

BAB I PENDAHULUAN

A. LATAR BELAKANG PENELITIAN

Tujuan belajar adalah untuk mengadakan perubahan didalam diri seperti mengubah kebiasaan dari yang buruk menjadi baik, mengubah sikap dari yang negatif menjadi yang positif (Dalyono, 2010:49). Dengan belajar seseorang dapat menambah pengetahuan dalam berbagai bidang ilmu, dapat berfikir kritis, kreatif dan lain sebagainya. Belajar memiliki tiga prinsip, yaitu *pertama*, prinsip belajar sebagai perubahan perilaku. *Kedua*, belajar sebagai proses. *Ketiga*, belajar sebagai bentuk pengalaman (Suprijono, 2009:4)

Dalam pemahaman umum, belajar merupakan proses untuk mendapatkan ilmu pengetahuan yang biasanya didapat melalui pendidikan formal yaitu sekolah. Dalam lingkungan sekolah, tentu terjadi proses belajar mengajar yang tidak lepas dari peran gurunya. Pada hakikatnya, guru dianggap komponen paling penting untuk menentukan keberhasilan peserta didik, karena ditangan gurulah kurikulum, sumber belajar, sarana dan prasarana, dan iklim pembelajaran dapat menjadi berarti bagi kehidupan peserta didik.

Karena guru dianggap menjadi unsur paling menentukan keberhasilan peserta didik, maka sudah sewajarnya dalam proses pembelajaran guru selalu melakukan berbagai evaluasi, pengukuran, serta penilaian sebagai umpan balik dalam upaya memperbaiki proses belajar mengajar, mengukur sejauhmana keberhasilan siswa, maupun sebagai informasi yang berharga dalam mengambil keputusan.

Tentunya informasi tersebut bisa didapat dengan melakukan prosedur pengukuran, karena hanya dengan melakukan pengukuran dapat dikuantifikasikan secara benar, yaitu melalui pengamatan indikator-indikatornya yang operasional. Keakuratan, kecermatan serta keobjektifan informasi tentu hanya akan diperoleh dengan pengukuran yang akurat, cermat dan objektif pula. Oleh karena itu, untuk mendapatkan semua informasi hasil ukur yang tidak menyesatkan, tentu

Dwi Putri Musdansi, 2014

**PERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA SMA BERDASARKAN TEKNIK
PENSKORAN DAN UKURAN SAMPEL**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

memerlukan alat ukur yang berkualitas tinggi sehingga informasi yang diperoleh dapat dipercaya.

Apabila alat ukur yang dibuat memberikan informasi yang tidak akurat digunakan dalam pengambilan keputusan, maka dapat dipastikan kesimpulan atau keputusan tersebut pastilah juga merupakan keputusan yang tidak tepat. Keputusan yang tidak tepat memang tidak langsung dapat dirasakan akibatnya, namun dapat menimbulkan hal buruk yang tidak dikehendaki. Sebagai contoh seseorang dapat ditolak oleh perguruan tinggi berdasarkan tes masuk yang diselenggarakan, padahal mahasiswa tersebut sangat potensial. Dari hal tersebut terlihat bahwa dengan menggunakan tes yang tidak mampu menghasilkan data yang reliabel dan valid, bukan saja calon mahasiswa yang bersangkutan yang dirugikan tetapi juga Universitas yang menolak boleh jadi kehilangan mahasiswa yang sangat potensial. Contoh tersebut memberi pelajaran bahwa betapa besar dampak yang ditimbulkan jika instrumen yang digunakan tidak reliabel dan valid jika digunakan dalam mengambil keputusan.

Oleh karenanya, para ahli psikometrika telah menetapkan beberapa kriteria penting bagi setiap alat ukur psikologi untuk dapat dinyatakan sebagai alat ukur yang baik, yaitu mampu menghasilkan data dan memberikan informasi yang akurat. "Kriteria yang dimaksud adalah valid, reliabel, objektif, standar, ekonomis dan praktis" (Azwar, 2012:2). Naga (2013) mengatakan alat ukur yang baik memberikan hasil yang konstan bila digunakan berulang-ulang, asalkan kemampuan yang diukur tidak berubah. Hal tersebut berkaitan dengan reliabilitas tes tersebut. Makin tinggi reliabilitasnya maka akan tinggi pula tingkat kepercayaan skor amatannya. Skor amatan akan mendekati skor tulennya jika komponen keliru dalam pengukuran semakin kecil.

Ada beberapa hal yang dapat mempengaruhi reliabilitas tes antara lain perbedaan kondisi individu, perbedaan kondisi tes, variasi pengadministrasian tes, serta kesalahan dan perbedaan pemberian skor (Sax,1980:281 dalam Busnawir). Disisi lain juga perlu dipertimbangkan faktor panjang tes, kecepatan, homogenitas kemampuan siswa, tingkat kesukaran butir yang dapat mempengaruhi reliabilitas tes (Mehren dan Lehman, 1973:115-117 dalam Busnawir). Menurut Harun

Dwi Putri Musdansi, 2014

PERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA SMA BERDASARKAN TEKNIK PENS KORAN DAN UKURAN SAMPEL

(2007:131) reliabilitas juga dipengaruhi oleh cara penyajian tes, suasana hati, sikap subjek terhadap tes, motivasi dan kondisi subjek, keadaan ruangan pengujian, cara memberikan tes, dan sebagainya. Seperti juga dikemukakan Thorndike dalam Surapranata (2009:87) bahwa salah satu dari enam faktor penyebab terjadinya perbedaan skor (interkonsistensi) adalah faktor yang tidak pernah diperhitungkan seperti menerka dan melihat soal yang dilihat sebelumnya.

Sebagaimana diketahui bahwa untuk menetapkan nilai yang berkaitan dengan kinerja dan hasil karya siswa dilakukan melalui evaluasi. Cara evaluasi yang umum dilakukan oleh pelaku pendidikan adalah dengan menggunakan tes. Tes dianggap sebagai alat ukur yang paling praktis untuk mengetahui kemampuan siswa. Seperti yang dikemukakan oleh Anastasi (2007) bahwa tes pada dasarnya adalah alat ukur yang objektif dan dibakukan atas sampel tertentu. Mengutip pendapat Frederick G. Brown dari buku Azwar: Frederick (Azwar, 2011:3) mengemukakan bahwa tes adalah “prosedur yang sistematis guna mengukur sampel perilaku seseorang. Tes dapat disusun dalam berbagai bentuk dan tipe sesuai dengan tujuan dan maksud penyusunan tes”.

Bentuk tes yang sampai saat ini paling banyak digunakan dalam dunia pendidikan adalah tes bentuk pilihan ganda. Tes bentuk pilihan ganda banyak digunakan karena memiliki kekuatan tersendiri. Zainul (1993:62-64) mengatakan bahwa *pertama*, butir soal tipe pilihan ganda dapat dikonstruksi dan digunakan untuk mengukur segala level tujuan intruksional, mulai dari yang sederhana sampai yang paling kompleks. *Kedua*, karakteristik dari butir pilihan ganda hanya menuntut waktu kerja peserta tes sangat minimal sehingga dapat menggunakan jumlah butir soal relatif banyak yang berimplikasi pada penarikan sampel pokok bahasan yang akan diujikan dapat lebih luas. *Ketiga*, penskoran hasil kerja peserta dapat dikerjakan secara objektif. *Keempat*, tipe butir soal dapat dikonstruksi sehingga menuntut kemampuan peserta tes untuk membedakan berbagai tingkatan kebenaran sekaligus. *Kelima*, jumlah *option* yang dapat disediakan lebih dari dua. *Keenam*, tipe butir soal pilihan ganda memungkinkan dilakukan analisis butir soal secara baik. *Ketujuh*, tingkat kesukaran butir soal dapat dikendalikan dengan

mengubah tingkat homogenitas alternatif jawaban. *Kedelapan*, informasi yang diberikan lebih kaya.

Selain kekuatan-kekuatan tersebut, tes pilihan ganda juga memiliki keterbatasan yaitu sukar dikonstruksi, ada kecenderungan pembuat tes mengkonstruksi butir soal tipe ini hanya mengukur ranah ingatan, atau aspek yang paling rendah yaitu ranah kognitif saja. Selain itu setiap jawaban butir tipe objektif memiliki peluang peserta tes untuk melakukan terkaan atau tebakan. Tingkat terkaan adalah satu per jumlah *option* dalam butir soal itu. Terkaan pada soal pilihan ganda sebenarnya dikarenakan tidak adanya antisipasi agar pengikut tes tidak melakukan tebakan. Antisipasi ini berkaitan dengan teknik penskoran. Teknik penskoran yang saat ini umum digunakan adalah penskoran menjumlahkan jawaban betul atau disebut juga *number right score*. Dengan teknik ini, bisa saja soal sukar direspon benar oleh siswa yang tidak memiliki kecukupan pengetahuan, dan dijawab salah oleh yang memiliki kecukupan pengetahuan. Jawaban benar yang diberikan oleh peserta tes yang memiliki kemampuan rendah dan jawaban salah oleh kemampuan tinggi tentu berakibat pada reliabilitas yang dihasilkan. Apalagi dalam pelajaran Matematika yang oleh kebanyakan siswa dijadikan momok tersendiri, sehingga besar kemungkinan siswa untuk melakukan terkaan terhadap tes yang diberikan.

Karena adanya faktor terkaan yang ditimbulkan oleh tes bentuk pilihan ganda terutama dalam pelajaran Matematika, tentu hal tersebut memberikan masalah, yaitu hasil tes yang diperoleh peserta tes tidak menggambarkan kemampuan peserta tes yang sesungguhnya sehingga timbul ketidakpercayaan orang terhadap hasil tersebut. Selain itu juga tidak akan terdeteksi apakah siswa memahami tentang soal yang ada atau tidak. Oleh karenanya, untuk meminimalisir terkaan siswa pada tes pilihan ganda maka dibuatlah teknik penskoran untuk mengurangi tebakan. Teknik penskoran tersebut adalah penskoran pinalti. “Pemberian skor dengan pinalti (hukuman) dalam tes prestasi dikenal dengan istilah *Correction for guessing*”(Azwar, 2011 :114). Metoda penskoran dengan pinalti ini diyakini dapat meningkatkan reliabilitas dan validitas tes. Hal ini dikarenakan skor yang terkoreksi sudah memiliki tingkat estimasi

Dwi Putri Musdansi, 2014

PERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA SMA BERDASARKAN TEKNIK
PENSKORAN DAN UKURAN SAMPEL

yang lebih baik sebab meliputi pengukuran karakteristik adanya terkaan (Crocker dan Algina: 401). Selain penskoran dengan hukuman, upaya lain untuk meningkatkan reliabilitas dengan mengurangi efek terkaan adalah adanya tambahan nilai terhadap butir soal yang diragukan dengan cara mengosongkan jawaban tersebut, dikenal dengan *reward score*. Kedua teknik *Correction for guessing* tersebut belum atau tidak pernah digunakan untuk oleh para guru disekolah sebagai antisipasi agar siswa tidak melakukan tebakan. Karena guru belum tahu persis besar pengurangannya maupun kebermanfaatannya mengenai penggunaan teknik ini.

Selain itu, mengutip pendapat Nikto (1996:72-73) dalam Busnawir (2006:4) mengatakan bahwa “hal yang perlu diperhatikan dalam estimasi kestabilan koefisien reliabilitas adalah ukuran sampel, karena indeks reliabilitas dihitung berdasarkan sampel responden sehingga akan berfluktuasi berdasarkan besar sampel yang ditarik dari suatu populasi”. Namun seberapa besar perbedaan fluktuasi itu belum dapat ditentukan secara pasti. Azwar menyebutkan bahwa semakin heterogen suatu kelompok maka reliabilitas yang dihasilkan semakin tinggi. Sebaliknya, semakin homogen suatu kelompok maka semakin kecil reliabilitasnya. Heterogenitas ataupun homogenitas kelompok diperlihatkan oleh besar kecilnya varians distribusi skor subjek pada variabel yang diungkap oleh tes yang bersangkutan.

Dari uraian diatas dikemukakan bahwa jika menulis tes dalam bentuk tes pilihan ganda maka akan berpeluang untuk melakukan tebakan. Ada atau tidaknya tebakan tergantung pada intervensi yang diberikan pada tes, yaitu berupa petunjuk soal pada jenis teknik penskoran. Disisi lain reliabilitas juga dipengaruhi oleh ukuran sampel yang ditarik dari populasi, karena reliabilitas dihitung berdasarkan ukuran responden. Artinya reliabilitas akan berfluktuasi sesuai dengan jumlah sampel yang ditarik. Hal itu disebabkan karena adanya perbedaan sebaran skor yang menyebabkan jika semakin besar ukuran sampel maka akan diperoleh variasi skor yang lebih beragam daripada ukuran sampel yang kecil. Oleh sebab itu melalui penelitian ini, peneliti ingin mengkaji lebih jauh mengenai bagaimanakah

koefisien reliabilitas tes hasil belajar yang dihasilkan oleh teknik penskoran dan ukuran sampel.

B. IDENTIFIKASI MASALAH

Berdasarkan latar belakang masalah diatas, maka diperoleh masalah yang diidentifikasi sebagai berikut:

1. Tes bentuk pilihan ganda memberikan peluang besar kepada peserta tes untuk melakukan penebakan.
2. Bentuk tes pilihan ganda yang dihitung reliabilitasnya dapat berupa hasil belajar Matematika, Bahasa Indonesia, Sejarah, Bahasa Inggris dan lain sebagainya.
3. Penebakan berpengaruh terhadap reliabilitas
4. Reliabilitas dipengaruhi oleh berbagai faktor diantaranya ukuran sampel, teknik penskoran, kondisi subjek, kondisi ruangan, panjang tes dan lain sebagainya.
5. Teknik penskoran banyak jenisnya, diantaranya *number right score*, *reward score*, *punishment score*, *pembobotan nominal*, *confiden-weight*, dan sebagainya.

C. PEMBatasan MASALAH

Agar penelitian ini lebih efektif, efisien, terarah dan dapat dikaji lebih mendalam, maka diperlukan pembatasan masalah. Dalam penelitian ini, peneliti akan fokus pada reliabilitas tes pilihan ganda hasil belajar Matematika serta dua faktor yang diduga mempengaruhi reliabilitas yaitu:

1. Teknik penskoran yang terdiri dari *number-right score*, *reward score*, dan *punishment score*
2. Ukuran sampel yang dibatasi dengan ukuran sampel 30 dan 40.

D. RUMUSAN MASALAH

Berdasarkan identifikasi dan pembatasan masalah diatas maka dirumuskan permasalahan sebagai berikut:

Dwi Putri Musdansi, 2014

PERBANDINGAN RELIABILITAS TES HASIL BELAJAR MATEMATIKA SMA BERDASARKAN TEKNIK PENS KORAN DAN UKURAN SAMPEL

1. Apakah ada perbedaan reliabilitas tes pilihan ganda hasil belajar Matematika antara teknik penskoran *number-right score* dengan *reward score*?
2. Apakah ada perbedaan reliabilitas tes pilihan ganda hasil belajar Matematika antara teknik penskoran *number-right score* dengan *punishment score*?
3. Apakah ada perbedaan reliabilitas tes pilihan ganda hasil belajar Matematika kelompok *reward* dengan ukuran sampel 30 dan 40?
4. Apakah ada perbedaan reliabilitas tes pilihan ganda hasil belajar Matematika kelompok *punishment* dengan ukuran sampel 30 dan 40?

E. TUJUAN PENELITIAN

Tujuan penelitian ini adalah untuk menganalisis ada atau tidaknya:

1. Perbedaan reliabilitas tes pilihan ganda hasil belajar Matematika antara teknik penskoran *number-right score* dan *reward score*
2. Perbedaan reliabilitas tes pilihan ganda hasil belajar Matematika antara teknik penskoran *number-right score* dan *punishment score*
3. Perbedaan reliabilitas tes pilihan ganda hasil belajar Matematika kelompok *reward* dengan ukuran sampel 30 dan 40
4. Perbedaan reliabilitas tes pilihan ganda hasil belajar Matematika kelompok *punishment* dengan ukuran sampel 30 dan 40

F. MANFAAT PENELITIAN

Dengan adanya penelitian ini maka akan memberikan manfaat:

1. Memberikan informasi tentang koefisien reliabilitas yang akan diperoleh dengan teknik penskoran *punishment score* dan *reward score* jika melakukan penebakan.
2. Mendapatkan bukti empirik bahwa ukuran sampel yang berbeda dengan teknik penskoran sama akan memberikan kontribusi yang berbeda pula terhadap kestabilan reliabilitas.
3. Memberikan informasi kepada evaluator pendidikan tentang teknik penskoran yang baik digunakan untuk memperoleh koefisien reliabilitas.