

THE ACCURACY OF CONCORDANCE METHOD

Dwining Bintarawati

SMAN 1 Serang Baru

Perumkota Serang Baru Blok E, Bekasi

dwining286@gmail.com

Abstract

This research was aimed to examine the accuracy of concordance method on linear method, linear parallel method and equipercentile method. The accuracy of those three concordance methods is revealed through the average score of the root mean square deviation (RMSD). This research used two sets of test instrument, namely the Final Examination of Chemistry, in which each was developed by SMAN 1 Tambun Selatan and SMAN 1 Cikarang Utara. Both of these two sets of test instrument were developed based on similar construct (similar standards of competence and basic competences). The method used was an experimental method. Random Sampling with replacement technique was used. Research hypotheses were tested by using analysis of one way variance (ANOVA). The study concluded that: (1) there was no difference in the average score of the RMSD between the linear methods, parallel linear methods, and equipercentile method when viewed statistically, and (2) viewed empirically, the difference of the average score of the RMSD of the three concordance methods was very small. The average score of RMSD when perceived by the parallel linear method is 3,0964, when perceived by the linear method is 3,1644, and when perceived by the equipercentile method is 3,2943. This indicated that the parallel linear method was little more accurate compared to both linear method and equipercentile method.

Keywords: *accuracy, concordance method, linear method, parallel linear method, equipercentile method*

AKURASI METODE *CONCORDANCE*

Dwining Bintarawati

SMAN 1 Serang Baru

Perumkota Serang Baru Blok E, Bekasi

dwining286@gmail.com

Abstrak

Penelitian ini bertujuan untuk meneliti akurasi metode *concordance* antara metode linier, metode paralel linier, dan metode ekipersentil. Akurasi ketiga metode *concordance* tersebut dilihat melalui nilai rerata dari deviasi akar kuadrat rata-rata (RMSD). Penelitian ini menggunakan dua perangkat tes, yaitu perangkat tes Ulangan Akhir Semester mata pelajaran kimia yang dikembangkan oleh dua sekolah yang berbeda, yaitu SMAN 1 Tambun Selatan dan SMAN 1 Cikarang Utara. Kedua perangkat tes tersebut disusun berdasarkan kisi-kisi yang sama (standar kompetensi dan kompetensi dasar yang sama). Metode yang digunakan adalah eksperimen. Pengambilan sampel dilakukan melalui metode acak berulang dengan pengembalian. Hipotesis penelitian diuji dengan analisis varian (ANOVA) satu jalan. Hasil penelitian disimpulkan bahwa: (1) secara statistika tidak terdapat perbedaan nilai rerata RMSD antara metode linier, metode paralel linier, dan metode ekipersentil, dan (2) secara empirik, perbedaan nilai rerata RMSD dari ketiga metode *concordance* tersebut sangat kecil. Pada metode paralel linier diperoleh nilai rerata RMSD sebesar 3,0964, sedangkan dengan metode linier diperoleh nilai rerata RMSD sebesar 3,1644, dan dengan metode ekipersentil diperoleh nilai rerata RMSD sebesar 3,2943. Hal ini menunjukkan bahwa metode paralel linier sedikit lebih akurat dibandingkan metode linier dan metode ekipersentil.

Kata kunci: akurasi, metode *concordance*, metode linier, metode paralel linier, metode ekipersentil

PENDAHULUAN

Pendidikan nasional yang berdasarkan Pancasila dan Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 berfungsi mengembangkan kemampuan dan membentuk watak serta peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa, bertujuan untuk mengembangkan potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan menjadi warga negara yang demokratis serta bertanggung jawab. Untuk mengemban fungsi tersebut pemerintah menyelenggarakan suatu sistem pendidikan nasional sebagaimana tercantum Undang-Undang Nomor 20 tahun 2003 serta Peraturan Pemerintah Nomor 19 tahun 2005 tentang Standar Nasional Pendidikan. Hal tersebut berimplikasi terhadap sistem dan penyelenggaraan pendidikan termasuk pengembangan dan pelaksanaan

kurikulum. Kebijakan pemerintah tersebut mengamanatkan kepada setiap satuan pendidikan dasar dan menengah untuk mengembangkan Kurikulum Tingkat Satuan Pendidikan (KTSP).

Pengembangan Kurikulum Tingkat Satuan Pendidikan (KTSP) mengacu pada Standar Nasional Pendidikan (SNP) yang merupakan kriteria minimal tentang sistem pendidikan di seluruh wilayah hukum Negara Kesatuan Republik Indonesia. Ruang lingkup SNP meliputi standar: (1) isi, (2) proses, (3) kompetensi lulusan, (4) pendidik dan tenaga kependidikan, (5) sarana dan prasarana, (6) pengelolaan, (7) pembiayaan, dan (8) penilaian pendidikan.

Kurikulum adalah seperangkat rencana dan pengaturan mengenai tujuan, isi dan bahan pelajaran serta cara yang digunakan sebagai pedoman penyelenggaraan kegiatan pembelajaran untuk mencapai tujuan pendidikan tertentu. Kurikulum tingkat satuan pendidikan jenjang pendidikan dasar dan menengah dikembangkan oleh sekolah dan komite sekolah berpedoman pada standar kompetensi lulusan dan standar isi serta panduan penyusunan kurikulum yang dibuat oleh BSNP. Struktur kurikulum merupakan pola dan susunan mata pelajaran yang harus ditempuh oleh peserta didik dalam kegiatan pembelajaran. Kedalaman muatan kurikulum pada setiap mata pelajaran pada setiap satuan pendidikan dituangkan dalam kompetensi yang harus dikuasai peserta didik sesuai dengan beban belajar yang tercantum dalam struktur kurikulum. Kompetensi yang dimaksud terdiri atas standar kompetensi dan kompetensi dasar yang dikembangkan berdasarkan standar kompetensi lulusan.

Kompetensi adalah pengetahuan, keterampilan, sikap, dan nilai-nilai yang direfleksikan dalam kebiasaan berpikir dan bertindak. Peserta didik dianggap kompeten apabila dia memiliki pengetahuan, keterampilan, sikap, dan nilai-nilai untuk melakukan sesuatu setelah mengikuti proses pembelajaran. Untuk mengetahui pencapaian kompetensi peserta didik dalam kegiatan pembelajaran sesuai dalam muatan kurikulum dilakukan melalui kegiatan pengukuran. Menurut Wiersma (1990: 8), pengukuran secara teknis adalah pemberian angka pada obyek atau peristiwa menurut aturan-aturan yang memberikan arti kuantitatif secara numerik.

Pengukuran di bidang pendidikan memegang peranan yang sangat penting. Data hasil pengukuran dalam bidang pendidikan memiliki arti penting baik bagi sekolah atau lembaga pendidikan, guru, maupun bagi siswa dan orang tua siswa atau masyarakat. Bagi guru misalnya hasil pengukuran berfungsi untuk membandingkan tingkat kemampuan siswa dengan siswa-siswa lain dalam kelompok yang diajarnya. Menurut Djaali dan Muljono (2004: 6-7), di sekolah pengukuran dilakukan guru untuk menaksir prestasi siswa. Banyak alat atau instrumen yang dapat digunakan dalam kegiatan pengukuran. Pengukuran prestasi siswa di sekolah pada umumnya menggunakan instrumen tes yang disebut tes hasil belajar.

Istilah tes tidak hanya populer di lingkungan sekolah, tetapi juga di luar sekolah bahkan masyarakat umum. Di sekolah, tes sering digunakan untuk

mengukur prestasi belajar peserta didik dalam bidang kognitif, seperti pengetahuan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi. Penggunaan tes dalam dunia pendidikan sudah dikenal sejak dahulu kala, sejak orang mengenal pendidikan itu sendiri. Menurut Cronbach (1984: 26), tes adalah suatu prosedur yang sistematis untuk mengamati dan mendeskripsikan karakteristik seseorang dengan menggunakan skala numerik atau sistem kategori. Tes yang digunakan untuk mengukur prestasi belajar siswa, haruslah memenuhi syarat-syarat tes yang baik agar dapat berfungsi secara tepat dan akurat. Syarat-syarat tes yang baik menurut Suryabrata (2006: 303-306), terdiri atas enam faktor yaitu reliabel, valid, objektif, diskriminatif, komprehensif, dan mudah digunakan.

Sekolah sebagai suatu institusi pengelola pendidikan senantiasa berurusan dengan program tes. Terutama guru yang biasanya terlibat secara langsung dalam penyusunan tes hasil belajar. Tes buatan guru ini biasanya digunakan untuk ulangan harian, formatif, dan ulangan umum (sumatif). Tes buatan guru ini dimaksudkan untuk mengukur tingkat penguasaan siswa terhadap materi pelajaran yang sudah disampaikan. Untuk itu, guru harus membuat soal secara logis dan rasional mengenai pokok-pokok materi yang seharusnya ditanyakan sebagai bahan pengetahuan penting untuk diketahui dan dipahami oleh peserta didiknya.

Tes dibuat berdasarkan kurikulum yang digunakan di sekolah. Tes memiliki kesesuaian dengan materi pembelajaran. Oleh karena itu, sebelum dilakukan pembuatan tes, terlebih dahulu dilakukan penyusunan kisi-kisi soal, agar materi penilaian betul-betul representatif dan relevan dengan materi pembelajaran yang sudah diberikan oleh guru kepada peserta didik. Menurut Arifin (2009: 93), kisi-kisi yang baik akan memperoleh perangkat soal yang relatif sama sekalipun penulis soalnya berbeda. Dalam konteks penilaian hasil belajar, kisi-kisi soal disusun berdasarkan silabus mata pelajaran. Jadi, guru harus melakukan analisis silabus terlebih dahulu sebelum menyusun kisi-kisi soal.

Banyaknya terbitan buku-buku berisi kumpulan soal, menyebabkan gurupun cenderung dengan mudah memilih atau mengadaptasi soal-soal tersebut untuk digunakan dalam tes formatif maupun tes sumatif di sekolah. Hal itu biasanya dilakukan untuk mempermudah dan mengefisienkan waktu dalam pembuatan tes. Padahal, tingkat kesukaran soal-soal tersebut belum tentu sesuai dengan sifat-sifat dan karakteristik peserta didiknya.

Arifin (2009: 119-120) mengatakan bahwa kualitas tes buatan guru atau tingkat kesahihan dan keandalannya juga belum menjamin keobjektifannya, sebab hanya diberikan kepada sekelompok peserta didik, kelas dan sekolah tertentu saja. Jadi, tes masih bersifat sektoral, karena belum diujicobakan kepada sekelompok besar peserta tes sehingga pengukurannya masih belum meyakinkan. Selain itu, tes buatan guru juga bersifat temporer, artinya hanya berlaku pada saat dan situasi tertentu. Pada kesempatan lain belum tentu tes tersebut dapat digunakan, karena mungkin berubah, baik bentuk butirnya maupun kapasitas peserta didiknya.

Guru dalam menyusun dan melaksanakan tes untuk kepentingan penilaian hasil belajar cenderung lemah dalam menafsirkan hasil pengukuran dari dua atau beberapa perangkat tes yang berbeda tetapi mengukur hal yang sama. Mereka sering memperlakukan hal tersebut sama, tanpa memperhatikan tingkat kesukaran tes yang digunakan. Tindakan demikian dapat dikatakan menyimpang dari teori pengukuran. Menurut Hambleton dan Swaminathan (1985: 123) sebenarnya tidak pernah ditemukan dua perangkat tes dengan butir soal berbeda yang sama tingkat kesukarannya. Selain itu, walaupun dua perangkat tes disusun berdasarkan kisi-kisi yang sama, jarang sekali bahkan hampir tidak ditemukan dua perangkat tes yang benar-benar setara dalam sebaran serta peringkat indeks kesukaran. Dengan demikian, diperlukan untuk *linking assessment* dari dua perangkat tes yang berbeda.

Pada akhir proses pembelajaran setiap semester, setiap sekolah melaksanakan tes sumatif, atau yang lebih dikenal dengan istilah ulangan umum atau ulangan akhir semester (UAS). Tes sumatif adalah tes hasil belajar yang dilaksanakan setelah sekumpulan materi pelajaran atau satuan program pengajaran selesai diberikan. Djaali dan Muljono (2004: 13) menambahkan bahwa tujuan dilaksanakannya tes sumatif adalah untuk menentukan nilai yang menjadi lambang keberhasilan siswa setelah mereka menempuh proses pembelajaran dalam jangka waktu tertentu. Hasil tes sumatif berguna untuk: (a) menentukan kedudukan atau rangking masing-masing siswa di dalam kelompoknya, (b) menentukan dapat atau tidaknya siswa melanjutkan program pembelajaran berikutnya, dan (c) menginformasikan kemajuan siswa untuk disampaikan kepada pihak lain seperti orang tua, sekolah, masyarakat, dan lapangan kerja.

Pelaksanaan ulangan akhir semester yang banyak dilakukan pada sekolah-sekolah menengah atas di kabupaten Bekasi dengan proses penyusunan soal ulangan akhir semester secara mandiri di sekolah masing-masing. Soal ulangan akhir semester ini dibuat oleh guru mata pelajaran berdasarkan Musyawarah Guru Mata Pelajaran (MGMP) di sekolah tersebut. Hal ini tentunya, mengakibatkan karakteristik tes yang berbeda pada perangkat tes yang disusun oleh sekolah satu dengan sekolah lainnya.

Ulangan akhir semester dilaksanakan dengan tujuan yang sama, yakni mengukur kompetensi siswa di akhir pembelajaran, dan dikembangkan dari standar kompetensi dan kompetensi dasar pokok yang sama yang terdapat dalam kurikulum dari setiap mata pelajaran. Terkait dengan hal ini, tes-tes yang dibuat diharapkan mengukur konten yang sama, dan memenuhi syarat tes yang baik, ditinjau dari karakteristik butir-butir penyusunannya. Mengingat banyak sekolah di kabupaten Bekasi menyiapkan sendiri perangkat ujiannya, mengakibatkan butir-butir penyusunannya mempunyai tingkat kesulitan yang berbeda-beda pula. Sehubungan dengan hal ini, perlu dilakukan *linking assessment* untuk dapat terjadi adanya pembakuan tingkat kemampuan yang

peroleh siswa, dan mengetahui pembakuan pencapaian pembelajaran di sekolah-sekolah.

Menurut Dorans (2004: 227), *linking* antartest dapat dikelompokkan menjadi 3, yakni *equating*, *concordance*, dan prediksi (*prediction*). Ketiga hubungan tersebut dibedakan berdasarkan konstruk tes dan distribusinya. Jika tes-tes tersebut secara statistik dan konseptual dapat saling menggantikan, maka hubungannya dapat diketahui dengan *equating*, jika sama distribusinya (mengukur konstruk yang sama) maka hubungannya dapat diketahui dengan *concordance*, dan jika kondisi untuk penyetaraan dan *concordance* tidak terpenuhi, digunakan prediksi skor harapan.

Yin (2004: 276) menambahkan bahwa ada dua pendekatan yang dapat ditempuh untuk melakukan *concordance* pada dua perangkat tes, yaitu dengan pendekatan teori tes klasik dan teori respons butir. Berdasarkan teori tes klasik, ada 3 metode yang dapat digunakan, yakni metode linier, metode paralel linier, dan metode ekipersentil. Lebih lanjut, menurut Hambleton dan Swaminathan (1985: 129), dengan menggunakan teori respons butir, ada 4 metode yang dapat ditempuh, yakni metode regresi, metode rerata dan sigma, metode tegar dan sigma, dan metode kurva karakteristik.

Pendekatan pada penelitian ini digunakan metode *concordance* yang termasuk pendekatan teori tes klasik. Dalam metode linier, digunakan dua macam tes X dan tes Y serta dua kelompok peserta yakni kelompok K1 dan K2. Dalam hal ini, kelompok K1 mengerjakan tes X serta kelompok K2 mengerjakan tes Y. Hasil tes tersebut menghasilkan rerata dan simpangan baku, μ_x dan σ_x serta μ_y dan σ_y . Selanjutnya dalam metode linier ini dilakukan penyamaan distribusi probabilitas di antara skor X dan skor Y itu. Menurut Naga (1992: 353), dengan distribusi probabilitas yang sama, rerata dan simpangan baku yang berbeda dari tes tersebut dapat digunakan untuk mentransformasi nilai dari satu distribusi probabilitas ke distribusi probabilitas berikutnya. Dalam metode ekipersentil, persentil skor dibagi pada dua atau lebih tes serta memadakkannya menurut persentil yang sama.

Kolen dan Brennan (1995: 215) menambahkan bahwa keakuratan metode-metode penyetaraan dapat diketahui melalui kesalahan pengukuran standarnya. Kesalahan pengukuran pada suatu kasus dalam hal ini *concordance* dinyatakan sebagai selisih antara hasil *concordance* sebenarnya dibandingkan dengan nilai harapannya. Karena kedua tes memiliki skala yang sama, maka nilai harapan merupakan nilai asli hasil pengamatan. Untuk menentukan keakuratan metode, kriteria metode diindikasikan dengan nilai RMSD (*Root Mean Square Deviation*). Menurut Livingston (1993: 34), rata-rata RMSD yang kecil menunjukkan akurasi yang lebih tinggi daripada rata-rata RMSD yang besar.

Penelitian bertujuan untuk mengetahui *linking assessment* pada dua perangkat tes ulangan akhir semester (UAS) kimia kelas XII-IPA SMA di kabupaten Bekasi dengan menggunakan metode *concordance* serta membandingkan

ketepatan ketiga metode *concordance* yaitu metode linier, metode paralel linier, dan metode ekipersentil dengan pendekatan teori tes klasik.

METODE PENELITIAN

Pendekatan dalam penelitian ini menggunakan pendekatan kuantitatif. Variabel bebas dalam penelitian ini adalah metode *concordance* yang terdiri atas metode linier, metode paralel linier, dan metode ekipersentil sedangkan variabel terikatnya adalah RMSD.

Populasi terjangkau dalam penelitian ini adalah siswa kelas XII IPA SMA Negeri 1 Tambun Selatan dan siswa kelas XII IPA SMA Negeri 1 Cikarang Utara. Pengambilan sampel dilakukan dengan metode acak berulang dengan pengembalian (*random sampling with replacement*). Instrumen dalam penelitian ini menggunakan dua instrumen tes ulangan akhir semester (UAS) mata pelajaran kimia yang dikembangkan oleh dua sekolah yang berbeda (SMAN 1 Tambun Selatan dan SMAN 1 Cikarang Utara) berdasarkan kisi-kisi yang sama. Dari hasil ujicoba diambil 30 butir soal valid dari 35 soal ujicoba dari dua instrumen tes kimia tersebut. Reliabilitas instrumen tes kimia I memiliki koefisien 0,721, sedangkan reliabilitas instrumen tes kimia II memiliki koefisien 0,752, sehingga kedua instrumen tes kimia tersebut dinyatakan reliabel. Teknis analisis data meliputi analisis deskriptif, uji persyaratan analisis, dan analisis inferensial. Sebelum dilakukan analisis inferensial untuk pengujian hipotesis penelitian terlebih dahulu dilakukan uji prasyarat analisis yang meliputi uji normalitas dan uji homogenitas sedangkan analisis data inferensial adalah untuk menguji hipotesis menggunakan Analisis Varian (ANOVA) Satu Jalan. Untuk mengetahui signifikansi perbedaan dari masing-masing kelompok perlakuan dilakukan uji lanjut menggunakan Uji Tukey.

HASIL PENELITIAN

Analisis varians dilakukan pada nilai RMSD yang dihasilkan dari ketiga metode *concordance*, yaitu metode linier, metode paralel linier, dan metode ekipersentil, masing-masing sebanyak 31 replikasi. Dalam penelitian ini diajukan tiga hipotesis penelitian yang perlu diuji secara empirik, yaitu hipotesis yang membandingkan nilai RMSD dengan metode linier dan nilai RMSD dengan metode paralel linier, nilai RMSD dengan metode paralel linier dan nilai RMSD dengan metode ekipersentil, serta nilai RMSD dengan metode linier dan nilai RMSD dengan metode ekipersentil. Hipotesis dalam penelitian ini secara inferensial dapat diuji dengan Analisis Varian (ANOVA) satu jalan dengan pendekatan statistik parametrik karena prasyarat normalitas dan homogenitas terpenuhi.

Berdasarkan hasil perhitungan ANOVA satu jalan, pada taraf signifikansi $\alpha = 0,05$, memberikan nilai $F_{hitung} = 3,0786$ lebih kecil dibandingkan $F_{tabel} = 3,0977$.

Hal ini berarti hipotesis nol (H_0) diterima, artinya tidak terdapat perbedaan rerata nilai RMSD dari metode linier, metode paralel linier, dan metode ekipersentil secara statistika.

Tabel 1. Hasil Perhitungan Analisis Varian (ANOVA) Satu Jalan pada $\alpha=0,05$

Sumber Varians	JK	Db	RJK	Fhit	Ftabel $\alpha=0,05$
Antar	0,6267	2	0,3133	3,0786	3,0977
Dalam	9,1600	90	0,1018		
Total	9,7867	92			

Nilai rerata RMSD dari ketiga metode *concordance* dapat dilihat pada tabel berikut.

Tabel 2. Nilai rerata RMSD dari Ketiga Metode *Concordance*

Metode Concordance	RMSD
Metode Linier	3,1644
Metode Paralel Linier	3,0964
Metode Ekipersentil	3,2943

Pada tabel 2, perbedaan nilai rerata RMSD antara ketiga metode *concordance* sangat kecil. Nilai rerata RMSD terkecil terdapat pada metode paralel linier sebesar 3,0964 disusul nilai rerata RMSD pada metode linier sebesar 3,1644, dan nilai rerata RMSD terbesar terdapat pada metode ekipersentil yaitu sebesar 3,2943.

PEMBAHASAN

Perangkat tes yang digunakan dalam penelitian ini adalah perangkat tes Ulangan Akhir Semester (UAS) mata pelajaran kimia yang dikembangkan oleh dua sekolah yang berbeda, yaitu SMAN 1 Tambun Selatan dan SMAN 1 Cikarang Utara, melalui MGMP/MKKS di sekolah masing-masing. Kedua perangkat tes tersebut disusun berdasarkan kisi-kisi yang sama (standar kompetensi dan kompetensi dasar yang sama). Sebelum digunakan untuk penelitian, kedua perangkat tes ini sebelumnya telah diujicobakan terlebih dahulu di sekolah lain. Kemudian, butir-butir yang valid digunakan sebagai instrumen tes pada penelitian ini.

Skor perolehan hasil tes siswa menunjukkan bahwa rerata hasil belajar siswa pada mata pelajaran Kimia di SMAN 1 Tambun Selatan lebih tinggi

dibandingkan dengan SMAN 1 Cikarang Utara. Hal tersebut diperkuat oleh hasil konversi skor dengan ketiga metode *concordance*. Hasil konversi skor dari SMAN 1 Tambun Selatan ke SMAN 1 Cikarang Utara (X ke Y) dengan menggunakan metode linier, metode paralel linier, dan metode ekipersentil, menunjukkan bahwa nilai hasil konversi skornya menjadi lebih rendah dibandingkan nilai asli hasil tesnya. Jadi, nilai siswa di SMAN 1 Tambun Selatan akan lebih rendah jika dibawa ke SMAN 1 Cikarang Utara, begitupun sebaliknya, nilai SMAN 1 Cikarang Utara akan lebih tinggi jika dibawa ke SMAN 1 Tambun Selatan.

Tujuan utama dari penelitian ini adalah untuk mengetahui perbandingan akurasi dari ketiga metode *concordance*, yaitu metode linier, metode paralel linier, dan metode ekipersentil. Akurasi ketiga metode *concordance* tersebut dilihat melalui nilai rerata RMSD. Hasil pengujian hipotesis penelitian di atas dengan menggunakan Anava satu jalan menyatakan bahwa H_0 diterima, artinya tidak terdapat perbedaan rerata nilai RMSD dari metode linier, metode paralel linier, dan metode ekipersentil secara statistika.

Perbedaan yang tidak terdapat antara nilai rerata RMSD dari metode linier, metode paralel linier, dan metode ekipersentil berdasarkan hasil pengujian hipotesis, dimungkinkan karena terdapatnya data pencilan pada data hasil skor tes UAS mata pelajaran kimia di kedua sekolah yang dijadikan tempat penelitian. Adanya data pencilan (*outlier*) akan memperbesar kisaran data skor tes, hal ini tentunya akan mempengaruhi hasil penarikan kesimpulan akhir.

Nilai rerata RMSD tersebut menunjukkan akurasi pengukuran dari metode *concordance* yang digunakan. Akurasi pengukuran merupakan ketepatan atau kedekatan hasil pengukuran dengan nilai yang sebenarnya atau nilai yang dianggap benar, ditandai dengan rendahnya kesalahan dalam pengukuran (*standard error*). Rerata RMSD yang kecil menunjukkan akurasi yang lebih tinggi daripada rerata RMSD yang besar.

Pada tabel 2 menunjukkan secara empirik, perbedaan nilai rerata RMSD antara ketiga metode *concordance* sangat kecil. Nilai rerata RMSD terkecil terdapat pada metode paralel linier sebesar 3,0964 disusul nilai rerata RMSD pada metode linier sebesar 3,1644, dan nilai rerata RMSD terbesar terdapat pada metode ekipersentil yaitu sebesar 3,2943. Nilai RMSD yang kecil menunjukkan akurasi yang tinggi, begitu pun sebaliknya nilai RMSD yang besar menunjukkan akurasi yang lebih rendah. Walaupun perbedaan nilai rerata RMSD dari ketiga metode *concordance* sangat kecil, hasil penelitian menunjukkan bahwa metode paralel linier dan metode linier sedikit lebih akurat dibandingkan dengan metode ekipersentil.

Berdasarkan hal tersebut, asumsi awal mengenai metode ekipersentil memiliki akurasi yang lebih tinggi dibandingkan dengan metode *concordance* yang lain tidak terpenuhi. Hal ini disebabkan karena kedua skor tes responden pada penelitian ini memiliki distribusi skor yang berbeda, distribusi tersebut terkait dengan rerata dan simpangan bakunya. Sedangkan metode ekipersentil cukup baik digunakan, jika distribusi skor pada kedua tesnya sama. Hal ini

menyebabkan konversi skor menggunakan metode ekipersentil menunjukkan akurasi yang lebih rendah dibandingkan dengan metode paralel linier, dan metode linier.

SIMPULAN

Akurasi pengukuran merupakan ketepatan atau kedekatan hasil pengukuran dengan nilai yang sebenarnya atau nilai yang dianggap benar, ditandai dengan rendahnya kesalahan dalam pengukuran (*standard error*). Rerata RMSD yang kecil menunjukkan akurasi yang lebih tinggi daripada rerata RMSD yang besar.

Perbedaan yang tidak terdapat di antara nilai rerata RMSD dari metode linier, metode paralel linier, dan metode ekipersentil, berarti untuk melakukan penyetaraan dapat memilih salah satu dari tiga metoda tersebut.

DAFTAR PUSTAKA

- Arifin, Zaenal. (2009). *Evaluasi Pembelajaran Prinsip dan Teknik Prosedur*. Bandung: Remaja Rosdakarya.
- Cronbach, Lee J. (1984). *Essentials of Psychological Testing*. New York: Harper and Row Publisher.
- Djaali dan Pudji Muljono. (2004). *Pengukuran dalam Pendidikan*. Jakarta: Program Pascasarjana Universitas Negeri Jakarta.
- Dorans, Neil J. (2004). "Equating, Concordance, and Expectation". *Jurnal Applied Psychological Measurement*, Vol. 28 (4).
- Hambleton, R. K., dan H. Swaminathan. (1985). *Item Response Theory: Principles and Applications*. Boston: Kluwer Academic Publisher.
- Kolen, M. J dan R. L. Brennan. (1995). *Test Equating Methods and Practice*. New York: Springer Verlag Inc.
- Livingston, Samuel A. (1993). "Small-Sample Equating with Log-Linier Smoothing". *Journal of Educational Measurement*, Vol. 30(1).
- Naga, Dali S. (1992). *Pengantar Teori Skor pada Pengukuran Pendidikan*. Jakarta: Gunadarma.
- Suryabrata, Sumadi. (2006). *Psikologi Pendidikan*. Jakarta: PT Raja Grafindo Persada.

Wiersma, Willian dan Stephen G. Jurs. (1990). *Educational Measurement and Testing*. Massachusetts: Allyn and Bacon.

Yin, Ping, R. L. Brennan, dan M. J. Kolen. (2004). "Concordance Between ACT and ITED Scores From Different Population". *Journal Applied Psychological Measurement*, Vol. 28(4).