

**DIMENSIONAL VALIDITY OF MATHEMATICS TEST
IN THE NATIONAL EXAM FOR JUNIOR SECONDARY
SCHOOLS (SMP) 2003-2006**

Badrun Kartowagiran

Abstract

This research aimed at investigating the dimension of mathematics test in the 2003, 2005 and 2006 National Exams for Junior Secondary Schools. Dimension in this research referred to the sub-skills in the mathematic skills. This research was an explorative descriptive research with the sources of data in the form of students responses from high, middle and low categories of schools in Yogyakarta province, with N = 2700, 3400, 3020 respectively for the 2003, 2005, and 2006. Data were analyzed using the exploratory factor analysis to find the number of loading factors in the tests. Findings suggested that the 2003, 2005, and 2006 Mathematics National Exams had three dimension covering the algebra, geometry, and measurement. Such tests can only explain 35% or about one third of all variances in the learners' mathematic ability. Findings implied that test developers should add more factors and variation in terms of the levels of difficulty in the test items so that the test could explain more variances of mathematic ability. On the other hand, practitioners should take the results of this research as feedback in their teaching of mathematics at schools.

Key words: *dimensional validity*

**VALIDASI DIMENSIONALITAS PERANGKAT TES UJIAN
AKHIR NASIONAL SMP MATA PELAJARAN
MATEMATIKA 2003-2006**

Badrun Kartowagiran

Abstrak

Penelitian ini bertujuan untuk mengetahui banyaknya dimensi perangkat tes UAN Matematika tahun 2003, 2005, dan 2006. Dalam penelitian ini, yang dimaksud dengan dimensi adalah subkemampuan yang ada dalam kemampuan matematik. Penelitian ini bersifat deskriptif eksploratif. Data utama penelitian ini adalah respon siswa peserta Ujian Akhir Nasional yang ada di Provinsi D.I. Yogyakarta. Jumlah lembar jawaban siswa yang dianalisis sebanyak 2700 untuk tahun 2003, 3400 untuk tahun 2005, dan 3020 untuk tahun 2006. Respon siswa diambil dari SMP kelompok tinggi, sedang, dan rendah. Data dianalisis menggunakan teknik analisis faktor eksploratori, kemudian diestimasi banyaknya muatan dimensi/faktor dalam perangkat tes. Disimpulkan bahwa perangkat tes UAN Matematika tahun 2003, 2005, dan 2006 mengukur 3 dimensi subkemampuan yakni aljabar, geometri, dan pengukuran. Perangkat tes tersebut, hanya mampu menjelaskan sekitar 35% atau sekitar sepertiga varians kemampuan matematika siswa. Terkait dengan hal ini, pengembang tes UAN mata pelajaran matematika perlu mengusahakan agar menambah faktor dan variasi tingkat kesulitan butir tes, sehingga tes lebih dapat menjelaskan variansi kemampuan peserta tes. Sebaliknya bagi praktisi, sebaiknya memanfaatkan hasil penelitian ini untuk umpan balik dalam pembelajaran matematika di sekolah.

Kata kunci: *validasi dimensionalitas*

Pendahuluan

Ujian Akhir Nasional (UAN), sebagai pengganti Evaluasi Belajar Tahap Akhir (Ebtanas) merupakan salah satu proses pengukuran hasil belajar yang telah dilaksanakan secara nasional di Indonesia mulai tahun 1985. UAN ini dilaksanakan di berbagai jenjang pendidikan, yakni Sekolah Menengah Pertama dan sampai Sekolah Menengah Atas. Adapun tujuan UAN sebagai berikut: (1) untuk memperoleh informasi tentang mutu hasil pendidikan secara nasional, (2) mengukur pencapaian hasil belajar siswa baik sekolah/madrasah negeri maupun swasta, (3) memperoleh gambaran perbandingan mutu pendidikan pada sekolah madrasah, antar sekolah/madrasah, dan antar wilayah dari tahun ke tahun, (4) menjadi bahan penentuan kebijakan pembinaan sekolah/madrasah, (5) sebagai bahan pertimbangan dalam memberikan Surat Tanda Tamat Belajar dan seleksi masuk ke jenjang pendidikan yang lebih tinggi.

Pada perancangan, perakitan tes, dan analisis butir tes, pendekatan teori yang digunakan pendekatan tes unidimensi, yang hanya mengukur dimensi tunggal saja. Pada teori respons butir ini, ada asumsi yang harus dipenuhi, yakni independensi lokal dan unidimensi (Hambleton, Swaminathan dan Rogers, 1991 serta Hulin dkk., 1983). Independensi lokal terjadi jika faktor-faktor yang mempengaruhi prestasi menjadi konstan, maka respons subjek terhadap pasangan butir yang manapun akan independen secara statistik satu sama lain. Unidimensi, artinya setiap butir tes hanya mengukur satu kemampuan. Asumsi unidimensi dapat ditunjukkan hanya jika tes mengandung hanya satu komponen dominan yang mengukur prestasi suatu subyek.

Pada kenyataannya di lapangan, asumsi unidimensi sulit terpenuhi. Hal ini sesuai dengan pendapat bahwa kebanyakan tes pendidikan dan psikologi pada beberapa tingkat bersifat multidimensi (Bolt dan Lall, 2003; Ackerman, dkk., 2003). Analisis secara unidimensi pada data yang realitasnya multidimensi akan mengakibatkan terjadinya kesalahan sistematis dalam administrasi tes. Sebagai akibatnya, informasi yang diperoleh akan menyesatkan dan tentunya merugikan peserta didik.

Di satu sisi, matematika merupakan pengetahuan yang memiliki objek dasar yang abstrak, yang berdasarkan kebenaran konsistensi, tersusun secara hirarkis dan sesuai dengan kaidah penalaran yang logis. Matematika dipandang sebagai salah satu unsur instrumental dalam sistem proses belajar mengajar untuk mencapai tujuan pendidikan dan berperan dalam perkembangan ilmu dan teknologi. Namun di sisi lain, berdasarkan pengamatan yang dilakukan tahun 2005 dan tahun 2006, matematika merupakan mata pelajaran yang dianggap sulit. Hal ini dikuatkan dengan hasil Ujian Akhir Nasional (UAN) peserta didik SMP dan SMU, ketidakilulusannya banyak disebabkan oleh pencapaian pada pelajaran matematika yang rendah. Pada tahun 2005, ada sebanyak 4.567 peserta didik SMP/MTs yang nilai matematika pada UAN tidak mencapai 4,01. Pada tahun 2006, ada 3.291 peserta didik yang nilai UAN untuk mata pelajaran matematika tidak mencapai 4,26. Informasi selengkapnya tentang peserta didik yang tidak lulus menempuh UAN berdasarkan nilai matematikanya disajikan pada Tabel 1.

Tabel 1. Peserta Didik SMP/MTS di DI Yogyakarta yang Tidak Lulus Menempuh UAN Berdasarkan Nilai Matematika

UAN tahun	Banyaknya peserta didik yang tidak lulus	Jumlah peserta didik SMP/MTs peserta UAN	Persentase peserta didik yang tidak lulus	Batas nilai lulus
2005	4.567	49.454	9,61%	4,01
2006	3.291	47.545	4,86%	4,26

Sumber: Laporan Pelaksanaan UAN Depdiknas Prov. DI Yogyakarta 2005 dan 2006

Untuk dapat menjawab benar butir-butir dalam tes matematika, peserta didik perlu memiliki beberapa kemampuan. Permasalahan matematika yang dikemas dalam soal cerita diperlukan kemampuan membaca dan kemampuan matematika untuk mendeterminasikan jawaban benar (Reckase, 1997). Pernyataan ini juga diperkuat oleh Bolt dan Lall (2003) juga (Ackerman, dkk., 2003). Hasil penelitian Thulber, Shinn dan

Smolkowski (2002) juga menunjukkan bahwa dalam penyelesaian permasalahan dalam matematika, paling tidak ada 3 dimensi yang termuat, yakni dimensi aplikasi, dimensi bahasa dan dimensi komputasi. Berdasarkan hal ini, diperoleh bahwa tes matematika sebenarnya bersifat multidimensi atau mengukur beberapa kemampuan, tidak hanya kemampuan matematika peserta didik.

Adanya muatan multidimensi pada perangkat tes yang dianalisis dengan model unidimensi menyebabkan estimasi kemampuan yang tidak tepat dan memberikan informasi yang menyesatkan. Terkait dengan hal ini, diperlukan suatu penelitian tentang banyaknya muatan dimensi (validasi dimensi) perangkat tes UAN Matematika SMP, yang juga dapat dimanfaatkan sebagai sarana diagnostik dalam rangka meningkatkan kompetensi matematika peserta didik.

Tujuan penelitian ini adalah untuk mengetahui banyaknya dimensi pada perangkat tes UAN mata pelajaran matematika 2002/2003, 2004/2005, dan 2005/2006. Penelitian ini penting untuk dilaksanakan karena penelitian tentang validasi dimensi yang termuat dalam suatu tes masih sangat jarang dilakukan pada perangkat tes yang diujikan baik dalam skala lokal maupun nasional. Oleh karena itu, diharapkan hasil penelitian ini memberikan wacana baru dalam kegiatan peningkatan kualitas berbagai perangkat soal khususnya Ujian Akhir Nasional. Terlebih lagi analisis butir dari suatu perangkat tes dengan pendekatan teori respons butir multidimensi, mengingat teori ini merupakan teori yang relatif baru dan sedang berkembang.

Dalam tes, representasi domain isi (*content*) merupakan hal yang penting dalam validitas tes. Melalui suatu evaluasi tentang tes yang dapat dilakukan oleh ahli materi suatu subjek (*subject-matter expert, SME*), dapat diketahui butir-butir yang menyusun suatu tes dan relevansinya pada domain yang direncanakan (Sireci dan Geisinger, 1995). Hasil evaluasi ini akan menunjukkan dimensi-dimensi yang konsisten dengan struktur isi, maupun dimensi-dimensi yang tidak konsisten dengan struktur isi. Hasil ini juga mendasari suatu penyekalaan multidimensi (*multidimensional scaling, MDS*).

Penyekalaan multidimensi memposisikan butir-butir pada suatu ruang (*space*) pada lokasi koordinat tertentu. Ruang ini ditentukan oleh dimensi-dimensi tertentu sebagai sumbunya. Jarak relatif antar pasangan butir menyatakan perbedaan-perbedaan antar butir (Bolt, 2001). Semakin dekat jarak suatu butir dengan butir lainnya, semakin besar persamaan sifat kedua butir tersebut. Berdasarkan kedekatan jaraknya atau persamaan sifat-sifatnya ini, butir-butir dapat dikelompokkan sesuai dengan substansinya. Analisis ini dikenal dengan analisis validitas kluster secara hierarki (Sireci dan Geisinger, 1995).

Tes dalam pendidikan dan psikologi yang mengukur variabel-variabel laten bersifat multidimensi. Jika pada analisis butirnya menggunakan pendekatan unidimensi, maka yang diperoleh adalah ketidaktepatan dalam ukurannya (Wang, Chen, dan Cheng, 2004). Hal ini disebabkan karena pendekatan unidimensi mengabaikan korelasi antar kemampuan laten. Pendekatan pengukuran secara multidimensi memperhatikan hubungan antar kemampuan laten yang menyebabkan meningkatnya ketepatan pengukuran. Kelebihan lain dari teori respons butir multidimensi disampaikan oleh De la Tore dan Patz (2005) bahwa analisis dengan pendekatan ini memberikan tambahan informasi yang meningkatkan ketepatan estimasi parameter butir. Pada keadaan ini, unidimensi merupakan kasus dari multidimensi, yakni ketika korelasi antar variabel laten sama dengan nol.

Ada dua tipe struktur dimensi, yakni tipe dominan ganda dengan korelasi antar dimensi dan tipe satu dimensi dominan dengan beberapa dimensi minor (Kirisci, Hsu dan Yu, 2001). Senada dengan itu, Wang, Chen dan Cheng (2004) menyatakan bahwa ada dua jenis multidimensi, yakni multidimensi antar butir dan multidimensi dalam butir.

Tipe dominan ganda dengan korelasi antar dimensi sama dengan multidimensi dalam butir (*within-item multidimensional*). Pada tipe ini butir dapat memuat beberapa dimensi dominan, dan dimensi dominan ini saling berkorelasi. Sebagai contohnya, tes prestasi matematika memuat pemahaman konsep matematika, tes kemampuan berhitung, dan tes kemampuan membaca.

Tipe yang kedua, yakni satu dimensi dominan dengan beberapa dimensi minor, yang juga disebut dengan multidimensi antar butir (*betwitem multidimensional*). Pada suatu tes, ada dimensi dominan yang termuat. Dimensi-dimensi ini memiliki dimensi-dimensi minor, antar dimensi minor terdapat korelasi.

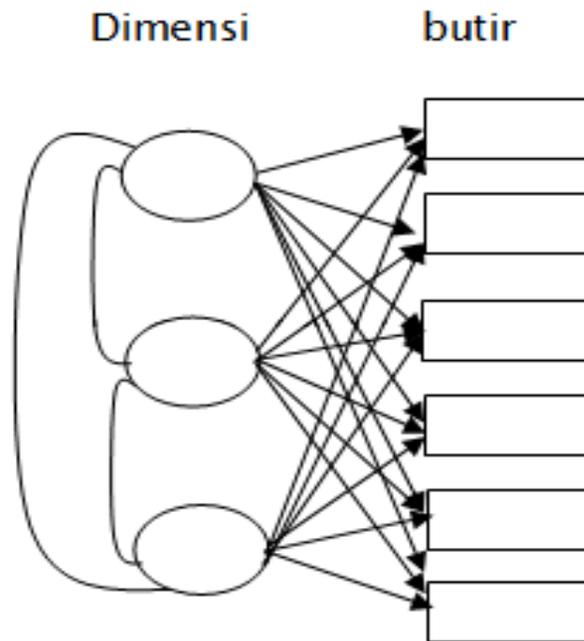
Banyak cara yang dapat dilakukan untuk mengetahui banyaknya dimensi yang termuat dalam tes. Cara-cara tersebut yakni dengan analisis faktor yang merupakan bagian dari persamaan model struktural (*structural equation modeling, SEM*), dengan prosedur DETECT ataupun analisis faktor non linear *Normal Ogive Harmonic Analysis Robust Method* (NOHARM) (Finch dan Habing, 2005). Cara lain yang dapat digunakan yakni dengan prosedur Statistik T-Stoud (Sam Elias, John Hattie, Graham Douglas, tth).

Dalam suatu penelitian, biasanya digunakan instrumen yang melibatkan butir-butir yang banyak. Untuk memahami data seperti ini, biasanya digunakan analisis faktor. Analisis faktor digunakan untuk mereduksi data, dengan menemukan hubungan antar variabel yang saling bebas (Stapleton, 1997), yang kemudian terkumpul dalam variabel yang jumlahnya lebih sedikit untuk mengetahui struktur dimensi laten (Anonim, 2001; Garson, 2006) , yang disebut dengan faktor. Faktor ini merupakan variabel yang baru, yang disebut juga dengan variabel laten, variabel konstruk dan memiliki sifat tidak dapat diketahui langsung (*unobservable*). Analisis faktor dapat dilakukan dengan dua cara, yakni analisis faktor eksploratori (*eksploratory factor analysis*) dan analisis faktor confirmatory (*confirmatory faktor analysis*).

Analisis faktor eksploratori merupakan suatu teknik untuk mendeteksi dan mengases sumber laten dari variasi atau kovariansi dalam suatu pengukuran (Joreskog & Sorbom, 1993). Analisis faktor eksploratori bersifat mengeksplor data empiris untuk menemukan dan mendeteksi karakteristik dan hubungan antar variabel tanpa menentukan model pada data. Pada analisis ini, peneliti tidak memiliki teori *a priori* untuk menyusun hipotesis (Stapleton, 1997). Mengingat sifatnya yang eksplorasi inilah, hasil analisis faktor eksploratori ini lemah. Hasil analisis, yang menjelaskan hubungan antar variabel semata, juga tidak didasarkan pada teori yang ada. Hasil analisis juga hanya tergantung data empiris, dan jika variabel

terobservasinya banyak, hasil analisis akan sulit dimaknai (Stapleton, 1997). Biasanya analisis faktor terkait erat dengan pertanyaan tentang validitas (Nunnally, 1978). Ketika faktor-faktor teridentifikasi dihubungkan, analisis faktor eksploratori menjawab pertanyaan tentang validitas konstruk, apakah suatu skor mengukur apa yang seharusnya diukur.

Analisis faktor konfirmatori didasarkan pada premis bahwa masing-masing variabel manifest atau variabel yang dapat diamati secara sendiri tidak dapat menggambarkan secara sempurna suatu konsep atau suatu variabel laten atau variabel konstruk. Terkait dengan hal ini, dengan berlandaskan teori, satu konsep atau variabel laten atau variabel konstruk dapat digambarkan secara bersama oleh beberapa variabel manifest. Analisis faktor eksploratori yang akan dilakukan pada penelitian ini ditampilkan dalam Gambar 1 berikut.



Gambar 1. Hubungan Antara Butir dan Dimensi/Faktor pada Analisis Faktor Eksploratori pada Penelitian ini

Tampak dalam Gambar 1 di atas bahwa butir-butir itu tidak jelas, masuk pada dimensi mana. Dengan penelitian ini diharapkan ditemukan butir mana yang masuk dimensi 1, butir mana yang masuk dimensi 2, dan butir mana saja yang masuk dalam dimensi 3.

Penelitian ini merupakan penelitian yang mengkaji perangkat tes UAN mata pelajaran matematika 2002/2003, 2004/2005, dan 2005/2006. Menurut James dan James (1976), matematika adalah ilmu tentang logika mengenai bentuk, susunan, besaran, dan konsep-konsep berhubungan dengan jumlah, terbagi dalam tiga bidang yaitu aljabar, analisis dan geometri. Matematika dapat pula dipandang sebagai suatu struktur dari hubungan-hubungan yang mengaitkan simbol-simbol. Pandangan ini berdasarkan pemikiran tentang bagaimana terbentuknya matematika. Berkaitan dengan hal ini, Ruseffendi mengemukakan bahwa matematika terbentuk sebagai hasil pemikiran manusia yang berhubungan dengan ide, proses dan penalaran (Ismail, dkk., 1998: 1, 4).

Terkait dengan proses terbentuknya, matematika merupakan pengetahuan yang dimiliki manusia. Pengetahuan ini timbul karena kebutuhan manusia untuk memahami alam sekitar. Alam dijadikan sumber-sumber ide untuk memperoleh konsep matematika melalui abstraksi dan idealisasi.

Mula-mula dibuat model, dan dari model itu dibuat definisi-definisi dan aksioma-aksioma. Definisi merupakan sebuah persetujuan untuk menggunakan sesuatu sebagai pengganti sesuatu yang lain, biasanya berupa suatu ekspresi bahwa hal tersebut terlalu sulit untuk ditulis dengan mudah (James dan James, 1976). Adapun aksioma merupakan pernyataan yang diterima tanpa pembuktian. Melalui proses berfikir yang disebut dengan logika deduktif, diperoleh suatu teorema-teorema (Allendoerfer, 1969: 7). Teorema hasil proses berfikir ini merupakan suatu kesimpulan umum yang dapat dibuktikan (James dan James, 1976). Definisi-definisi, aksioma-aksioma dan teorema-teorema ini merupakan kesatuan yang menyusun suatu konsep matematika.

Konsep-konsep matematika bersifat abstrak, yang saling berkorelasi membentuk konsep baru yang lebih kompleks dan tersusun secara hierarkis, konsep yang satu menjadi dasar untuk mempelajari konsep

selanjutnya (Herman Hudoyo, 1988). Akhirnya konsep matematika yang ditemukan diterapkan kembali ke alam, dan manusia memanfaatkannya untuk memenuhi kebutuhan hidupnya.

Berkaitan dengan diterapkannya konsep matematika untuk memenuhi kebutuhan hidup manusia, matematika sering digunakan sebagai bahasa atau alat untuk menyelesaikan masalah, seperti masalah-masalah sosial, ekonomi, fisika, kimia, biologi dan teknik. Peran inilah yang menyebabkan matematika mendapat julukan sebagai ratunya ilmu (*queen of science*). Mengenai bagaimana seseorang menggunakan matematika untuk memecahkan masalah di berbagai bidang ilmu, tergantung pada kemampuan orang tersebut menguasai matematika dan menerapkannya.

Matematika perlu dikomunikasikan dari satu orang kepada orang lain, atau dari satu generasi ke generasi selanjutnya agar dapat bermanfaat bagi orang atau generasi lain. Selain itu juga dapat bermanfaat bagi perkembangan matematika. Pembelajaran matematika di sekolah merupakan bagian dari komunikasi ini. Proses komunikasi ini merupakan bagian dari pendidikan matematika.

Menurut James dan James (1976), matematika terdiri dari tiga cabang utama, yakni aljabar, geometri dan analisis. Dari ketiga cabang ini, dalam pembelajaran matematika, aljabar dipelajari oleh siswa terlebih dahulu pada pendidikan formal. Pada sekolah dasar, konsep matematika yang dipelajari masih berkisar pada aljabar dan geometri. Di SMP, materi yang dipelajari menjadi semakin kompleks, tidak hanya aljabar dan geometri saja, namun juga termasuk relasi dan fungsi yang merupakan bagian dari analisis.

Materi matematika pada tingkat yang lebih rendah akan lebih mudah dipelajari peserta didik, dibandingkan dengan materi pada tingkat yang lebih tinggi. Hal ini terkait dengan kemampuan untuk mempelajari materi yang lebih tinggi, diperlukan kemampuan yang lain, yaitu kemampuan berabstraksi. Pernyataan ini didukung Heri Retnawati, Pardi, dan Prastawa Rahmawan (1994), bahwa konsep aljabar lebih mudah dipahami siswa SMU yang telah menjadi mahasiswa baru dibandingkan dengan konsep geometri.

Matematika tersusun secara hierarkis, konsep yang satu menjadi dasar untuk mempelajari konsep selanjutnya (Herman Hudoyo, 1988). Sifat

ini menyebabkan penguasaan matematika siswa pada proses pembelajaran dipengaruhi oleh kemampuannya menguasai konsep matematika sebelumnya. Hal ini mengakibatkan kemampuan matematika siswa pada jenjang SMP dipengaruhi oleh penguasaan konsep matematika selama di sekolah dasar, dan penguasaan matematika di SMA dipengaruhi oleh penguasaan konsep matematika di SMP, begitu seterusnya.

Keberhasilan dalam pembelajaran matematika diketahui melalui evaluasi. Bagi peserta didik, evaluasi pembelajaran dilakukan melalui tes. Tes prestasi belajar matematika siswa dapat berupa tes objektif maupun tes uraian. Setelah direspons peserta didik, selanjutnya dilakukan penskoran. Tes objektif diberi skor dikotomi sedangkan tes uraian biasanya diberi skor berjenjang.

Tes matematika disusun oleh pendidik atau oleh pengembang tes tentu saja mempunyai tujuan mengukur kemampuan matematika peserta didik (bersifat unidimensi). Namun kenyataannya, yang terjadi lain. Permasalahan matematika yang dikemas dalam soal cerita diperlukan kemampuan membaca dan kemampuan matematika untuk mendeterminasikan jawaban benar (Reckase, 1997). Pernyataan ini juga diperkuat oleh Bolt dan Lall (2003) juga (Ackerman, dkk., 2003). Hasil penelitian Thulber, Shinn dan Smolkowski (2002) juga menunjukkan bahwa dalam penyelesaian permasalahan dalam matematika, paling tidak ada 3 dimensi yang termuat, yakni dimensi aplikasi, dimensi bahasa dan dimensi komputasi.

Metode Penelitian

Penelitian ini dilaksanakan di DI Yogyakarta, dan pengumpulan data dilakukan di Kantor Dinas Pendidikan Provinsi Daerah Istimewa Yogyakarta. Kegiatan penelitian dilaksanakan selama enam bulan yakni dari bulan Mei 2007 sampai dengan bulan Oktober 2007.

Penelitian ini bersifat deskriptif eksploratif, karena dalam penelitian ini dilakukan estimasi muatan multidimensi perangkat tes UAN mata pelajaran matematika SMP. Data utama dalam penelitian ini adalah respons siswa peserta Ujian Akhir Nasional pada mata pelajaran Matematika SMP

yang ada di Provinsi D.I. Yogyakarta tahun pelajaran 2002/2003, 2004/2005, 2005/2006. Jumlah lembar jawaban siswa yang diambil sebanyak 2700 untuk tahun pelajaran 2002/2003, 3400 untuk tahun pelajaran 2004/2005, dan 3020 untuk tahun pelajaran 2005/2006. Respon siswa diambil dari SMP kelompok tinggi, SMP kelompok sedang, dan SMP kelompok rendah.

Data yang berhasil dikumpulkan dianalisis dengan menggunakan teknik analisis faktor eksploratori. Dengan teknik analisis ini dapat diestimasi banyaknya muatan dimensi dalam perangkat tes. Banyaknya dimensi ini diketahui dengan menghitung banyaknya faktor yang termuat dalam perangkat tes pada analisis faktor. Analisis faktor dilakukan dengan bantuan komputer program SPSS.

Hasil Penelitian dan Pembahasan

1. Karakteristik Butir

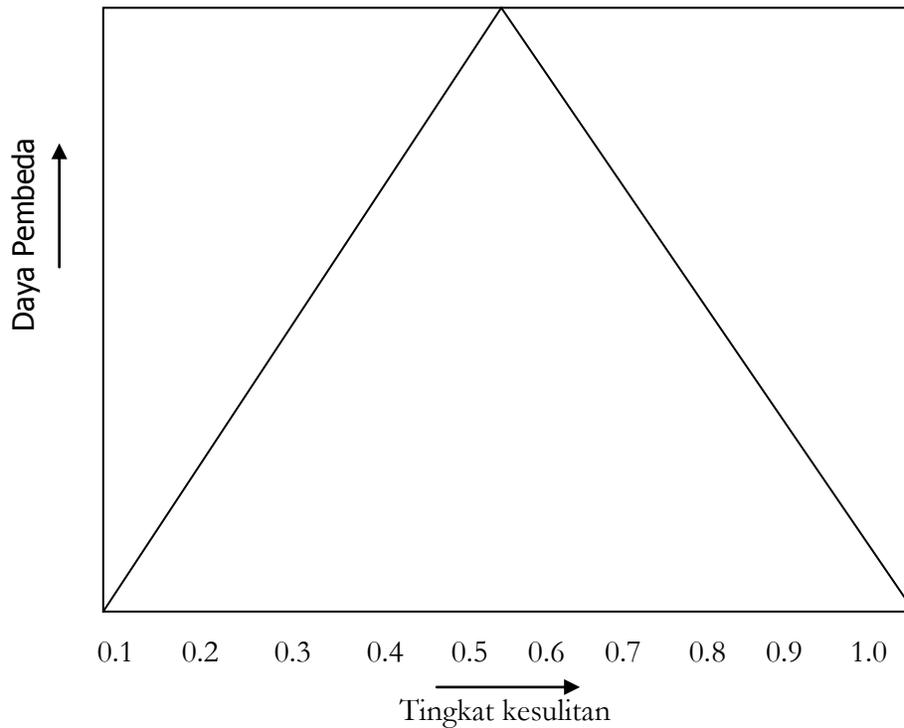
Pada penelitian ini, analisis untuk karakteristik perangkat soal Ujian Nasional SLTP berdasarkan teori tes klasik dilakukan dengan bantuan program *Bilog* versi 3.07 fase pertama. Karakteristik perangkat tes berupa estimasi parameter butir yang meliputi tingkat kesukaran dan daya pembeda. Hasil analisis dapat dilihat pada Tabel 2 berikut.

Tabel 2. Karakteristik Perangkat Soal UAN Matematika SLTP Tahun 2002/2003 – 2005/2006 berdasarkan Teori Tes Klasik

Tahun	Jumlah butir	Tingkat kesukaran (p)	Daya Pembeda (r bis)	Jumlah butir yang baik
2002/2003	40	0,24 – 0,842	0,178– 0,758	38
2004/2005	30	0,284 – 0,937	0,179– 0,775	30
2005/2006	30	0,448 – 0,958	0,323– 0,924	30

Mencermati Tabel 2 di atas, tampak bahwa bila dianalisis dengan menggunakan pendekatan teori klasik, ternyata hampir semua butir baik. Dalam tabel itu juga ditunjukkan bahwa untuk UAN tahun 2005/2006

tingkat kesukaran butir naik, demikian juga daya pembedanya. Memang, menurut teori klasik, ada hubungan antara tingkat kesukaran dan daya pembeda butir. Hal ini dapat dilihat pada Gambar 2.



Gambar 2. Hubungan Antara Daya Pembeda dan Tingkat Kesulitan

Hal penting yang juga harus diperhatikan dalam menganalisis empiris butir tes adalah kemampuan distraktor atau alternatif jawaban yang disediakan menarik peserta tes untuk memilihnya. Jangan sampai tidak seorang peserta tes-pun memilih alternatif jawaban yang disediakan. Diusahakan agar paling tidak, ada 2% peserta memilih alternatif jawaban yang disediakan. Hal ini sesuai dengan pendapat Fernandes (1984) yang menjelaskan distraktor dikatakan baik apabila paling tidak dipilih oleh 2% dari seluruh peserta.

Sementara itu, Nitko (1996) mengatakan distraktor dikatakan berfungsi manakala paling tidak dipilih oleh seorang peserta tes dari kelompok rendah. Pemilih dari kelompok rendah harus lebih banyak daripada kelompok atas. Distraktor juga dapat dikatakan berfungsi manakala peserta tes (siswa) dari kelompok atas dapat membedakan antara distraktor dan kunci jawaban sehingga yang memilih kunci jawaban lebih banyak daripada yang memilih distraktor.

2. Dimensionalitas Perangkat Tes UAN Matematika

Setelah diketahui karakteritiknya, butir-butir tes UAN Matematika diuji unidimensionalitasnya. Yang dimaksud dengan dimensi dalam penelitian ini adalah sub kemampuan dalam kemampuan matematika. Selanjutnya, untuk melakukan uji unidimensionalitas butir digunakan teknik analisis faktor. Hasil analisis dengan uji Keiser-Meyer-Oldkin dan uji Barlett diperoleh hasil seperti Tabel 3.

Tabel 3. Besarnya KMO untuk Perangkat Soal UAN Matematika SLTP Tahun 2002/2003 – 2005/2006

Tahun	Jumlah Butir	Besarnya KMO	Jumlah Komponen	Nama Komponen	Varians Total
2002/2003	40	0,963	5	geometri, aljabar, pengukuran, garis lurus dan fungsi	33,654%
2003/2004	30	.957	5	aljabar, geometri, pengukuran, statistik dan fungsi	36,023%
2005/2006	30	.968	3	aljabar, geometri, dan peng ukuran	32,650%

Hasil tersebut menunjukkan bahwa nilai KMO ketiga soal tes UAN Matematika $> 0,5$, yang menunjukkan bahwa ukuran sampling untuk ketiga tes ini telah sesuai. Hasil uji bartlett juga menunjukkan bahwa nilai p dari ketiga tes itu kurang dari 5%, yang menunjukkan bahwa butir-butir UAN matematika 2003, 2005, 2006 merupakan butir-butir yang saling

independen. Selanjutnya, dengan memperhatikan nilai eigen yang lebih dari 1, dapat diketahui varians total yang dapat dijelaskan oleh perangkat UAN mata pelajaran matematika 2002/2003, 2004/2005, dan 2005/2006, masing-masing 33,654%, 36,032%, dan 32,650%.

Hasil di atas memberikan gambaran bahwa baru sekitar 33% atau sekitar sepertiga kemampuan matematika yang dapat dijelaskan dengan soal tes Matematika tahun 2003, 2005, dan 2006. Hal ini mengindikasikan bahwa masih banyak materi yang belum diujikan dalam ujian nasional matematika. Subkemampuan lain dalam matematika yakni analisis, yang meliputi fungsi dan grafiknya belum nampak jelas pada penelitian ini. Hal ini mungkin dikarenakan fungsi yang dipelajari masih sedikit dan pada tes ini juga hanya memperoleh porsi yang sedikit pada UAN matematika SMP.

Tabel di atas juga menunjukkan bahwa soal UAN Matematika SMP Tahun 2003 terdiri dari 5 jenis kemampuan. Setelah dilakukan pencermatan bersama *Subject Matter Expert (SME)* dalam bidang matematika, yang terdiri dari dosen matematika murni, dosen pendidikan matematika, dan 2 orang guru matematika senior, dimensi-dimensi itu berturut-turut dinamai: geometri, aljabar, pengukuran, garis lurus dan fungsi. Untuk soal tes Matematika Tahun 2004 juga terdiri dari 5 kemampuan, yakni geometri, aljabar, pengukuran, statistik, dan fungsi. Sementara itu, soal UAN Matematika SMP Tahun 2006 terdiri dari 3 jenis kemampuan, yakni geometri, aljabar, dan pengukuran.

Penutup

Berdasarkan hasil analisis, dapat disimpulkan bahwa perangkat tes UAN Matematika SMP tahun 2002/2003 geometri, aljabar, pengukuran, garis lurus dan fungsi, tahun 2004/2005 geometri, aljabar, pengukuran, statistik, dan fungsi, sedangkan tahun 2005/2006 mengukur 3 faktor yakni aljabar, geometri dan pengukuran. Setelah dicermati untuk soal UAN Matematika tahun 2002/2003 dan tahun 2004/2005 ternyata hanya tiga faktor yang dominan, yaitu aljabar, geometri dan pengukuran. Hal ini sama dengan soal UAN Matematika SMP tahun 2005/2006.

Mencermati hasil analisis penelitian ini, diperoleh bahwa perangkat tes UAN matematika SMP dalam kurun waktu 3 tahun, hanya mampu menjelaskan sekitar 35% atau sekitar sepertiga varians kemampuan matematika siswa. Hal ini mengindikasikan bahwa masih banyak kemampuan matematika yang belum dimunculkan dalam soal UAN Matematika. Bila dicermati lebih jauh, ternyata soal juga kurang bervariasi tingkat kesulitan butir yang termuat di dalamnya, karena lebih banyak memuat butir dengan tingkat kesulitan yang rendah. Terkait dengan hal ini, pengembang tes UAN mata pelajaran matematika perlu mengusahakan agar menambah faktor dan variasi tingkat kesulitan butir tes, sehingga tes lebih dapat menjelaskan variansi kemampuan peserta tes.

Adanya hasil penelitian yang berdasarkan pada data empiris, bahwa perangkat tes UAN matematika SMP 2002/2003, 2004/2005, dan 2005/2006 mengukur 3 faktor yakni aljabar, geometri dan pengukuran, perlu menjadi perhatian sehingga dimanfaatkan untuk *washback-effect* dalam pembelajaran matematika di sekolah. Subkemampuan matematika yang terukur perlu dipelajari siswa secara detail, sehingga siswa dapat menguasai keseluruhan konsep matematika selama di SMP untuk keperluan hidup dan kehidupannya, maupun untuk melanjutkan ke jenjang yang lebih tinggi.

Daftar Pustaka

- Ackerman, T.A., dkk. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement*, 22, 37-53.
- Allendoerfer, O. (1969). *Principles of mathematics*. Auckland, NY: McGraw-Hill Book Company.
- Anonim. (2001). Factor analysis. *Journal of Consumer Psychology*, 10(1&2), 75-82. Lawrence Erlbaum.
- Bolt, D.M. dan Lall, V.M. (2003). Estimation of compensatory and noncompensatory multidimensional item response models using Marcov chain Monte-Carlo. *Applied Psychological Measurement*, 27, 395-414.

- Bolt, D.M. (2001). Conditional covariance-based representation of multidimensional teststructure. *Applied Psychological Measurement*, Vol. 27, No. 3, 244-257.
- De la Torre, J. dan Patz. (2005). Making the most of what we have: a practical application of multidimensional item response theory in scoring. *Educational and Behavioral Statistics*, 30, 295-311.
- Depdiknas DI Yogyakarta. (2005). *Laporan pelaksanaan UAN di SMP/MTs 2004/2005*. Tidak dipublikasikan.
- Depdiknas DI Yogyakarta. (2006). *Laporan pelaksanaan UAN di SMP/MTs 2005/2006*. Tidak dipublikasikan.
- Garson, D. (2006). *Factor analysis*. Diambil tanggal 24 September 2006 dari <http://www2.chass.ncsu.edu/garson/pa765/index.htm>.
- Finch, H. dan Habing, B. (2005). Perbandingan NOHARM dan DETECT dalam Penemuan Klaster Butir: Menghitung Dimensi dan Mengalokasikan Butir. *Journal of Education and Measurement*, Summer 2005 Vol. 42 No. 2 p. 149-169.
- Fernandes, H.J. X. (1984). *Evaluation of educational program*. Jakarta: National Education Planning, Evaluating and Curriculum Development.
- Hambleton, R.K., Swaminathan, H & Rogers, H.J. (1991). *Fundamental of item response theory*. Newbury Park, CA: Sage Publication Inc.
- Herman Hudojo. (1988). *Mengajar belajar matematika*. Jakarta: Dirjen PTPPLPTK.
- Heri Retnawati, Pardi & Prastawa Rahmawan. (1994). Kesiapan mahasiswa baru pendidikan matematika untuk mempelajari matematika di perguruan tinggi. *Laporan penelitian yang disusun dalam rangka lomba karya tulis ilmiah kependidikan tingkat nasional 1994*. Februari 1994.
- Hullin, C. L. , et al. (1983). *Item response theory: Application to psychological measurement*. Homewood, IL: Dow Jones-Irwin.

- Ismail, dkk. (1998). *Kapita selekta pembelajaran matematika*. Jakarta: Universitas Terbuka.
- James, G. & James, R.C. (1976). *Mathematics dictionary*. New York: Van Nostrand Reinhold Company.
- Joreskog, K. & Sorbom, D. (1993). *Lisrel 88: Structural equation modeling with the SIMPLIS command language*. Hillsdale, NJ: Scientific Software International.
- Kirisci, L., Hsu, T. dan Yu, L. (2001). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Applied Psychological Measurement*, 25, 146-162.
- Nitko, A.J. (1996). *Penilaian berkelanjutan berdasarkan kurikulum (PB2K): Kerangka, konsep, prosedur, dan kebijakan* (terj. AM. Ahmad) Jakarta: Pusat Pengembangan Agribisnis.
- Nunally, J. (1978). *Psychometric theory* (2nd ed.). New York: McGraw Hill.
- Reckase, M.D. (1997). A linear logistic multidimensional model for dichotomous item response data. In W.J. Linden & R.K. Hambleton (Eds), *Handbook of modern item response theory* (pp. 271-286). New York: Springer.
- Sireci, S. G., Geisinger, K. F. (1995). Using subject-matter experts to assess content representation: an MDS analysis. *Applied Psychological Measurement*. Vol. 19. No. 3, 241-255.
- Stapleton. (1997). *Basic concepts and procedures of confirmatory factor analysis*. Diambil tanggal 25 September 2006 dari <http://ericae.net/ft/Cfa.HTM>.
- Thulber, R.S., Shinn, M.R., dan Smolskowsky. (2002). What is measured in mathematics tests? Construct validity of curriculum-based on mathematics measures. *School Psychological Review*. Vol. 31. No.4, 498-514.

Van de Geer, J.P. (1971). *Introduction to multivariate analysis for the social sciences*. San Francisco : W.H. Freeman and Company.

Wang, W.C., Chen,P.,H., dan Cheng,Y.Y. (2004). Improving measurement precision of test batteries using multidimensional item response models. *Psychological Methods* Vol. Vol 9 No. 1,116-136.

Biodata

Dr. Badrun Kartowagiran, lahir di Bantul 25 Juli 1953 beralamat di Gejayan Jl. Mangga Gg. Apel 101 RT 07 RW 31 Condongcatur Depok Sleman. Sampai saat ini sebagai staf pengajar Fakultas Teknik dan Pascasarjana Universitas Negeri Yogyakarta. Pendidikan terakhir Program Doktor (S3) Universitas Gadjah Mada bidang studi Psikometri lulus tahun 2005.