

# Model Evaluasi Keberlanjutan SD Ke SMP Dalam Rangka Wajib Belajar 9 Tahun

Dr. Endang Mulyatiningsih<sup>1</sup>

**Abstrak.** *Studi ini bertujuan untuk: (1) menguji model pengukuran kapabilitas siswa SD untuk belajar ke SMP; (2) menguji model struktural alat keberlanjutan SD ke SMP; dan (3) mengevaluasi sistem keberlanjutan SD ke SMP dari aspek input, proses, dan produk belajar. Model evaluasi menggunakan pendekatan responsif dari Stake dan CIPP (context, input, process and product) dari Stufflebeam. Data yang diumpulkan terdiri dari data kuantitatif dan kualitatif yang diambil secara cross sectional. Sampel diambil menurut strata mutu sekolah. Data dikumpulkan dengan metode: tes, dokumentasi, observasi, kuesioner dan focus group discussion. Temuan yang diperoleh: (1) Kapabilitas siswa SD untuk belajar ke SMP berada pada kategori tidak mampu belajar yaitu nilai < 40 pada m.p. Matematika sebesar 16,42%; Kapabilitas siswa untuk belajar ke SMP dipengaruhi oleh variabel eksogen Tes Potensi Belajar (0,62) dan rerata rapor (0,49). (2) Hasil analisis prestasi siswa selama tujuh semester menunjukkan ada beda antar semua pengukuran yang diulang pada rentangan nilai F antara 38,238 sampai dengan 265,164. Hasil analisis post hoc anova menunjukkan nilai tes standar mutu SMP kelas 7 semester 1 tidak berbeda nyata dengan nilai rapor kelas 4 semester 1 pada mata pelajaran IPS ( $p = 0,175$ ) dan IPA ( $p = 0,527$ ). (3) Anteseden (input) potensi awal yang diharapkan adalah siswa memiliki prestasi belajar minimal pada nilai 5 tetapi ada 6,39% siswa memiliki nilai matematika kurang dari 4; Transaksi (proses) PSB SMP diharapkan bervariasi tetapi PSB yang diterapkan serentak dalam satu waktu; Produk belajar yang diharapkan adalah siswa dapat mencapai standar kompetensi lulusan tetapi 5,02% siswa belum mencapai standar nilai 4,25.*

**Kata Kunci:** *Evaluasi Keberlanjutan; Wajib Belajar*

## 1. Pendahuluan

Ujian kelulusan pada tingkat SD dan sistem seleksi masuk ke SMP dikelola oleh masing-masing daerah sesuai dengan konteks desentralisasi pendidikan. Dinas Pendidikan Propinsi DIY pada tahun 2006 menerapkan kebijakan Ujian Sekolah Daerah (USD) yang hasilnya digunakan sebagai salah satu alat seleksi masuk ke SMP. Setelah sistem keberlanjutan SD ke SMP dilakukan dengan berbagai cara, maka perlu ada evaluasi untuk memperoleh informasi yang diperlukan dalam pengambilan keputusan kebijakan yang menyangkut keberlanjutan layanan wajib belajar 9 tahun.

Evaluasi keberlanjutan SD ke SMP termasuk dalam lingkup kebijakan pemerataan dan perluasan akses pendidikan. Dalam layanan wajib belajar 9 tahun diharapkan semua anak usia sekolah, khususnya anak perempuan yang berada dalam keadaan sulit dan kelompok etnik minoritas, mempunyai akses dan dapat menyelesaikan pendidikan dasar yang bebas, wajib dan berkualitas baik. Wajib belajar pendidikan dasar menjamin dampak belajar dapat dicapai oleh semua

---

<sup>1</sup> Staf Pengajar FT Universitas Negeri Yogyakarta

penduduk, khususnya dalam buta aksara, buta angka/menghitung dan keterampilan hidup (*life skills*) esensial (Unesco, 2003 : 8).

Model evaluasi keberlanjutan SD ke SMP dalam rangka wajib belajar 9 tahun bertujuan untuk: (1) menguji model pengukuran kapabilitas siswa SD untuk belajar ke SMP; (2) menguji model struktural alat keberlanjutan SD ke SMP; dan (3) mengevaluasi sistem keberlanjutan SD ke SMP dari aspek: (a) input alat seleksi PSB dan kemampuan awal siswa; (b) proses PSB dan layanan belajar setelah siswa diterima di SMP; (c) produk belajar yang berupa standar kompetensi yang ingin dan telah dicapai sekolah

Pemecahan masalah penelitian ini dilakukan dengan menggunakan model evaluasi responsif dari Stake dan CIPP (*context, input, process and product*) dari Stufflebeam. Data penelitian terdiri dari data kuantitatif dan kualitatif yang diambil secara *cross sectional*. Sampel diambil menurut strata mutu sekolah dengan ukuran 8 orang kepala SMP, 134 siswa SD kelas VI tahun 2005/2006, 764 siswa SMP kelas IX tahun 2005/2006 dan 630 siswa SMP kelas VII tahun 2006/2007.

Data penelitian dikumpulkan dengan metode: tes, dokumentasi, observasi, kuesioner dan *focus group discussion*. Tes terdiri dari tes prestasi belajar dan tes potensi belajar. Dokumen prestasi siswa meliputi nilai rapor siswa sejak kelas 4 SD, nilai USD, nilai tes seleksi SMP, nilai Ujian Nasional SMP dan nilai tes standarisasi mutu SMP. Observasi dilakukan untuk mengamati motivasi belajar dan kuesioner digunakan untuk mengetahui potensi ekonomi sebagai pendukung kapabilitas siswa SD untuk belajar di SMP. *Focus group discussion* dilaksanakan untuk mengevaluasi sistem keberlanjutan SD ke SMP berdasarkan respon Kepala Sekolah sebagai responden evaluasi responsif.

Analisis data model evaluasi keberlanjutan SD ke SMP dilakukan melalui beberapa tahap. Pada tahap awal, analisis data dilakukan menggunakan *Chi Square*, regresi linier, dan Greenhouse-Geisser untuk menguji asumsi analisis sebelum penggunaan statistik inferensial. Analisis data utama menggunakan *Structural Equation Modeling (SEM)* untuk menguji model pengukuran dan model struktural evaluasi kapabilitas siswa SD untuk belajar ke SMP dan evaluasi alat keberlanjutan SD ke SMP. *One factor repeated-measures analysis of varian* digunakan untuk menguji stabilitas prestasi siswa selama tujuh semester (dari SD kelas IV sampai SMP kelas VII semester 1). Pada tahap akhir, dilakukan analisis *post hoc* anova menggunakan metode Bonferroni untuk mengetahui perbedaan masing-masing nilai yang diulang setelah hipotesis alternatif didukung data Analisis data kualitatif dilakukan dengan membandingkan antara kondisi yang ada dengan yang diharapkan kemudian dibandingkan lagi dengan standar yang sudah ditetapkan.

## 2. Kajian Teori

Baumgartner (1973), menjelaskan bahwa evaluasi merupakan sebuah proses pengambilan keputusan dinamik yang difokuskan pada perubahan yang telah dibuat. Proses evaluasi melibatkan pengumpulan data pengukuran yang sesuai, pertimbangan nilai menurut standar yang ditetapkan dan membuat keputusan yang berbasis pada data. Evaluasi berfungsi untuk memfasilitasi keputusan yang rasional

dan objektif. Griffin (1991: 5) menjelaskan bahwa secara umum, evaluasi dilibatkan dalam pembuatan pertimbangan berharga.

Evaluasi keberlanjutan (*sustainability evaluation*) merupakan bagian dari evaluasi produk yang berada dalam lingkup model evaluasi CIPP (*contexs, input, procces, product*). Evaluasi produk itu sendiri terdiri dari empat jenis yaitu *impact, effectiveness, sustainability* dan *transportability*. *Sustainability evaluation assesses the extent to which a program,s constributions are successfully institutionalized and continued over time* (Stufflebeam, 2002: 12). Apabila konsep *sustainability evaluation* diadopsi dalam program pendidikan dapat bermakna penilaian sebuah program pendidikan yang mampu memberi kontribusi pada kesuksesan pendidikan dan keberlangsungannya dari waktu ke waktu.

Evaluasi keberlanjutan SD ke SMP yang dimaksud dalam penelitian ini adalah evaluasi yang digunakan untuk memprediksi kemampuan siswa dan menetapkan keberlangsungan pelayanan siswa pada studi tahap berikutnya sesuai dengan kemampuan yang dimilikinya. Prediksi keberlanjutan kemampuan siswa ditetapkan setelah siswa menunjukkan kemampuan yang stabil meskipun diukur menggunakan alat ukur yang berbeda secara berkesinambungan dan komprehensif. Penilaian berkesinambungan yaitu materi yang dinilai sesuai tingkat kompetensi yang harus dimiliki pada tiap – tiap jenjang kelas atau semester. Penilaian secara komprehensif yaitu penilaian yang menggunakan alat pengukuran ganda.

Seorang siswa yang akan memasuki tugas belajar baru mempunyai sejarah khusus pengembangan belajar sebelumnya. *Prerequisite* belajar diperlukan agar kemampuan siswa setara untuk memasuki tugas belajar baru. *Cognitive entry behavior* merupakan istilah lain untuk menjelaskan tipe-tipe *prerequisite* pengetahuan, keterampilan dan kompetensi yang esensial untuk belajar pada tugas-tugas baru. Dalam *responsive evaluation, entry behavior* termasuk dalam klasifikasi data *antecedent* yang dapat berupa status seorang siswa sebelum mengikuti pelajaran seperti: bakat, pengalaman sebelumnya, minat dan kemauan.

Bloom (1976) yang dikutip oleh Roid and Haladyna (1982: 16) menjelaskan bahwa prestasi yang diukur dengan tes acuan kriteria mempunyai fungsi untuk menetapkan *cognitive entry characteristics, affective entry characteristics* dan kualitas pembelajaran. Model linier yang digambarkan oleh Bloom untuk menjelaskan tentang pencapaian prestasi adalah sebagai berikut:

$$Y = F(X1, X2, X3)$$

Dimana **Y** adalah prestasi yang diukur dengan test acuan kriteria, **X1** adalah *cognitive entry characteristics*, **X2** adalah *affective entry characteristics* seperti sikap dan kepribadian, dan **X3** adalah kualitas pembelajaran. Bloom's menganalisis tiga konstruk yang berhubungan dengan belajar yaitu *cognitive entry characteristics* menyumbang varian 50%, *affective entry characterictic* menyumbang 25%, sedangkan sisanya dipengaruhi oleh kualitas pembelajaran sekitar 25%. Apabila semua faktor tersebut dijumlahkan sedikitnya ada 90% untuk semua varian, 10% sisanya adalah sejarah keluarga, lingkungan rumah, latar belakang lain yang berpengaruh pada pada belajar siswa. *Cognitive* dan *affective entry characteristics* bekerja secara unik dan bersama-sama membentuk basis kapabilitas siswa.

Konstruk belajar dari Bloom menjadi dasar teori dalam penelitian ini. Asumsi yang menjadi pedoman dalam konsep keberlanjutan adalah hasil belajar dibangun dari berbagai dimensi yaitu kemampuan akademik, sikap dan kualitas pembelajaran. Apabila salah satu dimensi tidak terpenuhi secara sempurna maka dimensi yang lain perlu ditingkatkan. Dalam rangka wajib belajar 9 tahun, layanan pendidikan perlu memperhatikan karakteristik input yang tidak hanya diukur dari aspek kognitif saja tetapi juga perlu mempertimbangkan aspek afektif. Untuk memperoleh hasil pendidikan yang lebih bermakna, anak-anak yang mempunyai kemampuan kognitif rendah perlu dilayani dengan pembelajaran yang berkualitas baik supaya mencapai hasil belajar yang optimal.

Keberlanjutan pendidikan dari suatu jenjang pendidikan menuju ke jenjang pendidikan berikutnya banyak menggunakan alat seleksi. Berikut ini dikutip beberapa hasil penelitian yang mendalami prediksi alat seleksi penerimaan siswa/mahasiswa baru. Hasil penelitian Soetarlinah (1997: 46) menemukan skor tes bakat sekolah (TBS) dapat memprediksi nilai rapor. Pada siswa SMU, TBS mempunyai koefisien korelasi sebesar 0,35 terhadap nilai rata-rata dua mata pelajaran yaitu matematika dan bahasa Indonesia. Pada siswa SMK, TBS berkorelasi positif terhadap mata pelajaran bahasa Indonesia sebesar 0,44. Heri Widodo (2004: 87) meneliti daya prediksi TPA terhadap Indeks Prestasi Kumulatif (IPK) mahasiswa Universitas Sanata Dharma Yogyakarta. Secara umum skor komposit TPA berkorelasi positif dengan IPK pada sembilan jurusan dari sepuluh jurusan yang diteliti.

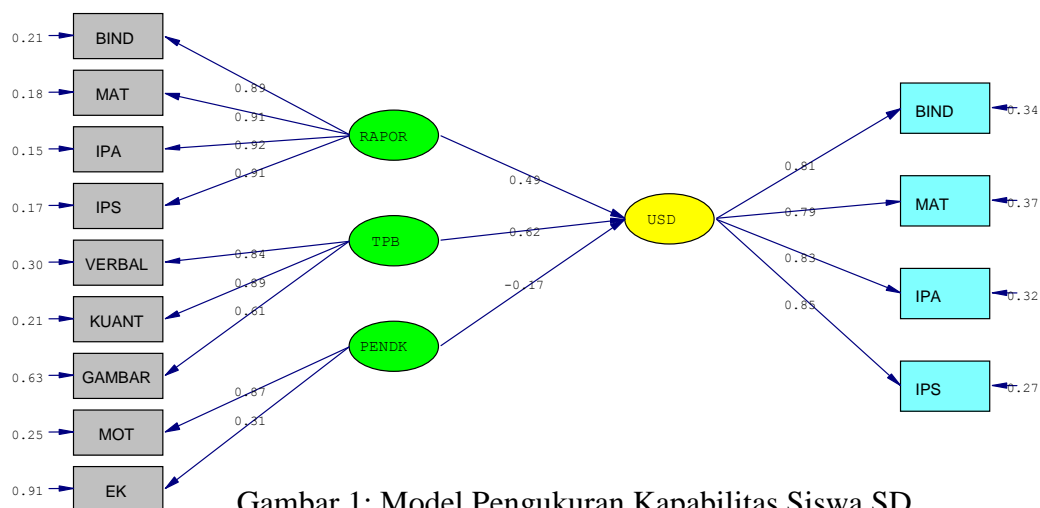
Friedenberg, (1995: 303) menulis ada empat dampak yang mungkin terjadi ketika sebuah keputusan dibuat melalui studi efisiensi seleksi dalam evaluasi keberlanjutan yaitu: (1) seseorang akan diterima pada basis skor prediktor dan berpenampilan baik sesuai kriteria (sebuah prediksi akurat atau hit, dinamakan *true positive*); (2) seseorang akan diterima pada basis skor prediktor dan berpenampilan di bawah kriteria (sebuah prediksi tidak akurat atau *miss*, dinamakan *false positive*); (3) seseorang akan ditolak pada basis skor prediktor dan secara potensial mempunyai penampilan baik pada kriteria (sebuah prediksi tidak akurat atau *miss*, dinamakan *false negative*); (4) seseorang akan ditolak pada basis skor prediktor dan secara potensial mempunyai penampilan lebih rendah dari kriteria (sebuah prediksi akurat atau *miss*, dinamakan *true negative*).

### **3. Hasil Dan Pembahasan**

#### **Pengukuran Kapabilitas Siswa SD untuk Belajar ke SMP**

Kapabilitas siswa SD yang akan belajar ke SMP dievaluasi dari tiga variabel laten eksogen yaitu nilai rerata rapor (RAPOR), tes potensi belajar (TPB), dan potensi pendukung (PENDK). Variabel laten endogen kapabilitas belajar adalah nilai Ujian Sekolah Daerah (USD) yang terdiri dari mata pelajaran Bahasa Indonesia (BIND), Matematika (MAT), IPA dan UAS IPS. Nilai rerata rapor (RAPOR) terdiri dari empat variabel manifes yaitu Bahasa Indonesia (BIND), Matematika (MAT), Ilmu Pengetahuan Alam (IPA) dan Ilmu Pengetahuan Sosial (IPS). Tes Potensi Belajar (TPB) dibangun dari tiga variabel manifes yaitu tes verbal (VERBAL), kuantitatif (KUANT) dan gambar (GAMBAR). Potensi pendukung terdiri dari dua variabel manifes yaitu motivasi belajar (MOT) dan potensi ekonomi (EK). Pengukuran kapabilitas belajar dilaksanakan terhadap 134 siswa SD kelas VI tahun ajaran

2005/2006 yang berasal dari 4 sekolah. Hasil analisis model pengukuran kapabilitas belajar dapat disimak pada Gambar 1 berikut ini:



Gambar 1: Model Pengukuran Kapabilitas Siswa SD untuk Belajar ke SMP

Kapabilitas siswa SD untuk belajar di SMP yang ditetapkan dengan nilai USD dipengaruhi oleh beberapa faktor yaitu nilai rerata rapor dan potensi belajar. Hasil analisis model pengukuran menunjukkan variabel laten potensi pendukung (ekonomi dan motivasi) berkorelasi negatif dengan USD, meskipun secara terpisah, motivasi belajar siswa menjadi indikator yang baik pada variabel laten pendukung.

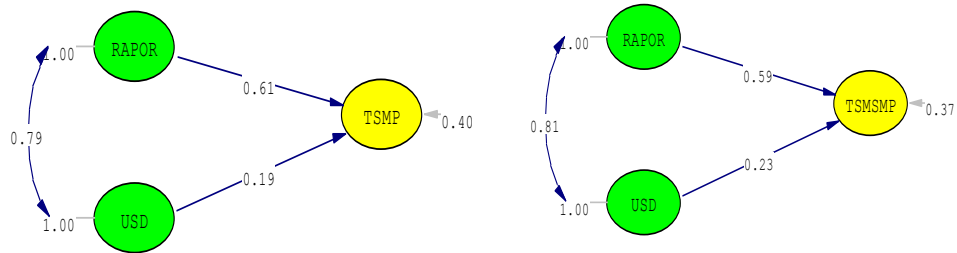
Hasil analisis model pengukuran kapabilitas belajar menunjukkan variabel manifes tes gambar dan potensi ekonomi memiliki muatan faktor yang rendah yaitu  $< 0,7$ . Hal ini menunjukkan bahwa dua variabel manifes tersebut kurang baik sebagai indikator pada variabel laten Tes Potensi Belajar (TPB) dan potensi pendukung (PENDK).

Analisis *Maximum Likelihood* dengan program LISREL memperoleh model persamaan struktural sebagai berikut:  $\eta_1 = \gamma_{11} \xi_1 + \gamma_{12} \xi_2 + \gamma_{13} \xi_3 + \zeta_1$  atau  $USD = 0,62*TPB + 0,49*Rapor - 0,17*PENDK + 0,077$ . Model persamaan struktural tersebut bermakna TPB mempunyai pengaruh yang lebih kuat terhadap USD (0,62) daripada pengaruh nilai rapor terhadap USD (0,49). Variabel laten pendukung (PENDK) berpengaruh negatif (-0,17) terhadap USD tetapi berpengaruh positif terhadap TPB (0,73) dan RAPOR (0,79). TPB dan RAPOR juga mempunyai hubungan yang sangat kuat yaitu 0,91. Hasil analisis *Goodness of Fit Statistic* menunjukkan model dinyatakan fit atau didukung data berdasarkan kriteria *Normed Fit Index* (NFI) = 0,90, *Comparative Fit Index* (CFI) = 0,94 dan *Incremental Fit Index* (IFI) = 0,94 yang lebih besar atau sama dengan kriteria model fit 0,9.

### Model Keberlanjutan SD ke SMP

Model keberlanjutan SD ke SMP disusun dari model struktural alat seleksi Penerimaan Siswa Baru (PSB) terhadap prestasi belajar SMP. Alat PSB menjadi variabel laten eksogen yang terdiri dari nilai rerata rapor SD (RAPOR) dan gabungan nilai Ujian Sekolah Daerah (USD). Prestasi belajar SMP menjadi

variabel laten endogen yang ditetapkan pada nilai tes standar mutu SMP (TSMP) semester 1. Hasil analisis model struktural prediksi nilai rerata rapor SD dan nilai USD terhadap nilai tes standar mutu SMP dapat ditampilkan pada Gambar 2.



Gambar 2a: Model Struktural Prediksi Nilai Rerata Rapor SD dan Nilai USD terhadap Tes Standar Mutu SMP

Gambar 2b: Model Struktural Prediksi Nilai Rerata Rapor SD dan Nilai USD terhadap Tes Standar Mutu SMP (+ data longitudinal)

Hasil analisis model struktural pada Gambar 2 dapat ditulis dengan menggunakan model persamaan matematis sebagai berikut:  $\eta_1 = \gamma_{11} \xi_1 + \gamma_{12} \xi_2 + \zeta_1$  atau  $TSMP = 0,61Rapor + 0,19USD + 0,40$ . Setelah model ditambah 73,3% data longitudinal dari sampel yang sama terjadi perubahan menjadi  $TSMP = 0,59Rapor + 0,23USD + 0,37$ . Hasil analisis model struktural menunjukkan rerata nilai rapor SD selama lima semester (RAPOR) mempunyai prediksi yang lebih tinggi ( $\gamma_{11} = 0,61$ ) terhadap nilai tes standarisasi mutu SMP kelas VII semester 1 (TSMP) sedangkan prediksi nilai USD terhadap nilai tes standarisasi mutu SMP sebesar ( $\gamma_{12} = 0,19$ ). Setelah 73,3% data longitudinal disertakan dalam analisis, koefisien  $\gamma_{12}$  meningkat 0,04 sehingga menjadi 0,23 dan  $\gamma_{11}$  menurun 0,02 sehingga menjadi 0,59.

Beberapa perubahan menyolok terjadi setelah penambahan data longitudinal adalah: (1) peningkatan *loading factor* pada variabel manifes nilai tes standarisasi mutu SMP mata pelajaran Bahasa Indonesia dari 0,42 menjadi 0,55; (2) Penurunan *loading factor* pada variabel manifes nilai tes standarisasi mutu SMP mata pelajaran IPA yaitu dari 0,70 menjadi 0,57. Hasil analisis *Goodness of Fit Statistic* menunjukkan *Normed Fit Index* (NFI) = 0,93, *Comparative Fit Index* (CFI) = 0,93 dan *Incremental Fit Index* (IFI) = 0,93 yang lebih besar dengan kriteria model fit (> 0,9). Pengecekan hasil analisis SEM dilakukan dengan metode analisis regresi dan korelasi *intersection test*. Hasil analisis regresi dapat disimak pada Tabel 1.

Tabel 1. Hasil Analisis Regresi Prediksi Rerata Nilai Rapor dan USD terhadap Tes Standar Mutu SMP

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0,578 <sup>a</sup>	0,334	0,332	2,11339

a. Prediktor: (Constant). USD, RAPORSD

Hasil analisis regresi menunjukkan prediksi USD dan rapor SD secara bersama-sama sebesar 0,578 atau berada pada kategori korelasi sedang. Hasil analisis regresi menunjukkan koefisien determinasi ( $R^2$ ) sebesar 0,334 atau dengan kata lain 33,4% prediksi nilai tes standar mutu SMP ditentukan oleh nilai rerata rapor dan USD. Hasil analisis ini menunjukkan pula bahwa 66,6% prediksi nilai tes standar mutu SMP ditentukan oleh faktor lain yang tidak diteliti. Selanjutnya, konsistensi hasil analisis dibuktikan dengan korelasi *intersection test* alat-alat pengukur prestasi siswa menggunakan program SPSS. Hasil analisis dilaporkan di Tabel 2.

Tabel 2. Validitas Prediksi *Intersection Test* Pengukur Prestasi Siswa

Variabel Penjelas	n Kasus	Variabel yang Dijelaskan		
		Rerata UN	USD	Tes SMP Sem 1
Tes Seleksi	763	0,639		
Rapor SD	626		0,70	0,546
USD	626		1	0,44
TPB	134		0,761	
Rapor SD	134		0,811	

Hasil analisis korelasi *intersection test* memberi gambaran yang menyeluruh terhadap alat seleksi yang digunakan di SMP. Menurut hasil analisis korelasi *intersection test* tersebut, tes potensi belajar (TPB) dan nilai rapor SD mempunyai korelasi tinggi sebagai prediksi USD, sedangkan USD mempunyai korelasi rendah sebagai prediksi tes standarisasi mutu SMP semester 1. Rerata rapor SD cukup konsisten sebagai prediksi TPB, USD dan tes standarisasi mutu SMP walaupun jumlah kasus yang digunakan berbeda.

Prediksi keberlanjutan sekolah dapat dimonitor dari stabilitas kemampuan atau prestasi belajar. Hasil analisis *General Linear Model* (GLM) multivariat *repeated-measure* GLM memperoleh *mean* tiap pengulangan pengukuran seperti pada Tabel 3.

Tabel 3. *Repeated* Nilai Rerata Rapor per Mata Pelajaran

Mata Pelajaran	Waktu Pengukuran (dalam dokumen nilai)						
	Kelas 4		Kelas 5		Kelas 6		Kelas 7
	Sem 1	Sem 2	Sem 1	Sem 2	Sem 1	USD	Sem 1
B. Indonesia	7,282	7,419	7,197	7,578	7,585	7,162	6,474
Matematika	7,034	7,233	6,884	7,042	7,131	6,853	5,669
IPA	7,084	7,244	7,204	7,530	7,657	7,869	7,005
IPS	6,899	7,066	6,952	7,217	7,300		6,999

Data pada Tabel 3 menunjukkan pola-pola penilaian yang diberikan guru pada semester ganjil lebih rendah dari pada nilai semester genap. Analisis multivariat dilakukan untuk menguji hipotesis nol yaitu tidak ada perbedaan rerata nilai tiap-tiap pengulangan pengukuran. Rangkuman hasil analisis *repeated measure* menggunakan metode *Sphericity Assumed* terhadap 4 mata pelajaran yang menjadi indikator prestasi belajar siswa dapat dilihat pada Tabel 4.

Tabel 4. Hasil *Test of Within-Subject Effect Measure: MEASURES\_1 Sphericity Assumed*

Source	Type III SS	df	MS	F	Sig.
B.Indonesia	542,442	6	90,407	236,565	0,00
Matematika	1064,991	6	177,498	265,164	0,00

IPA	298,887	6	49,814	129,269	0,00
IPS	77,252	5	15,45	38,238	0,00

Menurut hasil analisis *Sphericity Assumed* pada 4 mata pelajaran yang dilakukan pengulangan pengukuran diperoleh informasi bahwa prestasi nilai siswa kelas 7 SMP tidak stabil yang terbukti dari penerimaan  $H_a$  dengan  $\text{Sig.} < 0,05$ . Prestasi siswa yang paling tidak stabil terdapat pada mata pelajaran matematika dengan  $F$  paling tinggi yaitu 265,164. Penyebab perbedaan prestasi pada tiap-tiap pengukuran untuk mengetahui rerata kelompok nilai *repeated* yang berbeda dilakukan dengan metode Bonferroni. Hasil analisis *pair wise comparison* tersebut secara umum menunjukkan bahwa prestasi siswa kelas 4 SD tidak berbeda nyata dengan prestasi kelas berikutnya dalam beberapa kali pengulangan pengukuran.

### **Evaluasi Sistem Keberlanjutan SD ke SMP**

Pada tahun 2006, provinsi DIY menggunakan nilai ujian standar (USD) tingkat propinsi sebagai instrument ujian akhir SD dan alat PSB. Hasil evaluasi responsif dari para kepala sekolah terhadap model keberlanjutan SD ke SMP dengan digunakannya instruemn tersebut dapat dipaparkan sebagai berikut:

#### **Rasional**

Beberapa alasan rasional yang dipertimbangkan untuk mendukung sistem PSB SMP menggunakan alat seleksi menurut respon Kepala Sekolah antara lain:

- 1) SMP mengharapkan input siswa yang diterima cukup baik sesuai dengan keinginan dan kondisi sekolah.
- 2) Daya tampung SMP kurang memadai untuk menerima semua lulusan Sekolah Dasar.
- 3) SMP membutuhkan prasyarat belajar (*entry behaviour*) yang harus dipenuhi siswa untuk dapat mengikuti pendidikan agar kurikulum yang dibebankan ke SMP dapat dicapai.
- 4) SMP perlu mengetahui kondisi input siswa untuk menetapkan program keberlanjutan layanan pendidikan yang sesuai terutama bagi siswa yang belum memenuhi prasyarat belajar.

#### **Anteseden/input**

Input prasyarat belajar di SMP yang diharapkan memenuhi standar minimal nilai 5 tetapi hasil observasi menemukan 6,39% siswa SD mempunyai nilai matematika  $< 4$ . Judgment yang diambil adalah siswa yang tidak memenuhi prasyarat belajar perlu mengikuti matrikulasi, pendalaman materi atau penyetaraan kemampuan awal.

#### **Transaksi/proses**

- 1) SMP mengharapkan PSB dilakukan dengan beberapa variasi alat dan waktu pendaftaran. PSB yang diobservasi dilakukan serentak dengan alat seleksi yang sama. Judgment yang diambil : SMP favorit menggunakan tes prestasi belajar dengan tes potensi belajar dan diperkenankan membuka pendaftaran lebih awal. SMP yang mempunyai jumlah pendaftar kurang, dapat



menggunakan nilai rapor SD selama tiga tahun terakhir sebagai alat PSB dan memperpanjang waktu pendaftaran.

- 2) Proses pelayanan belajar yang diharapkan dilakukan dalam beberapa bentuk sesuai dengan kebutuhan siswa tetapi proses pelayanan belajar yang diobservasi hanya beberapa sekolah yang sudah menerapkan. Judgment yang diambil adalah sekolah menyelenggarakan layanan matrikulasi untuk siswa kurang mampu, *grouping atau tracking* sesuai kemampuan pada kegiatan belajar tambahan (*les*), *cooperatif learning* pada kegiatan belajar reguler, dan *peer teaching* pada kegiatan belajar kelompok.

### **Produk**

SMP mengharapkan penentuan kelulusan siswa sebagai produk belajar dilakukan secara komprehensif tetapi produk yang diobservasi penentuan kelulusan ditetapkan dengan nilai Ujian Nasional (UN). Hasil analisis deskriptif nilai UN SMP menunjukkan 5,02% siswa tidak dapat memenuhi standar kelulusan. Penyebab kegagalan terbesar ditemukan pada mata pelajaran Bahasa Inggris. Judgment yang diambil penentuan kelulusan ditetapkan dengan nilai UN dan siswa yang tidak lulus UN dapat mengikuti UN Pendidikan Kesetaraan.

### **Sustainability**

Keberlanjutan sistem PSB SMP menggunakan seleksi nilai USD dapat memenuhi keberlanjutan dari beberapa unsur yaitu: Produktivitas (*productivity*) perangkat evaluasi telah didukung Dinas Pendidikan dengan cara mengumpulkan soal-soal ujian dari tahun ke tahun. Produk alat evaluasi masa transisi SD ke SMP memenuhi indikator keamanan (*security*) karena telah menggunakan prosedur baku. Kebijakan evaluasi keberlanjutan SD ke SMP dilindungi (*protection*) oleh Keputusan Kepala Dinas Pendidikan Propinsi DIY. Kebijakan sistem PSB menggunakan USD memenuhi unsur *viability* karena lebih efisien dan dapat memberi manfaat tinggi bagi penggunanya. Model evaluasi keberlanjutan SD ke SMP menggunakan nilai USD dapat diterima (*acceptability*) dengan catatan sekolah diperkenankan menentukan kriteria lain terhadap calon siswa yang akan diterima.

### **Pembahasan**

Hasil analisis butir soal USD Bahasa Indonesia menemukan 60% butir soal mudah dan 30% butir soal memiliki daya pembeda rendah. Soal-soal yang mudah bukan merupakan alat prediksi yang baik oleh sebab itu sebaiknya dibedakan antara soal yang digunakan sebagai penentu kelulusan dan sebagai alat seleksi.

Hasil pengukuran kapabilitas siswa SD untuk belajar ke SMP pada mata pelajaran matematika diketahui 16,42% sampel siswa SD belum memenuhi prasyarat belajar karena memperoleh nilai  $< 4$  atau berada pada kategori E (tidak mampu). Menurut Bloom's (1973: 75) kesulitan tugas belajar pertama yang tidak diperbaiki dapat menyebabkan kesulitan tugas belajar selanjutnya.

Kebijakan pemerintah menetapkan standar kelulusan menggunakan nilai minimum 4,25 (tahun 2006) semestinya disertai dengan persyaratan nilai minimum yang sama pada sistem PSB. Apabila *output* ditetapkan dengan standar tetapi input tidak dikenakan peraturan yang sama, maka akan banyak terjadi penghuni sekolah

yang abadi karena mereka tidak memenuhi standar pada awal masuk maupun pada akhir sekolah.

Analisis kasus penambahan nilai rapor terjadi pada siswa yang mendapat nilai rapor 5 dan 6 terutama pada nilai rapor semester genap. Kasus ini terjadi karena guru tidak ingin siswa yang tidak mampu belajar tersebut tinggal kelas. Alasan yang dilontarkan guru cukup lugas yaitu '*biar mereka menjadi tanggungan kelas atau sekolah berikutnya yang lebih tinggi*'. Guru pada umumnya sudah berusaha melakukan perbaikan pembelajaran tetapi karena kemampuan siswa yang rendah maka hasil belajar sangat sulit ditingkatkan.

Hasil analisis reliabilitas butir subtes kuantitatif memperoleh koefisien Alpha sebesar 0,907 atau berada pada kategori sangat handal. Karakteristik butir tes kuantitatif 64% termasuk dalam kategori tingkat kesulitan sedang. Hasil analisis korelasi bivariat antara subtes kuantitatif dengan nilai mata pelajaran yang diujikan melalui USD menemukan koefisien korelasi yang cukup tinggi yaitu dengan matematika 0,663; Bahasa Indonesia 0,688 dan IPA 0,733. Berdasarkan temuan ini, subtes kuantitatif cukup baik apabila digunakan untuk mengetahui potensi belajar siswa.

Hasil penelitian ini didukung oleh Tritjahyo (2004: 70) yang meneliti tentang 'pengaruh IQ dan status sosial ekonomi (SSE) terhadap prestasi belajar siswa kelas V SD. Hasil penelitian ini menemukan bahwa IQ mempunyai korelasi yang lebih tinggi daripada SSE. Hasil analisis dengan menggunakan 246 sampel menemukan koefisien korelasi ( $r$ ) IQ dengan prestasi belajar mata pelajaran bahasa Indonesia sebesar 0,452, matematika (0,433), IPA (0,379) dan IPS (0,33). Koefisien korelasi antara SSE dengan prestasi belajar cukup rendah yaitu pada mata pelajaran bahasa Indonesia (0,319), IPA (0,180), IPS (0,158) dan matematika (0,123).

Koefisien hubungan nilai USD terhadap nilai standar mutu SMP kelas VII semester 1 lebih rendah (0,19) daripada rerata nilai rapor selama lima semester (0,61). Setelah 73,13% data longitudinal disertakan dalam analisis, koefisien  $\gamma_{12}$  meningkat 0,04 sehingga menjadi 0,23 dan  $\gamma_{11}$  menurun 0,02 sehingga menjadi 0,59. Menurut simulasi analisis yang dilakukan dengan cara menambah atau mengurangi jumlah sampel, hasil analisis selalu menunjukkan koefisien korelasi yang berubah-ubah namun gradasi korelasi tetap sama, yaitu prediksi nilai rapor selalu lebih tinggi dari nilai USD.

Prediksi rerata nilai rapor lebih tinggi terhadap nilai standar mutu SMP. Hal ini berarti nilai rapor yang diberikan guru SD dapat digunakan sebagai alat penerimaan siswa baru di SMP. Hasil penelitian ini didukung oleh Fishman dan Pasanella (1960), Hills (1964) dan Munday (1967) yang dikutip oleh Sumadi Suryabrata (1994: 26). Hasil penelitian tersebut menemukan bahwa angka rapor di sekolah menengah merupakan prediktor tunggal terbaik bagi keberhasilan belajar di perguruan tinggi. Hal ini disebabkan karena kemampuan yang diukur memiliki banyak kesamaan. Dalam artikel tersebut, beliau menjelaskan apabila nilai rapor digunakan sebagai alat seleksi maka perlu dilakukan *adjustment model* untuk meniadakan kesenjangan nilai yang diberikan antara sekolah satu dengan yang lain karena karakteristik mutu sekolah beragam.

Hasil analisis prediksi tes seleksi terhadap UN menemukan koefisien korelasi sebesar 0,639 atau berada di atas korelasi nilai rapor dan USD. Beberapa

kelemahan penggunaan tes seleksi sebagai alat penerimaan siswa baru telah diteliti oleh Ajisukmo (2004: 2). Hasil penelitian beliau antara lain menyebutkan bahwa tes seleksi memberi dampak psikologis bagi orangtua atau siswa yang memiliki nilai UAS tinggi tetapi tidak dapat diterima di sekolah negeri karena seleksi dilakukan pada waktu yang bersamaan.

Penelitian tentang prediksi alat seleksi terhadap prestasi belajar cukup sering dilakukan. Siswo Pratomo (1991: 523) mengambil beberapa kesimpulan penelitian sebagai berikut: (1) NEM SMA, TKU, dan ujian tulis Sipenmaru tahun 1988 merupakan prediktor yang meyakinkan terhadap prestasi belajar mahasiswa apabila digunakan sebagai prediktor tunggal; (2) STTB SMA bukan merupakan prediktor yang meyakinkan terhadap prestasi belajar mahasiswa; (3) sumbangan efektif masing-masing prediktor adalah ujian tulis Sipenmaru = 14,59%, NEM = 12,235%, TKU = 5,526% dan STTB = 2,96%; Hasil penelitian Sri Musrifah (1989: 796) mengambil beberapa kesimpulan yaitu: (1) intelegensi, kebiasaan belajar, pendidikan orangtua dan prestasi belajar mahasiswa non-PMDK lebih tinggi daripada mahasiswa PMDK; (2) tidak ada perbedaan prestasi belajar antara mahasiswa yang berasal dari sekolah swasta dan sekolah negeri; (3) sumbangan intelegensi, kebiasaan belajar, pendidikan orangtua terhadap prestasi belajar relatif kecil.

Pemanfaatan hasil ujian nasional sebagai dasar pembinaan dan pemberian bantuan kepada satuan pendidikan yang tidak mampu harus segera direalisasikan. Kasus yang ditemukan dalam penelitian ini, kegagalan ujian nasional ditemukan dari sekolah yang berperingkat rendah. Penyebab kegagalan berasal dari mata pelajaran Bahasa Inggris (37,5% dari 240 siswa). Kegagalan yang bersifat masal pada umumnya berasal dari guru yang kurang mampu mengajar dengan baik dan menarik. Berdasarkan temuan ini, peningkatan mutu pendidikan melalui ujian nasional perlu dibarengi dengan peningkatan mutu guru dalam inovasi pembelajaran.

## 4. Simpulan Dan Saran

### Simpulan

1. Model kapabilitas siswa untuk belajar di SMP ditetapkan dengan USD sebagai variabel kriterianya. Hasil analisis SEM memperoleh model persamaan struktural:  $USD = 0,62*TPB + 0,49*Rapor - 0,17*PENDINGK + 0,077$ . Kapabilitas siswa untuk belajar ke SMP dipengaruhi oleh variabel eksogen Tes Potensi Belajar (0,62), rerata rapor (0,49).
2. Keberlanjutan SD ke SMP ditetapkan dengan variabel kriteria nilai Standar Mutu SMP (TSMP). Model persamaan struktural prediksi alat keberlanjutan SD ke SMP adalah  $TSMP = 0,61Rapor + 0,19USD + 0,40$ . Nilai tes standar mutu SMP dapat diprediksi dari nilai USD sebesar 0,19 dan rerata nilai rapor SD 0,61. Setelah 73,3% data longitudinal disertakan dalam analisis, hubungan nilai USD meningkat menjadi 0,23 dan rerata rapor menurun menjadi 0,59.
3. Evaluasi responsif sistem keberlanjutan SD ke SMP

- a. SMP mendukung sistem PSB SMP menggunakan alat seleksi dengan pertimbangan: (a) SMP mengharapkan input siswa yang diterima cukup baik; (b) daya tampung SMP kurang memadai untuk menerima semua lulusan SD; (c) SMP membutuhkan prasyarat belajar (*entry behaviour*) yang harus dipenuhi oleh siswa baru.
- b. Anteseden (input) potensi awal yang diharapkan adalah siswa memiliki prestasi belajar minimal pada nilai 5 tetapi ada 6,39% siswa memiliki nilai matematika kurang dari 4 sedangkan standar ditetapkan pada nilai 6. *Judgment*: menyetarakan kemampuan awal siswa yang belum memenuhi prasyarat belajar melalui kegiatan matrikulasi.
- c. Transaksi (proses) PSB SMP diharapkan bervariasi tetapi PSB yang diterapkan serentak dalam satu waktu. Proses PSB SMP yang standar dapat memenuhi azas obyektif, transparan, akuntabel, tidak diskriminatif dan kompetitif. *Judgment*: (a) sekolah favorit membuka pendaftaran dan seleksi lebih awal; (b) SMP peringkat sedang membuka pendaftaran secara serentak; (c) SMP yang belum memperoleh siswa baru sesuai *quota* dapat memperpanjang waktu PSB.
- d. Produk belajar yang diharapkan adalah siswa dapat mencapai standar kompetensi lulusan tetapi 5,02% siswa belum mencapai standar nilai 4,25 (tahun 2006). *Judgment*: siswa yang tidak memenuhi standar kelulusan dapat mengikuti Ujian Nasional Pendidikan Kesetaraan (UNPK).

## Saran

### Saran untuk Dinas Pendidikan

- 1) Menyempurnakan sistem keberlanjutan SD ke SMP yang dapat memenuhi azas obyektif, transparan, akuntabel, tidak diskriminatif dan kompetitif dengan menggunakan alat pengukuran ganda dan standar seleksi ganda pada sekolah-sekolah khusus.
- 2) Dinas Pendidikan tidak menetapkan nilai USD sebagai alat seleksi tunggal pada SMP karena USD kekuatan prediksinya rendah.
- 3) Menambah jumlah daya tampung siswa baru di SMP negeri dan pemeratakan mutu pendidikan sampai ke wilayah desa supaya jumlah pendaftar SMP menyebar ke seluruh sekolah.
- 4) Nilai standar digunakan sebagai penentu kelulusan SMP bagi siswa yang mempunyai kemampuan normal sedangkan siswa yang mengalami kesulitan belajar (*learning difficulty*) sejak dini sebaiknya menggunakan standar kelulusan yang berbeda.

### Saran untuk Sekolah

- 1) SMP mempertimbangkan nilai rerata rapor SD selama tiga tahun terakhir sebagai persyaratan masuk dan evaluasi input siswa.
- 2) Tidak semua siswa usia sekolah yang menjadi sasaran wajib belajar 9 tahun memiliki kapabilitas untuk belajar ke SMP, oleh sebab itu SMP negeri wajib melayani kebutuhan belajar siswa dari beragam kemampuan supaya semua peserta didik dapat *survive* belajar di SMP. SMP perlu menyetarakan kemampuan awal siswa yang belum memenuhi prasyarat belajar melalui kegiatan pendalaman materi, matrikulasi dan penyataran kemampuan. SMP yang mempunyai animo

pendaftar tinggi perlu mendampingi alat PSB yang menggunakan tes prestasi belajar dengan tes potensi belajar; SMP yang mempunyai jumlah pendaftar kurang, dapat menggunakan nilai rapor SD selama tiga tahun terakhir sebagai alat PSB.

SMP dapat menerapkan empat model pembelajaran yaitu *remedial learning* bagi siswa kurang mampu, *grouping atau tracking* sesuai kemampuan pada kegiatan belajar tambahan, *cooperatif learning* pada kegiatan belajar reguler dan *peer teaching* pada kegiatan belajar kelompok.

- 3) SMP perlu menyaring calon siswa untuk mengetahui karakteristik *entry behaviour* dan menggunakan data hasil seleksi penerimaan siswa baru untuk menetapkan keberlanjutan layanan belajar.
- 4) Guru dapat meningkatkan ketelitian nilai rapor sampai dengan satu angka desimal di belakang koma supaya dapat membedakan kemampuan antar siswa.

### **Saran untuk Kalangan Akademisi**

- 1) Hasil analisis model pengukuran menemukan sub tes gambar kurang baik sebagai indikator potensi belajar. Apabila ingin mengukur potensi belajar siswa sebaiknya menggunakan subtes kuantitatif.
- 2) Peneliti lain dapat mengembangkan model struktural kapabilitas belajar dengan kombinasi beberapa macam variabel yang berbeda.

### **Pustaka Acuan**

- Bamberger, M. & Cheema, S. (1993). *Case study of project sustainability-Implication for policy and operation from Asian experience (2<sup>nd</sup>)*. Washington DC: The World Bank.
- Baumgartner, T. A., & Jackson, A. S. (1995). *Measurement for evaluation: In physical education and exercise science (5<sup>th</sup> ed.)*. Madison: WCB
- Bloom, B. S. (1976). *Human characteristic and school learning*, New York: McGraw-Hill book Company
- Friedenberg, L. (1995). *Psychological testing, design, analysis and use*. Boston: Allyn and Bacon.
- Griffin, P., & Peter, N. (1991). *Educational assessment and reporting*. Sidney: Harcourt Brace Javanovich Publisher.
- Heri, W. (2004). Kegunaan tes potensi akademik plus pada mahasiswa program studi Ilmu Pengetahuan Alam, Universitas Sanata Dharma Angkatan 1999/2000. *Jurnal Penelitian*. No. 14, Mei 2004.
- Djemari Mardapi. (1999). Evaluasi pelaksanaan ebtanas. *Laporan Penelitian*. Puslitbangsisjian, Balitbang Dikbud
- (2004). Dampak ujian akhir nasional. *Laporan Penelitian*. Jakarta: Puspendik Balitbang Diknas
- Ridaura, S. L., Leffelaar, P. A., Van Ittersum, et all. (2003). Sustainability evaluation: A systemic, multi-scale framework for design and evaluation of alternatives for peasant agriculture (versi elektronik). *The Evaluation Exchange*, Vol. 9, Number 3, Fall 2003. Publikasi oleh Harvard Family Research Project.

- Roid, G. H & Haladyna, T. M. (1982). *A technology for test – item writing*. New York: Academic Press, Inc.
- Siswo Pratomo & Sumadi Suryabrata. (1991). Validitas prediktif NEM SMA, STTB SMA, TKU, dan nilai ujian tulis Sipenmaru tahun 1988 sebagai prediktor prestasi belajar mahasiswa Fakultas non eksakta Universitas Gajah Mada. *Berkala penelitian Pasca Sarjana, UGM seri A: Kelompok Ilmu Pengetahuan Sosial dan Humaniora*. Jilid 4, Nomor 3A. 1991. hlm 517 – 525
- Stufflebeam, D. L. (2002). *CIPP evaluation model checklist*. Diambil pada tanggal 16 Maret 2005 dari <http://www.wmich.edu/evalctr/checklists>
- Sumadi, S. (1994). Seleksi calon mahasiswa baru di perguruan tinggi yang sekarang dan kemungkinannya untuk masa yang akan datang. *Laporan Seminar Pengkajian Ujian Saringan Masuk ke Perguruan Tinggi*. Jakarta: Depdikbud Puslitbangsisjian.
- Unesco. (2003). *Global education digest 2003: Comparing education statistic across the world*. Montreal: The UNESCO Institute for Statistics.