol)
6.	20 Agustus 2005	Pembelajaran IV (kelas eksperimen dan kontrol)
7.	23 Agustus 2005	Pembelajaran V (kelas eksperimen dan kontrol)
8.	27 Agustus 2005	Pembelajaran VI (kelas eksperimen dan kontrol)
9.	30 Agustus 2005	Pembelajaran VII (kelas eksperimen dan kontrol)
10.	3 September 2005	Pembelajaran VIII (kelas eksperimen dan kontrol)
11.	6 September 2005	Tes akhir pemecahan masalah matematik (kelas eksperimen dan kontrol)
12.	10 September 2005	Pengukuran Skala Sikap (kelas eksperimen)

F. Analisis Data

Setelah data terkumpul, tahap selanjutnya adalah menganalisis data dan interpretasi hasil. Dalam operasionalnya akan digunakan software SPSS 12.0 untuk mengolah data. Tahap analisis data meliputi:

1. Menguji normalitas

Uji normalitas diperlukan untuk menguji apakah sebaran data berdistribusi normal atau tidak. Apabila hasil pengujian menunjukkan bahwa sebaran data berdistribusi normal maka dalam menguji kesamaan dua rata-rata digunakan uji t . Apabila hasil pengujian menunjukkan bahwa sebaran data tidak berdistribusi normal maka untuk menguji kesamaan dua rata-rata digunakan statistik nonparametrik, yaitu uji U Mann-Whitney. Uji normalitas ini dilakukan terhadap skor pretes, postes, dan gain kemampuan pemecahan masalah matematik dengan strategi heuristik dari dua kelompok siswa (kelas eksperimen dan kontrol).

Untuk menguji normalitas digunakan uji Chi-kuadrat, dengan hipotesis sebagai berikut:

H_o : sebaran data mengikuti distribusi normal

H_A : sebaran data tidak mengikuti distribusi normal

Statistik uji Chi-kuadrat yang digunakan sebagaimana dalam Nugraha (1998):

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

χ^2 : Nilai statistik chi-kuadrat

O_i : Nilai skor observasi ke- i

E_i : Nilai skor yang diharapkan ke- i

Kriteria penolakan H_o :

Hipotesis nol ditolak jika $\chi^2_{hitung} > \chi^2_{(0,05, k-1)}$

Hipotesis nol diterima jika $\chi^2_{hitung} \leq \chi^2_{(0,05, k-1)}$,

2. Menguji homogenitas varians dari kedua kelompok

Uji homogenitas varians digunakan untuk menguji kesamaan varians dari skor pretes, postes, dan gain kemampuan pemecahan masalah matematika pada kedua kelompok (kelompok kontrol dan kelompok eksperimen). Apabila hasil pengujian menunjukkan kesamaan varians maka untuk uji kesamaan dua rata-rata digunakan uji t (apabila berdistribusi normal) dan digunakan varians gabungan. Apabila hasil pengujian menunjukkan tidak homogen maka untuk uji kesamaan dua rata-rata digunakan uji t (apabila berdistribusi normal) dan tidak digunakan varians gabungan. Operasionalnya pengujian ini akan dilakukan dengan menggunakan software SPSS. Adapun hipotesis statistik yang digunakan adalah:

$$\text{Hipotesis : } H_o : \sigma_t^2 = \sigma_c^2$$

(variens kelompok eksperimen sama dengan varians kelompok kontrol)

$$H_A : \sigma_t^2 \neq \sigma_c^2$$

(variens kelompok eksperimen tidak sama dengan varians kelompok kontrol)

Untuk menguji hipotesis tersebut digunakan rumus statistik uji F sebagai berikut:

$$F = \frac{S_t^2}{S_c^2}$$

S_t^2 : varians kelompok eksperimen

S_c^2 : varians kelompok kontrol

Kriteria uji homogenitas adalah:

Hipotesis nol ditolak jika $F_{hitung} > F_{daftar}$

Hipotesis nol diterima jika $F_{hitung} \leq F_{daftar}$

3. Menghitung Gain Ternormalisasi

Menyatakan gain dalam hasil proses pembelajaran tidaklah mudah. Mana yang sebenarnya dikatakan gain tinggi dan mana yang dikatakan gain rendah, kurang dapat dijelaskan melalui gain absolut (selisih antara skor postes dengan pretes). Misalnya, siswa yang memiliki gain 2 dari 3 ke 5 dan siswa yang memiliki gain 2 dari 5 ke 7 dari suatu soal dengan skor maksimal 7. Gain absolut menyatakan bahwa kedua siswa memiliki gain yang sama. Secara logis seharusnya siswa yang kedua memiliki gain yang lebih tinggi dari siswa yang pertama. Hal ini karena usaha untuk meningkatkan dari 5 ke 7 (yang juga 7 merupakan skor maksimal) akan lebih berat daripada meningkatkan dari 3 ke 5.

Menyikapi kondisi bahwa siswa yang memiliki gain absolut sama belum tentu memiliki gain hasil belajar yang sama, Meltzer (2002) mengembangkan sebuah alternatif untuk menjelaskan gain yang disebut *normalized gain* (gain ternormalisasi). Gain ternormalisasi (g) diformulasikan dalam bentuk seperti di bawah ini:

$$g = \frac{\text{Skor Postes} - \text{Skor Pretes}}{\text{Skor Ideal} - \text{Skor Pretes}}$$

Skor gain ternormalisasi dapat dikategorisasi kedalam tiga kategori, yaitu rendah, sedang, dan tinggi. Menurut Hake (Meltzer, 2002) kategori gain ternormalisasi adalah sebagai berikut:

$g < 0,3$: Rendah

$0,3 \leq g < 0,7$: Sedang

$g \geq 0,7$: Tinggi

4. Uji dua rata-rata

Uji hipotesis ini adalah untuk menguji apakah kedua skor rata-rata populasi siswa sama. Sebagai hipotesis alternatifnya adalah bahwa skor rata-rata populasi siswa dari kelompok eksperimen lebih besar. Langkah-langkah pengujian adalah sebagai berikut:

- a. Menghitung nilai rata-rata dari kedua kelompok untuk setiap aspek kemampuan matematika dengan rumus:

$$\bar{x} = \frac{\sum x}{n}$$

$\sum x$: Jumlah skor total seluruh siswa dari tiap kelompok

n : Banyaknya siswa untuk tiap kelompok

- b. Menentukan hipotesis statistik

$$H_0 : \mu_t = \mu_c$$

(skor rata-rata kemampuan siswa yang mendapat pembelajaran dengan strategi heuristik dalam kelompok kecil sama dengan skor rata-rata kemampuan siswa yang mendapat pembelajaran secara konvensional)

$$H_A : \mu_t > \mu_c$$

(skor rata-rata kemampuan siswa yang mendapat pembelajaran dengan strategi heuristik dalam kelompok kecil lebih besar daripada skor rata-rata kemampuan siswa yang mendapat pembelajaran secara konvensional)



c. Menghitung statistik uji

Apabila sebaran data berdistribusi normal dan varians populasinya homogen, sedangkan varians populasi tidak diketahui sehingga sebagai penggantinya ditaksir dari sampel, maka statistik uji yang digunakan adalah statistik uji t . Rumusnya adalah sebagai berikut:

$$t = \frac{\bar{x}_t - \bar{x}_c}{\hat{\sigma} \sqrt{\frac{1}{n_t} + \frac{1}{n_c}}}$$

dimana, \bar{x}_t : Rata-rata kelompok eksperimen

\bar{x}_c : Rata-rata kelompok kontrol

n_t : Banyaknya siswa pada kelas eksperimen

n_c : Banyaknya siswa pada kelas kontrol

$\hat{\sigma}$: Simpangan baku gabungan taksiran

$\hat{\sigma}$ adalah taksiran simpangan baku gabungan antara kedua kelompok kontrol dan eksperimen yang dihitung berdasarkan rumus:

$$\hat{\sigma} = \sqrt{\frac{n_t S_t^2 + n_c S_c^2}{n_t + n_c - 2}}$$

dimana, S_t^2 : Varians skor kelompok eksperimen

S_c^2 : Varians skor kelompok kontrol

Kriteria pengujiannya adalah sebagai berikut :

Hipotesis nol (H_0) ditolak jika nilai t -hitung berada diluar interval $(-t_{\text{daftar}}, t_{\text{daftar}})$. Nilai t -daftar diperoleh dari daftar dengan derajat kebebasan $dk = n_t + n_c - 2$, dan taraf signifikansi pada 0,05.

Apabila sebaran data tidak berdistribusi normal maka untuk menguji kesamaan dua rata-rata digunakan statistik uji nonparametrik yaitu uji Mann-Whitney (statistik U). Rumus statistik uji yang digunakan (Edward, 1993) adalah sebagai berikut:

$$U = (n_x)(n_y) + \frac{n_x(n_x + 1)}{2} - \sum R_x$$

di mana,

U : Statistik uji Mann-Whitney

n_x, n_y : Ukuran sampel pada kelompok 1 dan kelompok 2

R_x : Ranking yang diberikan pada kelompok yang ukuran sampelnya n_x

Untuk sampel berukuran besar ($n > 20$), Edward (1993) menyarankan untuk menggunakan pendekatan ke distribusi normal dengan bentuk statistik sebagai berikut:

$$z = \frac{U - \frac{n_x n_y}{2}}{\sqrt{\frac{n_x n_y (n_x + n_y + 1)}{12}}}$$

di mana,

z : statistik uji z yang berdistribusi normal $N(0,1)$

5. Menganalisis hasil pengukuran skala sikap

Validitas isi skala sikap

Analisis skala sikap meliputi analisis terhadap validitas isi, validitas butir (untuk menyeleksi butir skala sikap), dan reliabilitas skala sikap. Menurut Sumarmo (2002), estimasi validitas isi skala sikap dilakukan dengan cara menyusun terlebih dahulu kisi-kisi skala sikap yang memuat aspek sikap yang

diukur, menyusun butir pernyataan berdasarkan kisi-kisi yang telah dirumuskan, kemudian keseluruhan butir skala sikap ditimbang oleh pakar. Berkaitan dengan validitas isi skala sikap yang akan digunakan, validasi telah diberikan oleh dosen pembimbing yang dalam hal ini juga merupakan pakar dalam evaluasi pendidikan matematika.

Validitas butir (Seleksi butir skala sikap)

Sebelum melakukan penyeleksian butir skala sikap, terlebih dahulu dilakukan pemberian skor terhadap butir skala sikap. Model skala yang digunakan adalah skala Likert. Dalam skala ini akan digunakan empat skala sikap yaitu Sangat Setuju (SS), Setuju (S), Tidak Setuju (TS), dan Sangat Tidak Setuju (STS). Skala Netral (N) tidak digunakan dalam skala sikap ini untuk menghindari jawaban dengan respon netral. Statistik uji yang digunakan untuk menyeleksi butir skala sikap adalah statistik uji t , dengan rumus:

$$t = \frac{\bar{x}_T - \bar{x}_R}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_R^2}{n_R}}}$$

di mana,

\bar{x}_T : Rata-rata skor kelompok tinggi

\bar{x}_R : Rata-rata skor kelompok rendah

s_T^2 : Varians kelompok tinggi

s_R^2 : Varian kelompok rendah

n_T : Banyaknya subjek pada kelompok tinggi

n_R : Banyaknya subjek pada kelompok rendah

Validitas butir diestimasi dengan membandingkan nilai t hitung dengan t tabel pada taraf signifikansi 0,05.

Reliabilitas skala sikap

Selanjutnya setelah dipilih butir skala sikap yang memenuhi kriteria, dilakukan uji reliabilitas skala sikap. Koefisien reliabilitas dihitung dengan menggunakan Alpha-Cronbach.

Analisis deskriptif hasil skala sikap

Analisis deskriptif dilakukan untuk melihat sikap siswa terhadap variabel sikap yang diukur. Selain itu untuk melihat pernyataan mana yang lebih dominan (persentasi skor tertinggi) dan tidak dominan (persentasi skor terendah) bagi siswa.

6. Membuat kesimpulan secara umum dari hasil pengolahan data.

Setelah data diolah sehingga memberikan informasi yang bermakna, tahap selanjutnya adalah memberikan interpretasi atas nilai-nilai hasil pengolahan. Interpretasi yang diberikan sesuai dengan tujuan pengolahan data, kemudian membuat kesimpulan secara umum terhadap penelitian yang dilakukan.



