

diujicobakan. Kualitas soal bukanlah satu-satunya hal yang mempengaruhi valid atau tidak validnya soal.

Berbagai faktor lain dapat mempengaruhi validitas soal. Berdasarkan pengamatan peneliti, guru belum pernah memberikan latihan soal seperti soal pilihan ganda nomor 1 selama proses pembelajaran. Guru kebanyakan memberikan latihan soal-soal hitungan seperti “Hasil dari $3x + x = \dots$ ”. Begitu juga soal-soal dalam LKS dan buku paket.

Pada lembar penilaian aljabar, peneliti menemukan bahwa 4 peserta memberikan lingkaran/garis bawah pada nomor soal pilihan ganda nomor satu yang berarti siswa menganggap sulit soal atau siswa masih ragu dengan jawabannya. Ditemukan pula bahwa 6 peserta lain menghapus pilihan jawaban yang dianggap benar untuk pertama kali dan mengganti pilihan jawaban yang dianggap benar.

Hal ini menunjukkan bahwa sebagian peserta uji coba dimungkinkan menggunakan teknik *guessing* atau menduplikasi jawaban temannya dalam

aljabar, guru mengecek pengetahuan awal siswa tentang operasi bilangan bulat. Awalnya siswa memang terlihat berusaha mengingat kembali materi tersebut. Namun setelah diberi contoh soal beberapa kali, akhirnya sebagian besar dari siswa mampu mengikuti materi dengan baik.

Pengetahuan awal siswa sangat berpengaruh terhadap proses siswa dalam menerima materi yang akan disampaikan selanjutnya. Untuk menjawab soal nomor 2, gabungan pengetahuan tentang konsep perkalian dan penjumlahan bentuk aljabar sangat diperlukan. Siswa yang belum paham betul tentang perkalian dan penjumlahan bentuk aljabar akan kesulitan dalam menjawab soal ini.

Data menunjukkan bahwa 17 dari 20 peserta yang menjawab salah memilih pilihan jawaban b, 2 peserta lainnya memilih pilihan jawaban a, dan 1 peserta memilih pilihan jawaban d. Dengan demikian soal pilihan ganda memang dapat menjadi salah satu alat ukur yang tepat untuk mengetahui siswa yang sudah

aljabar menggunakan variabel yang terlihat jelas berbeda seperti “ Hasil dari $2a + 5b + a = \dots$ ” atau “Hasil dari $6x - 4y^2 - 2x = \dots$ ”. Soal – soal latihan pada Lembar Kerja Siswa dan buku paket pun juga demikian.

Hal ini menyebabkan sebagian siswa kebingungan ketika dihadapkan dengan soal pilihan ganda nomor 3. Sebagian siswa menganggap bahwa variabel a^2 sejenis dengan variabel a . Terbukti bahwa saat uji coba instrumen berlangsung, ada salah satu peserta yang menanyakan hal ini. Di sisi lain, fenomena ini bagus karena ini adalah satu bentuk bukti bahwa soal dapat mengukur kemampuan siswa yang sesungguhnya tentang penjumlahan bentuk aljabar.

Data hasil uji coba menunjukkan bahwa 55 dari 65 peserta menjawab soal dengan benar. 5 peserta memilih pilihan jawaban b dan 5 peserta memilih pilihan jawaban d. Dengan demikian soal pilihan ganda dirasa dapat membedakan siswa yang sudah atau belum paham tentang penjumlahan dan perkalian

bentuk aljabar. Namun ada 4 peserta yang ditemukan menghapus pilihan jawaban yang dianggap benar untuk pertama kali dan mengganti pilihan jawaban yang dianggap benar dan pilihan jawaban yang dipilih memang benar. Ada satu peserta lain yang memberikan tanda lingkaran pada soal nomor 3 yang berarti soal tersebut dianggap sulit. Namun, pada akhirnya peserta tersebut memilih pilihan jawaban yang benar.

Hal ini menunjukkan bahwa sebagian kecil peserta uji coba bisa jadi menggunakan teknik *guessing* atau menduplikasi jawaban temannya dalam menjawab soal. Tentu saja hal ini mempengaruhi hasil perhitungan uji validitas soal pilihan ganda nomor tiga, dan akhirnya soal termasuk dalam kategori validitas yang rendah (kurang).

valid. Namun, soal valid dalam kategori rendah (kurang).

Berdasarkan pengamatan peneliti, guru sering memberikan latihan soal yang *familiar* dengan soal nomor 4. Tentu saja hal ini berpengaruh terhadap hasil uji coba. Dengan demikian soal akan sulit digunakan dalam mengukur kemampuan peserta uji coba karena soal dapat dikerjakan oleh sebagian besar peserta. Terbukti bahwa 59 dari 65 peserta mampu menjawab soal tersebut dengan benar.

Selain itu, didapati bahwa 6 peserta memilih pilihan jawaban yang salah yaitu pilihan jawaban c. Jadi terlihat bahwa pengecoh a dan b kurang masuk akal. Kondisi soal tentu mempengaruhi kualitas validitas soal. Pengecoh yang juga merupakan bagian dari suatu soal pasti mempengaruhi hasil uji validitas. Pengecoh yang kurang baik menyebabkan peserta dapat menggunakan teknik *guessing* dengan lebih mudah dalam memilih jawaban yang benar. Dengan

peserta uji coba mampu menjawab soal ini dengan benar yaitu 62 dari 65 peserta. Sedangkan untuk sisanya, yaitu 3 peserta memilih pilihan jawaban yang salah.

Namun setelah diteliti ada 1 peserta yang ditemukan menghapus pilihan jawaban yang salah dan menggantinya dengan jawaban yang benar. Kemudian ditemukan pula satu peserta lain yang memberikan tanda garis bawah pada soal nomor 5 yang berarti soal tersebut dianggap sulit. Namun, pada akhirnya peserta tersebut memilih jawaban yang benar.

Hal ini menunjukkan bahwa peserta uji coba dimungkinkan menggunakan teknik *guessing* atau menduplikasi jawaban temannya dalam menjawab soal karena ragu dengan kemampuannya sendiri. Tentu saja hal ini mempengaruhi hasil perhitungan uji validitas soal pilihan ganda nomor 5, akhirnya soal tidak termasuk dalam kategori validitas yang tinggi, namun sedang (cukup baik).

dalam mengukur kemampuan peserta uji coba karena soal dapat dikerjakan oleh sebagian besar peserta. Terbukti bahwa 60 dari 65 peserta mampu menjawab soal tersebut dengan benar.

Peneliti mendapati bahwa 3 peserta ragu-ragu dalam memilih jawaban yang benar, padahal soal dapat dikerjakan dalam kurun waktu kurang dari 1 menit. Namun pada akhirnya peserta tersebut memilih pilihan jawaban yang benar. Jika terjadi demikian, maka ada kemungkinan bahwa sebagian kecil dari peserta menduplikasi jawaban temannya. Jika hal seperti demikian terjadi, berarti soal belum dapat maksimal dalam membedakan kemampuan siswa yang sudah paham atau belum terkait materi akan soal yang diujikan. Dengan penjelasan di atas dapat diterima bahwa perhitungan validitas soal masuk dalam kategori sedang (cukup baik).

latihan soal, memang sebagian peserta terlihat paham, namun sebagian peserta lainnya perlu bimbingan dalam mengerjakan soal. Untuk dapat menjawab soal dengan benar, dibutuhkan pemahaman konsep operasi aljabar dan konsep pengoperasian bilangan/variabel positif dan negatif. Jika siswa belum memahami konsep tersebut, maka peserta akan kesulitan atau kurang teliti dalam mengerjakan soal.

Menurut pengamatan peneliti selama siswa mengerjakan latihan soal, sebagian besar peserta masih banyak bertanya kepada guru konsep perkalian bilangan/variabel positif dan negatif. Suatu pilihan yang tepat apabila peneliti menampilkan soal nomor 7 sebagai salah satu soal dalam instrumen penilaian aljabar karena soal ini dirasa mampu membedakan peserta yang paham dan tidak paham.

Hal tersebut terbukti dengan didapatkannya data bahwa 28 dari 65 peserta mampu menjawab dengan benar. Dengan penjelasan di atas dapat diterima bahwa

Untuk dapat menjawab soal dengan benar, dibutuhkan pemahaman konsep perpangkatan bentuk aljabar. Dan siswa sangat antusias dalam mengikuti proses pembelajaran ketika materi ini disampaikan. Ketika guru mengecek pemahaman siswa dengan menunjuk salah satu siswa maju ke depan kelas, siswa tersebut dengan percaya diri maju dan mengerjakan soal dengan sempurna.

Soal nomor 8 dirasa terlalu mudah untuk peserta. Terbukti bahwa 54 dari 65 peserta mampu menjawab soal tersebut dengan benar. Namun, peneliti mendapati bahwa pengecoh c bekerja lebih dari baik. 11 peserta yang menjawab salah memilih pilihan jawaban c. Kemungkinan besar siswa kurang teliti dalam memilih pilihan jawaban. Fenomena ini menyebabkan soal tidak dapat membedakan kemampuan siswa dengan tepat sehingga soal valid dalam kategori rendah (kurang).

i. Soal Pilihan Ganda Nomor 9

$$9. \quad (y + 2)(y - 9) =$$

- $2y - 7$
- $y^2 - 7$
- $y^2 - 7y - 7$
- $y^2 - 7y - 18$

Gambar 4.15

Soal Pilihan Ganda Nomor 9 Sebelum Revisi

Menurut dua validator, yaitu validator ketiga dan validator keempat, soal pilihan ganda nomor 9 telah memenuhi 11 karakteristik soal dalam *Multiple-choice Item Review Checklist*. Sedangkan menurut validator pertama soal belum memenuhi karakteristik pertama yaitu apakah soal sudah benar? Validator pertama memberikan saran kepada peneliti untuk memperbaiki soal dengan memberikan kalimat pertanyaan soal sebagai berikut. “Hasil dari operasi aljabar $(y + 2)(y - 2)$ adalah...”

Menurut validator kedua, soal belum memenuhi karakteristik kedua yaitu apakah soal yang diberikan jelas? Validator kedua juga berpendapat bahwa soal

bilangan/variabel positif dan negatif. Selain itu, selama pengerjaan latihan soal, peneliti mendapati bahwa sebagian kecil siswa menganggap bahwa soal seperti $(y + 2)(y - 2)$ adalah soal penjumlahan bentuk aljabar, bukan soal perkalian bentuk aljabar.

Jika siswa belum memahami konsep perkalian bentuk aljabar untuk bilangan/variabel positif dan negatif, maka peserta akan kesulitan dalam mengerjakan penyelesaian soal atau kurang teliti dalam mengerjakannya.

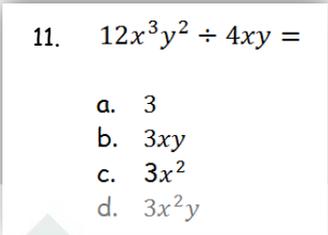
Suatu pilihan yang tepat apabila peneliti menampilkan soal nomor 9 sebagai salah satu soal dalam instrumen penilaian aljabar karena soal ini dirasa mampu membedakan peserta yang paham dan tidak paham konsep perkalian bentuk aljabar.

Data uji coba menunjukkan bahwa bahwa 46 dari 65 peserta mampu menjawab dengan benar. Dengan penjelasan di atas dapat diterima bahwa perhitungan validitas soal masuk dalam kategori cukup baik.

memberikan soal tentang menyatakan luas suatu bangun segi empat dengan bentuk aljabar. Saat guru memberikan latihan soal tersebut, sebagian besar siswa terlihat masih kesulitan dalam mengerjakan soal. Latihan soal-soal yang menuntut pemahaman memang sangat jarang diberikan guru sebagai latihan soal.

Dari penjelasan di atas, peneliti menyimpulkan bahwa soal nomor 10 terlalu sulit untuk dikerjakan peserta uji coba. Terbukti dengan data yang menunjukkan bahwa hanya 5 peserta saja yang mampu mengerjakan soal ini dengan benar. 44 peserta memilih pilihan jawaban c, 12 peserta memilih pilihan jawaban b, dan 4 peserta tidak memilih satu pun dari pilihan jawaban. Dengan demikian, soal hanya mampu membedakan pemahaman sebagian kecil dari peserta uji coba. Jadi dapat diterima bahwa perhitungan validitas soal masuk dalam kategori rendah (kurang baik).

k. Soal Pilihan Ganda Nomor 11



11. $12x^3y^2 \div 4xy =$

- a. 3
- b. $3xy$
- c. $3x^2$
- d. $3x^2y$

Gambar 4.18

Soal Pilihan Ganda Nomor 11 Sebelum Revisi

Menurut tiga validator, yaitu validator kedua, validator ketiga dan validator keempat, soal pilihan ganda nomor 8 telah memenuhi 11 karakteristik soal dalam *Multiple-choice Item Review Checklist*. Sedangkan menurut validator pertama soal belum memenuhi karakteristik kedua yaitu apakah soal yang diberikan jelas? Validator pertama memberikan saran kepada peneliti untuk memperbaiki soal dengan memberikan kalimat pertanyaan soal sebagai berikut. “Hasil dari operasi aljabar $12x^3y^2:4xy$ adalah...”. Kemudian peneliti merevisi soal dan didapat hasil revisi sebagai berikut.

para ahli. Hasil perhitungan validitas menggunakan korelasi *Pearson Product Moment* dengan taraf signifikansi 1% juga menunjukkan bahwa soal valid, namun termasuk dalam kategori sedang (cukup baik).

Berdasarkan pengamatan peneliti dalam proses pembelajaran, guru sering memberikan latihan soal yang *familiar* dengan soal nomor 12. Pada awal proses penyampaian materi, guru memberikan satu contoh soal untuk diselesaikan bersama di depan kelas. Sebagian siswa terlihat antusias dalam menerima materi ini. Sebagian siswa minta untuk diberi contoh soal lagi dan guru pun memberi contoh soal lain kepada siswa. Hingga akhirnya siswa paham dengan penjelasan yang disampaikan oleh guru.

Hal ini pasti berpengaruh terhadap hasil uji coba. Dengan demikian soal akan sulit digunakan dalam membedakan kemampuan peserta uji coba karena soal dapat dikerjakan oleh sebagian besar peserta.

Soal nomor 12 dirasa mudah untuk peserta. Terbukti bahwa 51 dari 65 peserta mampu menjawab

Menurut tiga validator, yaitu validator pertama, validator ketiga, dan validator keempat, soal pilihan ganda nomor 15 telah memenuhi 11 karakteristik soal dalam *Multiple-choice Item Review Checklist*. Sedangkan menurut validator kedua soal belum memenuhi karakteristik ketujuh yaitu apakah pengecoh masuk akal dan terbukti salah? Validator kedua memberikan saran kepada peneliti untuk mengganti pengecoh a dan b. Karena pengecoh a dan b dianggap kurang baik. Validator menyarankan agar peneliti menggunakan $8x$ sebagai pilihan jawaban a dan untuk pilihan jawaban b peneliti disarankan untuk menggunakan pilihan jawaban yang lebih memungkinkan di pilih peserta. Peneliti kemudian menggunakan $12x$ untuk pilihan jawaban b. Dan didapat soal hasil revisi sebagai berikut.

membedakan kemampuan peserta uji coba karena soal hanya dapat dikerjakan oleh sebagian peserta. Terbukti bahwa data menunjukkan bahwa hanya 19 siswa mampu mengerjakan soal dengan benar.

Namun, setelah peneliti mencoba meneliti kembali, didapat bahwa dalam soal tersebut perlu ditambahkan “Banyak buku Farhan adalah x ”. Ini menunjukkan bahwa kualitas soal kurang baik. Tentu hal ini menyebabkan soal tidak valid.

Penjelasan tersebut adalah faktor-faktor penyebab soal valid dan tidak valid yang sebagian besar berasal dari jawaban siswa. Adapun faktor lain yang mempengaruhi penilaian aljabar akan dijelaskan lebih rinci selanjutnya.

Pertama adalah kondisi peserta saat uji coba. Pada saat uji coba berlangsung, sebagian siswa terlihat kurang siap. Namun sebagian lainnya terlihat antusias dalam mengikuti uji coba ini. Ada beberapa siswa yang terlihat malas-malasan dalam mengerjakan instrumen penilaian, mengobrol dengan teman sebelahnya. Beberapa diantaranya juga ada yang sibuk melirik pekerjaan teman

sebelahnya. Ada beberapa anak yang sibuk melihat ke luar jendela, melihat siswa lain yang sedang olahraga. Beberapa diantaranya sibuk menukar tempat duduk demi dapat mendekati teman yang dianggapnya mampu mengerjakan penilaian aljabar. Namun, sebagian peserta lain mengerjakan penilaian dengan tekun dan sungguh-sungguh.

Kedua adalah proses pembelajaran. Proses pembelajaran tentu mempengaruhi hasil uji validitas soal. Latihan soal yang diberikan guru kepada siswa pasti mempengaruhi pengerjaan instrumen. Dari hasil pengamatan peneliti diketahui bahwa sebagian siswa lebih mudah dalam mengerjakan soal hitungan dibandingkan dengan soal yang menguji kemampuan pemahaman konsep sehingga soal-soal yang menguji kemampuan pemahaman kurang maksimal pengerjaannya.

Ketiga, jam pelajaran mata pelajaran matematika peserta uji coba penilaian aljabar kebanyakan bukan pada jam pertama tetapi jam ke-3, jam ke-4, jam ke-5, jam ke-6. Bahkan sebagian peserta uji coba mengikuti proses

pembelajaran matematika setelah mengikuti mata pelajaran olahraga. Tentu saja hal ini mempengaruhi kondisi fisik dan psikologis siswa dalam mengikuti proses pembelajaran pada jam berikutnya yaitu mata pelajaran matematika. Dengan demikian siswa dapat kurang fokus dalam menerima pembelajaran dan berakibat pada hasil pengerjaan penilaian aljabar. Peserta akan kesulitan dalam mengerjakan penilaian aljabar karena perhatian peserta pada saat penyampaian materi kurang maksimal.

Keempat, waktu pengerjaan soal selama 80 menit dianggap terlalu cepat oleh sebagian peserta. Ini menyebabkan peserta memberikan jawaban soal dalam situasi tergesa-gesa. Peserta yang memberikan jawaban dalam keadaan tergesa-gesa dimungkinkan menjawab tidak sesuai dengan kemampuan yang sesungguhnya.

Demikianlah analisis hasil uji validitas soal-soal pilihan ganda penilaian aljabar kelas VIII mengacu pada standar *TIMSS*.

soal dalam LKS dan buku Paket pun juga banyak yang sejenis dengan soal nomor 1.

Dalam soal ini siswa diminta untuk menunjukkan langkah-langkah dalam menyelesaikan soal. Skor yang didapat untuk tiap-tiap siswa sangat bervariasi. Sehingga soal ini lebih akurat dibanding soal pilihan ganda perihal pemberian skor. Data uji coba menunjukkan bahwa 31 peserta mendapat skor 5, 5 peserta mendapat skor 4, 3 peserta mendapat skor 3, 1 peserta mendapat skor 2, 4 peserta mendapat skor 1, dan 21 peserta mendapat skor 0. Terlihat bahwa soal dapat mengukur kemampuan peserta dalam menyelesaikan soal. Dalam lembar jawaban peserta dapat dilihat pengerjaan siswa yang paham tentang konsep penjumlahan bentuk aljabar, perkalian bentuk aljabar, dan operasi bilangan/variabel positif dan negatif. Dengan demikian dapat diterima bahwa soal valid dengan kategori tinggi (baik).

Menurut dua validator, yaitu validator ketiga dan validator keempat, soal uraian nomor 3 telah memenuhi 10 karakteristik soal dalam *Constructed-response Item Review Checklist*.

Menurut validator pertama, soal belum memenuhi karakteristik pertama yaitu apakah soal sudah benar? Validator pertama memberikan komentar bahwa "Kenapa harus ada 3b. Jawaban anak untuk keliling bisa berbeda". Validator pertama menganggap bahwa dengan adanya garis dengan panjang 3b akan menjadikan siswa kebingungan dalam mengartikan maksud dari soal.

Sedangkan menurut validator kedua, soal belum memenuhi karakteristik kedelapan yaitu apakah panduan pemberian skor jelas? Validator kedua menyarankan kepada peneliti untuk memperbaiki panduan dalam memberikan skor yaitu apabila jawaban salah maka skor 0.

datar dalam soal, sebagian peserta membagi bangun menjadi dua daerah dan menghitung keliling masing-masing daerah.

Sebagian peserta yang lain sudah mampu memahami alur perhitungan jawaban, namun kurang memahami konsep penjumlahan dan perkalian bentuk aljabar. Berdasarkan lembar jawaban peserta, sebagian kecil saja yang mampu mendapatkan skor maksimal yaitu 7 peserta, sedangkan sebagian besar peserta mendapatkan skor 3 dan sebagian kecil mendapat skor 1 dan 0.

Dengan demikian soal secara keseluruhan mampu membedakan kemampuan peserta dalam menyelesaikan soal. Dengan demikian dapat diterima bahwa soal valid dengan kategori sedang (cukup baik).

Product Moment dengan taraf signifikansi 1% menunjukkan bahwa soal valid dengan kategori sedang (cukup baik).

Berdasarkan pengamatan peneliti, dalam menyelesaikan soal ini diperlukan pemahaman konsep perkalian bentuk aljabar dan penjumlahan pecahan bentuk aljabar. Dengan soal ini dapat diketahui siswa yang paham tentang perkalian bentuk aljabar dan tidak. Sebagian peserta yang kurang paham akan perkalian bentuk aljabar seringkali menjumlahkan langkah penyelesaian yang seharusnya adalah perkalian bentuk aljabar. Sebagian peserta lainnya menambahkan pangkat pada langkah penyelesaian penjumlahan bentuk aljabar.

Dengan demikian dapat diterima bahwa soal valid dengan kategori sedang (cukup baik).

menunjukkan bahwa soal valid dengan kategori sedang (cukup baik).

Berdasarkan pengamatan peneliti, guru sering memberikan soal sejenis sebagai latihan. Dengan demikian, siswa telah familiar dengan soal nomor 5. Soal ini memang dirasa terlalu mudah untuk peserta uji coba, namun sebagian peserta masih kebingungan dalam menyatakan suatu soal cerita dalam bentuk aljabar.

Inilah tujuan dari soal nomor 5, membedakan siswa yang mampu menyatakan suatu soal cerita dalam bentuk aljabar dan tidak. Data uji coba menunjukkan bahwa, soal dapat digunakan untuk membedakan siswa yang mampu menyatakan suatu soal cerita dalam bentuk aljabar dan tidak. Sehingga dapat diterima bahwa soal valid dengan kategori sedang (cukup baik).

$$\sigma_2^2 = \frac{\sum X_2^2 - \frac{(\sum X_2)^2}{N}}{N} = \frac{45 - \frac{(45)(45)}{65}}{65} = 0,213$$

$$\sigma_3^2 = \frac{\sum X_3^2 - \frac{(\sum X_3)^2}{N}}{N} = \frac{55 - \frac{(55)(55)}{65}}{65} = 0,1302$$

$$\sigma_4^2 = \frac{\sum X_4^2 - \frac{(\sum X_4)^2}{N}}{N} = \frac{59 - \frac{(59)(59)}{65}}{65} = 0,0838$$

$$\sigma_5^2 = \frac{\sum X_5^2 - \frac{(\sum X_5)^2}{N}}{N} = \frac{62 - \frac{(62)(62)}{65}}{65} = 0,044$$

$$\sigma_6^2 = \frac{\sum X_6^2 - \frac{(\sum X_6)^2}{N}}{N} = \frac{60 - \frac{(60)(60)}{65}}{65} = 0,071$$

$$\sigma_7^2 = \frac{\sum X_7^2 - \frac{(\sum X_7)^2}{N}}{N} = \frac{28 - \frac{(28)(28)}{65}}{65} = 0,2452$$

$$\sigma_8^2 = \frac{\sum X_8^2 - \frac{(\sum X_8)^2}{N}}{N} = \frac{54 - \frac{(54)(54)}{65}}{65} = 0,1406$$

$$\sigma_9^2 = \frac{\sum X_9^2 - \frac{(\sum X_9)^2}{N}}{N} = \frac{46 - \frac{(46)(46)}{65}}{65} = 0,2069$$

$$\sigma_{10}^2 = \frac{\sum X_{10}^2 - \frac{(\sum X_{10})^2}{N}}{N} = \frac{5 - \frac{(5)(5)}{65}}{65} = 0,071$$

$$\sigma_{11}^2 = \frac{\sum X_{11}^2 - \frac{(\sum X_{11})^2}{N}}{N} = \frac{57 - \frac{(57)(57)}{65}}{65} = 0,1079$$

soal. Banyak butir soal mempegaruhi tingkat reliabilitas. Semakin banyak soal, maka reliabilitas semakin tinggi. Soal-soal pilihan ganda terdiri atas 15 butir soal.

Kedua, homogenitas soal. Homogenitas soal adalah kesejensan isi soal-soal. Homogenitas soal-soal pilihan ganda termasuk tinggi karena hanya memuat soal-soal yang mengukur kemampuan dalam menyelesaikan operasi aljabar. Homogenitas yang tinggi cenderung mengarah pada tingginya tingkat reliabilitas. Ketiga, heterogenitas kelompok. Heterogenitas kelompok adalah keberagaman kemampuan subyek yang digunakan untuk uji coba. Subyek yang digunakan dalam uji coba ini tergolong heterogen karena diikuti oleh peserta dengan kemampuan beragam.

Namun, kondisi pada saat tes diberikan tidak seragam karena menyesuaikan jadwal mata pelajaran matematika tiap kelas. Selain itu peneliti mendapati bahwa motivasi antar individu begitu beragam. Sebagian siswa begitu antusias dalam mengerjakan. Sebagian lainnya tidak begitu tekun dalam mengerjakan.

$$\sigma_2^2 = \frac{\sum X_2^2 - \frac{(\sum X_2)^2}{N}}{N} = \frac{1201 - \frac{(243)(243)}{65}}{65} = 4,501$$

$$\sigma_3^2 = \frac{\sum X_3^2 - \frac{(\sum X_3)^2}{N}}{N} = \frac{485 - \frac{(147)(147)}{65}}{65} = 2,347$$

$$\sigma_4^2 = \frac{\sum X_4^2 - \frac{(\sum X_4)^2}{N}}{N} = \frac{602 - \frac{(136)(136)}{65}}{65} = 4,884$$

$$\sigma_5^2 = \frac{\sum X_5^2 - \frac{(\sum X_5)^2}{N}}{N} = \frac{1309 - \frac{(273)(273)}{65}}{65} = 2,498$$

Dan didapat jumlah varians untuk seluruh item sebagai berikut.

$$\sum \sigma_i^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + \sigma_5^2 = 19,378$$

Kemudian akan dihitung varians total. Dari data dalam lampiran hs didapat $\sum Y_i^2 = 18107$ dan $\sum Y_i = 989$. Berdasarkan data tersebut, berikut ini adalah perhitungan varians total.

$$\begin{aligned} \sigma_t^2 &= \frac{\sum_{i=1}^n Y_t^2 - \frac{(\sum_{i=1}^n Y_t)^2}{N}}{N} = \frac{18107 - \frac{(989)(989)}{65}}{65} \\ &= 47,06 \end{aligned}$$

Perhitungan uji reliabilitas menggunakan rumus Cronbach's Alpha menunjukkan bahwa reliabilitas soal-soal uraian termasuk dalam kategori korelasi tinggi (baik). Ini berarti, hasil pengukuran soal-soal uraian dapat dipercaya.

Hal yang paling mempengaruhi reliabilitas adalah kualitas validitas soal. Validitas soal yang baik menjadikan tingkat reliabilitas semakin tinggi. Namun, berbagai faktor turut mempengaruhi reliabilitas soal. Pertama, jumlah butir soal. Banyak butir soal mempengaruhi tingkat reliabilitas. Semakin banyak soal, maka reliabilitas semakin tinggi. Soal-soal uraian terdiri atas 5 butir soal. Meskipun hanya terdiri dari 5 butir soal saja, namun ini bukanlah satu-satunya faktor yang mempengaruhi tingkat reliabilitas.

Kedua, homogenitas soal. Homogenitas soal adalah kesejajaran isi soal-soal. Homogenitas soal-soal uraian termasuk tinggi karena hanya memuat soal-soal yang mengukur kemampuan dalam menyelesaikan operasi aljabar. Homogenitas yang tinggi cenderung mengarah pada tingginya tingkat reliabilitas.

soal yang kurang baik. Karena soal yang baik akan mendukung didapatkannya hasil uji validitas dengan kategori validitas yang tinggi. Namun, hasil uji validitas tersebut perlu dianalisis kembali karena banyak sekali faktor yang dapat mempengaruhi kevalidan soal pilihan ganda. Berbagai faktor yang besar pengaruhnya terhadap kevalidan soal pilihan ganda adalah sebagai berikut.

Pertama, soal pilihan ganda sangat rentan akan teknik *guessing*. Teknik *guessing* adalah suatu teknik dalam memilih pilihan jawaban yang dianggap benar dengan cara menebak. Dalam artian, peserta uji coba tidak menggunakan kemampuannya dalam mengerjakan soal, namun peserta uji coba hanya memperkirakan satu pilihan jawaban yang dianggap berpotensi besar sebagai jawaban yang benar. Pemilihan jawaban didasarkan dari berbagai macam sudut pandang, sesuai dengan batas kemampuan masing-masing peserta.

Dapat dipastikan, ketika peserta tidak dapat mengerjakan soal, maka sebagian besar peserta akan lebih memilih teknik *guessing* dalam menjawab daripada tidak

memilih satu pun pilihan jawaban. Ini semua karena ketika jawaban salah maka tidak ada pengaruh terhadap hasil penilaian. Dengan demikian teknik *guessing* yang digunakan peserta menyebabkan soal pilihan ganda tidak dapat mengukur kemampuan peserta dalam menyelesaikan soal dengan tepat.

Kedua, jawaban soal pilihan ganda lebih mudah diduplikasi. Soal pilihan ganda tidak mengharuskan peserta menuliskan langkah pengerjaan sehingga lebih mudah bagi peserta untuk melihat hasil pekerjaan temannya. Sebagian peserta menduplikasi jawaban peserta lainnya dengan cara saling bekerja sama, melirik tanpa sepengetahuan, dan lain sebagainya. Dengan demikian teknik duplikasi jawaban yang digunakan peserta menyebabkan soal pilihan ganda tidak dapat mengukur kemampuan peserta dalam menyelesaikan soal dengan tepat.

Dua faktor di atas adalah yang mempengaruhi kevalidan soal dari cara peserta mengerjakan penilaian aljabar. Berikut ini akan dipaparkan faktor kevalidan soal

dari kondisi peserta pada saat tes berlangsung dan proses pembelajaran materi aljabar.

Kondisi tiap peserta pada saat tes berlangsung mempengaruhi pengerjaan tes peserta itu sendiri dan peserta lain. Antusiasme dan motivasi masing-masing peserta tes menentukan seberapa tekun dan sungguh-sungguh tiap peserta mengerjakan tes. Antusiasme dan motivasi yang tinggi tentu menjadikan peserta tekun dan sungguh-sungguh dalam mengerjakan tes, begitu pula sebaliknya. Dengan demikian, antusiasme dan motivasi peserta yang tinggi mendukung didapatkannya hasil uji validitas dengan kategori validitas yang tinggi.

Proses pembelajaran materi aljabar tentu mempengaruhi pengerjaan tes peserta. Penanaman konsep materi operasi aljabar oleh guru kepada siswa harus selalu ditekankan. Selama ini siswa yang tidak paham betul dengan konsep operasi aljabar sebagian besar hanya mampu mengerjakan soal-soal yang sejenis dengan latihan soal yang biasa diberikan oleh guru. Selain itu dengan pemahaman konsep yang baik siswa dapat lebih mudah

Sebelum diujicobakan, soal harus sudah divalidasi oleh para ahli. Dengan demikian, soal yang akan diujicobakan telah dianalisis secara kualitatif. Yang dimaksud dengan analisis secara kualitatif disini adalah setiap butir soal diperiksa satu persatu berdasarkan karakteristik soal yang ditelaah berdasarkan instrumen validasi. Hasil validasi para ahli tentu berisi saran perbaikan bagi soal-soal yang tidak valid secara kualitatif. Dengan demikian soal direvisi dan didapatkan soal yang lebih baik. Dengan kata lain, hal tersebut adalah salah satu usaha untuk meminimalisir soal-soal yang kurang baik. Karena soal yang baik akan mendukung didapatkannya hasil uji validitas dengan kategori validitas yang tinggi.

Dibanding soal pilihan ganda, soal uraian memang lebih akurat dalam mengukur kemampuan peserta uji coba dalam mengerjakan soal. Soal uraian cenderung tersusun oleh lebih dari satu macam konsep. Artinya, soal uraian hanya dapat diselesaikan apabila peserta uji coba menguasai semua konsep yang tersusun dalam soal uraian

yang diujikan. Sehingga soal uraian tentu lebih mampu mengukur kemampuan peserta dengan tepat.

Perlu digaris bawahi bahwa soal dengan bobot soal yang sedang dan tinggi sangat sesuai dengan teknik penyekoran soal uraian karena teknik penyekoran soal uraian dapat dilakukan dengan detail. Maksud dari teknik penyekoran yang detail adalah tiap langkah penyelesaian soal yang dikerjakan oleh siswa jika benar akan mendapat skor, sesuai dengan panduan *skoring*. Tentu saja hal ini mendukung didapatkannya hasil uji validitas dengan kategori validitas yang tinggi karena dapat dilihat kemampuan tiap peserta dalam menyelesaikan konsep-konsep materi operasi aljabar yang tersusun dalam soal uraian tersebut. Sehingga didapat hasil yang valid akan kemampuan peserta tes.

Namun demikian, banyak sekali faktor yang dapat mempengaruhi kevalidan soal uraian. Kondisi tiap peserta pada saat tes berlangsung mempengaruhi pengerjaan tes peserta itu sendiri dan peserta lain. Antusiasme dan motivasi masing-masing peserta tes menentukan seberapa

teknik dan sungguh-sungguh tiap peserta mengerjakan tes. Antusiasme dan motivasi yang tinggi tentu menjadikan peserta tekun dan sungguh-sungguh dalam mengerjakan tes, begitu pula sebaliknya. Dengan demikian, antusiasme dan motivasi peserta yang tinggi mendukung didapatkannya hasil uji validitas dengan kategori validitas yang tinggi.

Proses pembelajaran materi aljabar tentu mempengaruhi pengerjaan tes peserta. Penanaman konsep materi operasi aljabar oleh guru kepada siswa harus selalu ditekankan. Selama ini siswa yang tidak paham betul dengan konsep operasi aljabar sebagian besar hanya mampu mengerjakan soal-soal yang sejenis dengan latihan soal yang biasa diberikan oleh guru. Selain itu dengan pemahaman konsep yang baik siswa dapat lebih mudah mengerjakan soal-soal baik dalam bentuk hitungan atau bentuk pemahaman. Pemahaman konsep yang baik oleh peserta mendukung didapatkannya hasil uji validitas dengan kategori validitas yang tinggi.

Maksud dari konsisten adalah bagaimana skor tes konsisten dari pengukuran satu ke lainnya. Hal tersebut dapat diketahui dengan menghitung reliabilitas hasil uji coba menggunakan rumus korelasi Cronbach's Alpha.

Hal yang paling menentukan reliabilitas soal adalah kualitas kevalidan soal. Soal dengan kategori validitas yang tinggi akan mendapatkan reliabilitas yang tinggi pula. Dengan demikian hal-hal yang mempengaruhi validitas soal pilihan ganda juga mempengaruhi reliabilitas soal uraian.

Namun, hasil uji reliabilitas perlu dianalisis karena banyak sekali faktor yang dapat mempengaruhi kereliabelan soal uraian. Pertama, jumlah butir soal. Jumlah butir soal mempengaruhi reliabilitas soal. Jumlah soal yang semakin banyak akan mendukung didapatkannya reliabilitas yang tinggi, begitu pula sebaliknya. Ini adalah salah satu tugas peneliti untuk membuat penilaian dengan jumlah soal yang banyak.

Kedua, homogenitas soal. Homogenitas soal adalah kesejenisan isi materi yang diujikan soal-soal.

Homogenitas yang tinggi cenderung mengarah pada tingginya tingkat reliabilitas.

Perlu digaris bawahi bahwa soal yang banyak dengan homogenitas soal yang tinggi akan saling mendukung didapatkannya reliabilitas yang tinggi. Namun, Soal yang banyak dengan homogenitas soal yang rendah akan menjadikan reliabilitas menjadi rendah pula.

Ketiga, heterogenitas kelompok. Heterogenitas kelompok adalah keberagaman kemampuan subyek yang digunakan untuk uji coba. Uji coba yang diikuti oleh peserta dengan kemampuan beragam mendukung didapatkannya hasil reliabilitas yang tinggi, begitu pula sebaliknya.

Keempat, keseragaman waktu tes. Waktu tes yang seragam untuk seluruh peserta juga mendukung didapatkannya reliabilitas soal yang tinggi, begitu pula sebaliknya.

Berdasarkan penjelasan di atas dapat disimpulkan bahwa kereliabelan soal pilihan ganda dipengaruhi oleh berbagai faktor. Faktor yang paling menentukan

