

BAB I

PENDAHULUAN

A. Latar Belakang

Peraturan Menteri Pendidikan dan Kebudayaan Nomor 59 Tahun 2011 tentang “Kriteria kelulusan peserta didik dari satuan pendidikan dan penyelenggaraan ujian sekolah/madrasah dan ujian nasional” menguraikan bahwa Ujian Nasional (UN) bertujuan untuk menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu dalam kelompok mata pelajaran ilmu pengetahuan dan teknologi. Tujuan Ujian Nasional (UN) tersebut secara eksplisit mengisyaratkan bahwa Ujian Nasional (UN) merupakan penentu kelulusan bagi setiap peserta didik pada satuan pendidikan tertentu. Dengan demikian, bagi satuan pendidikan Ujian Nasional (UN) merupakan beban berat karena menyangkut kredibilitas dan cerminan kualitas satuan pendidikan. Sehubungan dengan hal tersebut, maka setiap satuan pendidikan berupaya sekuat tenaga untuk mensukseskan pelaksanaan Ujian Nasional (UN) sehingga memperoleh hasil yang memuaskan.

Ujian Nasional merupakan salah satu perjuangan akhir dari siswa untuk mengakhiri setiap jenjang pendidikan tertentu, oleh karena itu diperlukan persiapan yang matang untuk menghadapinya, apalagi nilai Ujian Nasional sangat menentukan besar kecilnya peluang dalam memasuki jenjang pendidikan selanjutnya. Ujian ini bertujuan untuk menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran Matematika, Ilmu Pengetahuan Alam (IPA) Bahasa Indonesia, Bahasa Inggris serta mendorong tercapainya target wajib belajar pendidikan dasar yang bermutu (Depdiknas, 2007:6).

Dalam perkembangan ujian nasional, soal UN sejak ujian ini diperkenalkan sudah didesain lebih dari satu paket. Selama beberapa tahun soal UN menggunakan dua paket, yaitu paket A dan paket B dengan tujuan untuk memperkecil angka kecurangan. Pada pelaksanaan UN 2012 soal untuk tingkat SMP/MTs dibuat sekitar lima paket yakni paket A, B, C, D dan E dimana kelima paket naskah ujian untuk tiap mata pelajaran yang akan dibagikan secara acak

Ariani Arsad, 2014

KOMPARASI SKOR HASIL PENYETARAAN UN MATEMATIKA SMP/MTs ANTARA METODE LINIER DAN METODE EKUIPERSENTIL DI KOTA AMBON

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

pada para peserta UN. Mulai tahun 2013, Kemdikbud mengeluarkan format baru untuk naskah soal dan lembar jawaban UN 2013. Barcode digunakan pula dalam pelaksanaan UN tertulis bagi peserta didik jenjang pendidikan SMP/MTs dan SMA/MA/SMK sebagai sesuatu khas terbaru. Selain penggunaan barcode dan penyatuan naskah soal dengan lembar jawaban UN, Kemdikbud juga menyiapkan 20 variasi soal pada UN tahun ini sehingga siswa bisa berkonsentrasi pada soalnya masing-masing tanpa perlu melihat pekerjaan temannya. Penambahan jumlah paket soal ini merupakan cara pemerintah untuk menekan angka kecurangan dalam pelaksanaan ujian nasional. Maklum walaupun selama ini jumlah paket soal sudah cukup banyak, yakni lima, tetapi kebocoran masih terjadi di sejumlah daerah.

Meski naskah soal UN memiliki banyak paket, pihak penyelenggara tetap meyakinkan bahwa tingkat kesulitannya tetap sama antara satu soal dan soal yang lainnya. Komposisi soalnya sendiri terdiri dari 70 persen soal sedang, 20 persen soal sulit, dan 10 persen soal mudah.

Instrumen penilaian yang digunakan oleh pemerintah dalam bentuk Ujian Nasional (UN) menurut Peraturan Menteri Pendidikan Nasional RI Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan memenuhi persyaratan substansi, konstruksi, bahasa, dan memiliki bukti validitas empirik serta menghasilkan skor yang dapat diperbandingkan antar sekolah, antar daerah, dan antar tahun (Depdiknas, 2007:9). Keterbandingan skor antar sekolah, kabupaten/kota, provinsi, dan antar tahun dapat diperoleh jika semua peserta tes mengerjakan soal-soal (paket tes) yang sama. Perbedaan skor antar mereka menunjukkan perbedaan tingkat kemampuannya.

Pada program pengujian, khususnya dalam skala besar, penyusunan beberapa perangkat tes yang setara adalah salah satu kegiatan penting karena salah satu tugasnya adalah menjaga keamanan perangkat tes tersebut. Pada taraf tertentu kesetaraan beberapa perangkat tes dapat dilaksanakan pada saat mengembangkan tes itu sendiri, tetapi biasanya bervariasi antara perangkat tes yang satu dengan perangkat tes lainnya terutama dalam hal tingkat kesukaran. Hal ini dapat diatasi

dengan melakukan kegiatan penyetaraan skor antar perangkat tes dengan cara yang tepat dan benar.

Walaupun sampai taraf tertentu kesetaraan beberapa tes dapat diupayakan pada saat menyusun tes-tes itu sendiri, tetapi pada umumnya variasi taraf sukar antar tes tetap terjadi (Swediati,1997:1). Selain itu, penyetaraan tes perlu dilakukan mengingat bahwa menyusun tes yang benar-benar paralel tidaklah mudah. Jadi secara empirik membuat dua tes yang sama, tidak pernah secara sempurna paralel, terandalkan atau unidimensional, sehingga skor-skor yang dihasilkannya tidak dapat dibandingkan (Gronlund, 1985:169). Jika hasil tes digunakan untuk menentukan kenaikan kelas atau penjurusan program, tentunya hal itu menjadi tidak adil karena tidak dilakukan ekuivalensi skor untuk tes yang berbeda tersebut. Oleh karena itu, penting dilakukan penyesuaian skor-skor tersebut sehingga peserta tes yang berbeda, menggunakan tes yang berbeda tersebut dapat dibandingkan.

Ujian nasional menggunakan tes objektif yang terdiri dari butir-butir pilhan ganda yang secara luas dikenal dan digunakan. Soal pilihan ganda terdiri dari suatu masalah yang dinyatakan sebagai pertanyaan langsung atau suatu pernyataan tidak lengkap dan suatu daftar pilihan solusi yang merupakan alternatif atau pilihan jawaban. Keunggulan tes jenis ini diantaranya adalah dapat di skor dengan mudah dan cepat serta ruang lingkup bahan yang ditanyakan sangat luas. Pada soal UN mata pelajaran Matematika SMP/MTs terdapat soal inti (*anchor items*) sebanyak lima butir dari 40 butir soal atau sebanyak 12,5 persen dalam 20 paket soal. Tujuan *anchor items* ini diperlukan agar antar paket memiliki keterkaitan yang sama. Dengan menggunakan *anchor items* dapat dihilangkan bias penyetaraan yang ditimbulkan perbedaan kemampuan dua kelompok atau lebih. Penempatan *anchor items* dalam paket soal UN diletakkan pada nomor urutan yang sama.

Pada situasi demikian, terdapat beberapa paket tes yang digunakan untuk mengukur variabel yang sama, namun skor hasil tes atau skor mentah yang dihasilkan tes tidak dapat langsung diperbandingkan walaupun tes-tes tersebut dibuat dengan kisi-kisi yang sama. Alasannya, skor mentah yang dihasilkan tes

sering tidak memiliki titik nol yang sama atau ukuran yang sama. Selanjutnya, untuk menghasilkan skor tes yang dapat diperbandingkan dari beberapa paket tes, perlu disusun skor tes yang memiliki skala tunggal (*common scale*).

Masalah tersebut dapat diatasi dengan melakukan penyetaraan *skor* yang diperoleh dari peserta yang mengambil tes-tes tersebut yang dikenal sebagai metode penyetaraan (*equating*). Dengan kata lain penyetaraan adalah proses untuk menentukan hubungan antara skala skor dari dua tes atau lebih tes agar skor-skor tersebut diperlakukan secara adil. Kegiatan penyetaraan tes dapat dilakukan dengan mengembangkan konversi suatu sistem unit tes ke sistem unit tes yang lain sehingga setelah dikonversi skor yang berasal dari dua perangkat tes menjadi setara dan dapat dipertukarkan.

Penyetaraan skor dapat digunakan untuk memantau mutu pendidikan. Melalui penyetaraan skor dapat dilakukan pengukuran perkembangan siswa, komparasi hasil tes yang berasal dari perangkat tes yang berbeda. Penyetaraan ini perlu dilakukan agar jangan sampai seseorang dirugikan atau diuntungkan karena kebetulan mengambil perangkat tes yang lebih sukar atau lebih mudah.

Di negara maju seperti Amerika Serikat, penyetaraan tes sangat diperlukan oleh lembaga-lembaga yang menangani tes khusus yaitu *test battery*, misalnya *Iowa Test of Basic Skills* (ITBS) dan *Iowa Test of Educational Development* (ITED). Tes tersebut digunakan untuk mengukur perkembangan kemampuan atau prestasi siswa pada jenjang dan kelas tertentu, terdiri dari beberapa tingkat yang terkait dengan umur dan kelas (Hieronymus, et al, dalam Kartono, 2008). Kelompok siswa yang mempunyai umur dan kelas yang sama, dalam perkembangannya, terdapat siswa yang cepat belajarnya dan ada yang lambat belajarnya.

Khusus untuk siswa yang cepat belajarnya, siswa tersebut cocok diberi tes tingkat lanjut, tetapi untuk anak yang lambat belajarnya, siswa tersebut tidak cocok diberi tes tingkat lanjut. Tidak perlu diadakan tes tingkat lanjut untuk siswa yang lambat belajar, karena tes yang terlalu sulit dapat menyebabkan siswa frustrasi, cukup dilakukan penyesuaian skor pada tes tingkat di bawahnya untuk mengetahui kemampuan pada tes tingkat lanjut.

Ariani Arsad, 2014

KOMPARASI SKOR HASIL PENYETARAAN UN MATEMATIKA SMP/MTs ANTARA METODE LINIER DAN METODE EKUIPERSENTIL DI KOTA AMBON

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

Dengan memfasilitasi siswa-siswa dengan kemampuan yang berbeda, akan terdapat beberapa paket tes dengan tingkat yang berbeda, skala skor hasil tes berbeda. Perlu diadakan penyesuaian skor yang berasal dari hasil tes dengan tingkat yang berbeda pada skala yang sama, sehingga hasil tes dapat diperbandingkan. Dalam hal ini, untuk menyesuaikan skor-skor dari beberapa paket tes dengan tingkat yang berbeda agar skor tersebut dapat dibandingkan juga diperlukan penyetaraan tes.

Penyetaraan atau *equating* merupakan suatu prosedur secara empiris dalam rangka menyetarakan skor dari perangkat tes yang satu ke perangkat tes yang lain. Melalui proses penyetaraan yang dilakukan dengan benar, dapat diukur perkembangan dan memetakan berbagai kecenderungan, serta mendapatkan dan menggabungkan informasi dan perangkat tes yang berbeda. Hasil penyetaraan memungkinkan kita melakukan perbandingan atau konvensi secara langsung hasil-hasil tes individu yang mengikuti perangkat tes yang berbeda.

Terdapat dua cara penyamaan skor pada teori klasik yaitu penyamaan cara linier dan cara ekuipersentil. Penyetaraan secara linier melalui kesamaan nilai baku dapat dilakukan pada masing-masing rancangan penyetaraan dari rancangan A sampai rancangan D. Pada penyetaraan, gandengan internal atau eksternal adalah sama saja sehingga tidak dibedakan. Penyetaraan secara ekuipersentil adalah penyetaraan pada peringkat persentil yang sama. Peringkat persentil suatu skor responden adalah letak skor itu di dalam kumpulan skor responden yang menyatakan bahwa skor itu lebih tinggi dari sekian persen skor responden lainnya. Pada umumnya penyetaraan ekuipersentil adalah nonlinier. Hal ini berbeda dengan penyetaraan linier yang adalah linier. Menurut Hambleton, Swaminathan, dan Roger (1991:124) asumsi penyetaraan dengan cara linier adalah kedua skor tes distribusinya berbeda. Distribusi tersebut terkait dengan rerata dan simpangan bakunya.

Pengukuran dalam dunia pendidikan mengenal dua macam kekeliruan, yaitu kekeliruan acak atau kekeliruan sampel (*sampling error*) dan kekeliruan sistematis (*systematic error*). Kekeliruan sampel adalah perbedaan antara keadaan sebenarnya yang ada pada populasi (*true score*). Hal ini disebabkan oleh karena

hasil ukuran pada sampel tersebut hanya merupakan salah satu dari sekian banyak kemungkinan hasil pengukuran yang dapat dicuplik berulang-ulang dari suatu populasi. Kekeliruan sistematik menyebabkan munculnya hasil penyetaraan yang menjurus ke arah yang sama dengan jumlah tertentu. Misalnya, hasil penyetaraan selalu lebih tinggi dalam jumlah tertentu dibandingkan dengan skor sesungguhnya. Kekeliruan semacam ini terjadi karena adanya faktor kesalahan alat ukur (*instrumental error*) dan faktor kesalahan manusia (*human error*). Penelitian lebih lanjut yang disarankan oleh Livingston, Doran dan Wright (1990) terhadap kombinasi terbaik penarikan sampel dan metode-metode penyetaraan adalah terkait dengan banyaknya metode penyetaraan dan variansinya serta metode sampel yang dimungkinkan.

Dengan demikian, uraian di atas memberi gambaran perlu adanya penelitian tentang penyetaraan linier dan penyetaraan ekuipersentil terhadap perangkat tes dalam hal ini soal UN sebagai salah satu upaya pengembangan alat ukur dalam mengevaluasi hasil belajar dalam berbagai bidang ilmu. Matematika sebagai suatu bidang ilmu, diajarkan mulai dari tingkat SD sampai perguruan tinggi. Peran dan pentingnya matematika dalam berbagai ilmu dan juga dalam kehidupan sehari-hari cukup besar dan tidak dapat diabaikan. Di lain pihak prestasi Indonesia dibidang matematika masih tertinggal jauh, dimana pada PISA 2012 Indonesia masih berada pada sepuluh besar peringkat terbawah. Sehubungan dengan masalah pada penelitian ini, maka soal UN matematika yang digunakan untuk penelitian ini adalah tingkat SMP/MTs.

B. Identifikasi Masalah

Berdasarkan latar belakang masalah maka dapat diidentifikasi permasalahan yang muncul, sebagai berikut :

1. Proporsi *anchor items* dalam soal UN sebesar 12,5 persen dalam 20 paket atau sebanyak 5 butir soal dalam 40 butir soal.
2. Skor yang diperoleh dari dua perangkat yang berbeda tidak dapat langsung dibandingkan.

3. Terdapat berbagai metode penyetaraan tes, baik yang tercakup dalam teori tes klasik maupun tes modern (teori respon butir). Pada teori klasik terdapat metode penyetaraan linier dan metode penyetaraan ekuipersentil. Dalam teori respon butir, untuk jawaban dikotomi, terdapat metode penyetaraan regresi, rerata dan sigma, tegar rerata dan sigma, dan lengkung karakteristik.
4. Perangkat tes yang disamakan ukurannya harus memenuhi syarat, yaitu mengukur ciri yang sama pada peserta yang setara, reliabilitas yang sama dan merata, dan memiliki taraf kesukaran yang merata.
5. Kekeliruan acak dan kekeliruan sistemik juga mempengaruhi skor hasil penyetaraan.

C. Batasan Masalah

Dari uraian yang dikemukakan pada identifikasi masalah di atas, penelitian ini dilakukan dengan pendekatan teori tes klasik dengan membatasi masalah pada karakteristik perangkat tes, konversi skor penyetaraan dan skor hasil penyetaraan dengan menggunakan metode linier dan metode ekuipersentil pada dua paket soal UN Matematika SMP/MTs tahun pelajaran 2012/2013 di Kota Ambon yang mempunyai kisi-kisi yang sama dan *anchor items* yang sama sebesar 12,5%.

Terdapat tiga desain dasar yang dapat digunakan untuk pengumpulan data dalam melakukan penyetaraan tes. Desain yang dimaksud adalah desain grup tunggal, desain grup ekuivalen, dan desain tes jangkar. Pemilihan desain yang digunakan dalam penyetaraan sangat tergantung pada situasi pelaksanaannya. Dalam penelitian ini dibatasi pada desain tes jangkar karena merupakan desain yang digunakan dalam UN.

D. Rumusan Masalah

Berdasarkan uraian pada latar belakang di atas, maka rumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimanakah karakteristik perangkat tes dari dua instrumen yang digunakan berdasarkan teori tes klasik?

2. Bagaimanakah konversi skor hasil penyetaraan dengan menggunakan metode penyetaraan linear dan metode penyetaraan ekuipersentil pada *anchor items* 12,5%?
3. Apakah terdapat perbedaan hasil penyetaraan bila skor hasil UN pada dua perangkat tes yang berbeda disetarakan menggunakan metode penyetaraan linier dan metode penyetaraan ekuipersentil pada *anchor items* 12,5%?

E. Tujuan Penelitian

Sesuai dengan perumusan masalah di atas, maka tujuan penelitian ini adalah:

1. Menganalisis karakteristik dari dua instrumen yang digunakan berdasarkan teori tes klasik.
2. Mendeskripsikan konversi skor hasil penyetaraan dengan menggunakan metode linear dan metode ekuipersentil pada *anchor items* 12,5%.
3. Membandingkan skor hasil penyetaraan bila skor hasil UN pada dua perangkat tes yang berbeda disetarakan dengan menggunakan metode linear dan metode ekuipersentil pada *anchor items* 12,5%.

F. Manfaat Penelitian

Kegunaan hasil penelitian ini dapat memberikan manfaat baik secara teoritis maupun praktis. Secara teoritis, hasil penelitian ini memberikan kontribusi tentang pengukuran dan penilaian pendidikan, khususnya tentang bagaimana memperbandingkan skor ujian nasional. Hasil penelitian ini juga dapat dimanfaatkan sebagai masukan awal untuk melakukan penelitian lanjutan.

Secara praktis, penyetaraan diharapkan dapat digunakan sebagai bagian dalam kurikulum diklat oleh lembaga pelaksana diklat bagi tenaga pendidik dan kependidikan. Dengan diketahuinya perbandingan skor ujian nasional dapat digunakan untuk memperbaiki dan meningkatkan mutu pendidikan di daerah.