

## Prediksi Tes Bakat Skolastik Terhadap Prestasi Belajar Siswa Sekolah Menengah Atas

*Prediction of Scholastic Talent Test on Student Achievement in Senior High School*

**Idwin Irma Krisna**

Pusat Penilaian Pendidikan, Balitbang Kemdikbud  
[idwinirma@gmail.com](mailto:idwinirma@gmail.com)

Naskah diterima 02/05/2018; direvisi 06/05/2018; disetujui 09/05/2018

---

**Abstract.** This study aims to estimate predictive validity of Scholastic Aptitude Test (TBS) and to describe aptitudes scholastic in each subset of TBS (*benchmarking*). The data used in this study is secondary data from the Center for Education Assessment. The data was result from selection test result of scholarship for Madrasah Aliyah Negeri (MAN). Predictive analysis was carried out using structural equation modelling using Lisrel 8 and benchmarking process was begun by estimating person's ability with Item Response Theory (IRT), followed by item mapping using ASP.net web server technology with program c#. The results of this study are as follow. (1) predictive model of TBS is fit with empirical data. The variance that is described by scholastic aptitude to predict individual academic performance in high school is 14% and TBS tend to be a better predictor in social science program (high school) than in natural science. Benchmarking process generates four benchmarks on each subtest. Each bench has different potency characteristics. Characters which differentiate potency for each benchmark on verbal subtest are degree of generality of the word, word association pattern and cognitive activities. Characters which differentiate potency for each benchmark on quantitative subtest are number sequence patterns, mathematical operation structure and number, and geometric shapes. The nature of the premises, relation characteristics between concept and cognitive process are the main factors which discriminate each benchmark in reasoning subtest.

**Keywords:** TBS, benchmarking, potency characteristics

**Abstrak.** Penelitian ini bertujuan untuk mengestimasi validitas prediktif Tes Bakat Skolastik (TBS) dan mendeskripsikan bakat skolastik pada setiap subtes TBS (*benchmarking*). Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Pusat Penilaian Pendidikan berupa data seleksi penerimaan beasiswa masuk Madrasah Aliyah Negeri (MAN). Analisis prediksi menggunakan model persamaan struktural dengan program Lisrel 8.80 dan proses benchmarking yang diawali dengan estimasi kemampuan orang menggunakan model Item Response Theory, selanjutnya pemetaan butir soal menggunakan teknologi web server ASP.net dengan program c#. Hasil penelitian menunjukkan model prediksi TBS sesuai dengan data empiris. Varian yang dijelaskan oleh potensi bakat skolastik terhadap prestasi seseorang adalah 14 % dan prediksi pada jurusan IPS cenderung lebih tinggi dibandingkan dengan jurusan IPA. Proses benchmarking menghasilkan empat bench pada setiap subtesnya. Setiap bench mempunyai karakteristik potensi yang berbeda. Karakteristik yang membedakan potensi antar bench pada subtes verbal adalah perbedaan derajat keumuman pemakaian kata-kata, pola hubungan kata dan aktifitas kognitif. Sementara itu pada subtes kuantitatif yang membedakan karakteristiknya adalah pola deretan angka, bentuk operasi matematika dan bilangan, dan bentuk bangun geometri. Pada subtes penalaran karakteristik yang membedakan adalah sifat premis, karakteristik hubungan antar konsep dan proses

Kata kunci : TBS, benchmarking, karakteristik potensi

## PENDAHULUAN

Tes Potensi Belajar merupakan seperangkat alat tes yang mengukur kemampuan kognitif potensial umum yang dimiliki oleh seseorang. Tes potensi dirancang untuk memprediksi kemampuan seseorang jika yang bersangkutan diberikan kesempatan untuk melanjutkan belajar ke jenjang yang lebih tinggi. Alasan yang mendasari pengembangan Tes Potensi Belajar adalah penggunaan tes prestasi dalam sistem seleksi belum memberikan informasi potensi akademik secara adil. Tes prestasi dirancang untuk mengukur hasil dari program pembelajaran atau pelatihan yang dilaksanakan pada kondisi yang terkontrol (Anastasi, 1988: 411). Tes prestasi disusun berdasar silabus mata pelajaran pada suatu jenjang pendidikan atau pelatihan. Dengan demikian maka tes prestasi sangat tergantung dengan pencapaian kurikulum yang diterima di sekolah. Kondisi ini tentunya akan menjadi tidak adil pada siswa yang berasal dari sekolah dengan sarana dan prasarana yang terbatas.

Tes potensi lebih mengukur kemampuan yang bersifat spesifik yang terbentuk dari interaksi antara karakteristik individu dan kesempatan belajar di lingkungannya. Tes potensi memberikan kontribusi yang besar dalam memprediksi performans seseorang pada jenjang pendidikan yang lebih tinggi atau pada situasi yang baru (Lohman, 2004: 43). Pusat Penilaian Pendidikan (Puspendik), Badan Penelitian dan Pengembangan, Kementerian Pendidikan dan Kebudayaan telah mengembangkan Tes Potensi Belajar sejak tahun 1990 dan sejak tahun 2000 tes potensi akademik di beri nama Tes Bakat Skolastik (TBS). TBS digunakan sebagai alat seleksi penerimaan mahasiswa baru di beberapa PTN dan PT swasta serta penerimaan calon pegawai baru instansi swasta dan di beberapa kementerian. Pada jenjang Sekolah Menengah Atas (SMA), TBS dapat digunakan untuk keperluan peminatan karena tes tersebut mengukur bakat akademik. TBS yang dikembangkan oleh Puspendik sedikit banyak mengacu pada pengembangan *Scholastic Aptitude Test* (SAT). TBS terdiri atas tiga subtes yaitu: verbal, kuantitatif dan penalaran. Ketiga subtes mengukur kemampuan yang sama yaitu penalaran. Sebagai alat seleksi, aspek validitas prediktif menjadi hal yang penting sehingga TBS terbukti akurat dalam memprediksi prestasi seseorang pada jenjang yang lebih tinggi.

Pelaporan hasil seleksi akan menjadi lebih bermanfaat apabila disertai dengan gambaran kemampuan seseorang pada nilai yang diperoleh. Penetapan interpretasi hasil tes tergantung pada definisi isi, level dan *cutoff score*, secara spesifik

menggambarkan deskriptor kemampuan yang akan digunakan sebagai acuan bagi penentu kebijakan (Ferrara, et.al., 2011: 5). *Benchmarking* di rancang sebagai batasan seseorang yang ingin meningkatkan kualitas dan produktivitas dalam berkarir. Salah satu komponen penting dalam *benchmarking* adalah standar performans berdasarkan respon peserta tes dalam menjawab pertanyaan (Resnick, Nolan & Resnick, 1995: 454). Standar performansi akan memberikan gambaran proses kognitif yang dimiliki seseorang yang diperoleh berdasarkan respon pada butir soal yang diberikan.

Berdasarkan latar belakang yang telah diungkapkan di atas, pertanyaan yang diajukan dalam penelitian ini adalah:

1. Bagaimana validitas prediktif TBS terhadap keberhasilan seseorang pada jenjang pendidikan SMA?
2. Bagaimana *Performance-Level Descriptors* (PLDs) pada setiap subtes TBS?

## KAJIAN PUSTAKA

### 1. Inteligensi dan Aptitude (Bakat)

Inteligensi dan bakat merupakan kemampuan kognitif yang dimiliki oleh setiap individu (Cohen & Swerdlik, 2002: 257, 301). Inteligensi mengacu pada kemampuan intelektual secara umum yang berfungsi pada berbagai bidang prestasi sedangkan bakat merupakan kemampuan yang lebih spesifik yang hanya digunakan pada bidang prestasi tertentu saja (Berk, 2000: 316-319). Perkembangan teori intelegensi yang diawali oleh Raymond B.Cattell dan mendapat perhatian yang besar dari pengembang dan pengguna tes (dalam Cohen & Swerdlik, 2009: 283). Teori inteligensi ini diawali dengan pendapat Cattell mengenai dua tipe utama dari kemampuan kognitif yaitu fluid intelligence dan crystallized intelligence. Kedua faktor ini dapat dianalogikan dengan air. Air dengan wujud fluid jika diisikan pada suatu wadah (container) maka bentuknya akan mengikuti wadah tersebut. Berbeda halnya dengan air yang berbentuk kristal, bentuknya akan tetap.

Fluid intelligence (gf) merupakan kemampuan yang tidak tergantung pada budaya, pendidikan dan pengalaman, inteligensi ini bersifat bawaan. *Crystallized intelligence* (gc) merupakan ketrampilan dan pengetahuan seseorang yang dipengaruhi oleh pengalaman, budaya dan pendidikan baik formal maupun informal serta kemampuan untuk menerapkan pengetahuan tersebut pada situasi yang spesifik. Sebuah tugas gf bisa menjadi gc atau sebaliknya, sehingga kedua intelegensi ini mempunyai hubungan yang kuat. Sebuah fakta menunjukkan, kedua

inteligensi ini mempunyai hubungan yang kuat dengan  $r = 0,5$  (Gregory, 2007: 180). Kedudukan kedua inteligensi ini disajikan dalam sebuah spektrum kemampuan (Cronbach, 1984: 253), *g<sub>c</sub>* terletak pada spektrum tertinggi, *g<sub>c</sub>* pada spektrum terendah dan *aptitude* berada diantaranya. Berdasarkan penjelasan di atas maka antara inteligensi dan *aptitude* mempunyai hubungan yang sangat dekat. Inteligensi, *aptitude* dan pengetahuan yang dimiliki sebelumnya merupakan faktor kognitif yang mempunyai peran penting pada prestasi akademik seseorang (Horn, et.al, 1993).

## 2. Tes Bakat Skolastik

TBS dirancang untuk mengukur kemampuan kognitif secara umum yang dilandasi dengan pemahaman inteligensi. TBS mengukur kemampuan bernalar baik dalam verbal maupun numerik. Penelitian Olatoye & Aderogba (2011), numerik dan verbal secara bersama dapat menjelaskan varian sebesar 38,8% pada tes *aptitude* dan korelasi antara verbal dan numerik 0,713. Kemampuan numerik mempunyai kesamaan domain kemampuan verbal dan *aptitude* umum. Beberapa penelitian menunjukkan bahwa tes potensi mempunyai hubungan dengan tes intelegensi. Hasil penelitian Frey & Detterman, korelasi nilai SAT dengan beberapa tes IQ berkisar antara 0,53 - 0,83, dengan hasil ini memberikan bukti yang kuat bahwa SAT juga dapat sebagai tes inteligensi. Salah satu keunggulan dari tes potensi adalah kestabilan dari skor.

Tahap pengembangan butir soal TBS dimulai dari pengembangan butir soal yang meliputi penulisan kisi-kisi, validasi kisi-kisi, penulisan soal, telaah soal secara kualitatif, perakitan instrumen ujicoba dan selanjutnya mengujicobakan instrumen tersebut. Langkah selanjutnya adalah proses analisis data sampai dengan penyimpanan butir soal yang valid di bank soal.

## 3. Validitas Prediktif

Kecermatan prediksi merupakan suatu hal yang penting dalam pengembangan tes seleksi. Validitas prediktif akan menjadi penting jika tes digunakan sebagai prediktor. Pada jenjang SMA kriteria validasi yang digunakan adalah skor performansi yang diprediksi misalnya nilai pada setiap semester. Koefisien validitas prediktif merupakan korelasi antara skor tes dengan performans pada jenjang sekolah yang lebih tinggi. Keberhasilan di perguruan tinggi di dukung oleh deskriptor pada level kemampuan, *cutoff score* dan *benchmark* pada pengukuran di sekolah tinggi. Menurut Camara (2013), validasi

penilaian kesiapan seseorang di SMA adalah menentukan kriteria yang dibutuhkan untuk sukses di tahun pertama (pengetahuan, kreativitas, kemampuan) sebagai acuan penilaian.

## 4. Benchmarking

Istilah *benchmarking* pada awalnya dikenal dalam dunia ekonomi dan bisnis. Namun, saat ini istilah tersebut juga digunakan di pendidikan dalam rangka meningkatkan mutu pendidikan. *Benchmarking* merupakan suatu proses yang dilakukan oleh sebuah organisasi dalam rangka mengukur atau membandingkan kinerja yang dilakukan terhadap aktivitas atau kinerja yang sama pada organisasi lain baik secara internal ataupun eksternal sehingga diperoleh gambaran tentang kualitas kinerja organisasinya.

Proses *benchmarking* berkaitan dengan proses *standard setting*, penetapan standar yang tepat dalam pendidikan berdasarkan level *benchmark* akan memberikan arahan keberhasilan siswa pada jenjang pendidikan selanjutnya. Penetapan titik batas dapat dilakukan dengan menetapkan sebuah skor yang menjadi acuan, sebagai contoh SAT menetapkan titik batas berdasarkan skor gabungan adalah sebesar 1550 (Wyatt.et.al, 2011: 13). Peserta yang mencapai *benchmark* (1550) akan mempunyai kelebihan dibandingkan dengan peserta yang tidak mencapai titik tersebut.

## 5. Standard Setting

*Standard setting* merupakan metode yang digunakan untuk mendefinisikan level kemampuan seseorang dan penetapan *cutoff score* yang berhubungan dengan level tersebut (Bejar, 2008; Cizek, Bunch & Koons, 2004). Metode *standard setting* dapat dikategorikan sebagai model kontinum yang berfokus pada tes (*test centered*) atau model yang berfokus pada peserta tes atau *examinee centered* (Jaeger, 1989: 492). Dalam penelitian ini digunakan metode yang berdasarkan pada tes dan peserta yaitu *scale anchoring method*. Metode ini berkaitan dengan penyusunan deskripsi pada skala perilaku yang digambarkan dengan pemetaan soal. Pemetaan soal (*item mapping*) diawali dengan konsep "Content referencing" yang diperkenalkan oleh Bock, et.al (Kelly, 2002). *Content referencing* merupakan suatu prosedur yang menggunakan pendekatan IRT. Prosedur ini akan memetakan butir soal pada suatu skala berdasarkan probabilitas respon terhadap soal. Kumpulan soal (*item mapping*) yang berada pada nilai probabilitas sama akan menggambarkan kemampuan dan pemahaman siswa. Kelanjutan dari *content referencing* dan *item mapping* disebut sebagai *scale anchoring method* (Beaton & Allen, 1992). Pada

metode ini beberapa titik dipilih pada suatu skala dan selanjutnya mengidentifikasi soal yang sesuai pada titik tersebut. Sebuah soal akan dinyatakan sesuai dengan titik tersebut (*anchor*) jika sebagian besar siswa (misalnya dengan kriteria 65%) yang berada pada titik *anchor* dapat mengerjakan tetapi siswa yang berada pada titik di bawahnya tidak dapat mengerjakan. *Scale anchoring* dapat memberikan informasi secara acuan normatif ataupun konten mengenai pengetahuan yang dikuasai seseorang berdasarkan konstruk yang diukur (Beaton & Allen, 1992: 192). Pemilihan titik tersebut dinamakan dengan *anchor points* atau *anchor levels*. Salah satu cara penetapan *anchor points* adalah menggunakan persentil (Forsyth, 1991: 194). Empat angka yang ditetapkan sebagai *benchmark* internasional yang digunakan sebagai *anchor point* adalah persentil ke 25, 50, 75 dan 90. (Martin, Mullis, Beaton, gonzales, Smith & Kelly, 1997; Mullis, Martin, Beaton, gonzales, Smith & Kelly, 1998). Dengan analisis *scale anchoring*, *anchor item* pada setiap level diharapkan dapat membedakan antara *anchor point* yang berdekatan.

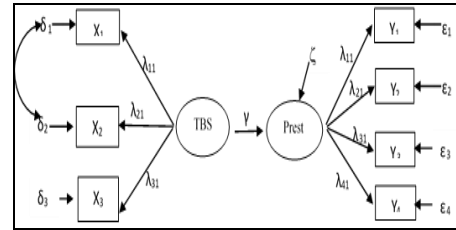
**METODE PENELITIAN**

**A. Sumber Data**

Penelitian menggunakan data sekunder yang diperoleh dari Pusat Penilaian Pendidikan Badan Penelitian dan Pengembangan Kemendikbud Jakarta. Data tersebut merupakan hasil seleksi penerimaan beasiswa di salah satu sekolah Madrasah Aliyah Negeri. Seleksi diadakan dalam rangka menjaring siswa dari pondok pesantren, madrasah tsanawiyah dan siswa SMP untuk mendapatkan beasiswa pada sekolah tersebut. Data kajian berupa respon jawaban 9840 peserta yang mengerjakan 4 paket tes. Data raport siswa semester 1 sampai dengan semester 4 diperoleh dari 349 siswa, jurusan IPA 260 orang dan jurusan IPS 62 orang.

**B. Model Penelitian**

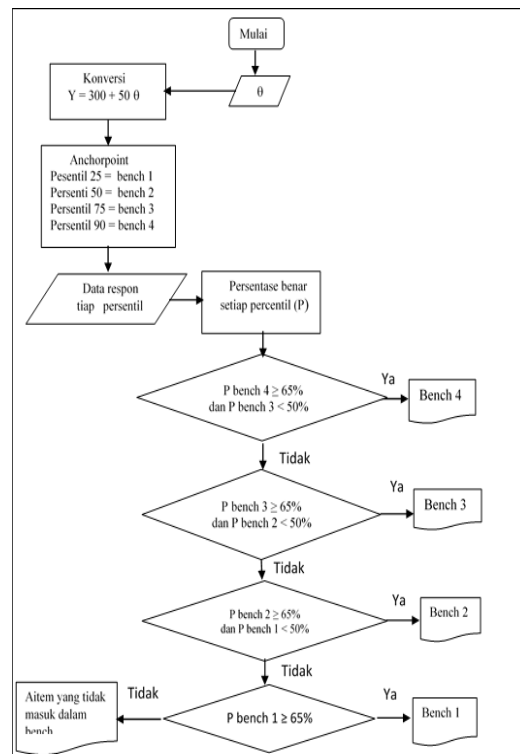
Penelitian ini menggunakan model persamaan struktural untuk membuktikan validitas prediktif TBS terhadap prestasi belajar seseorang. Analisis ini bertujuan untuk mengetahui bagaimana pengaruh variabel eksogenous TBS ( $\xi$ ) pada variabel endogenous prestasi ( $\eta$ ). Variabel yang teramati pada TBS adalah skor pada verbal (X1), kuantitatif (X2) dan penalaran (X3). Nilai IP semester 1 sampai semester 4 (Y1 – Y4 ) adalah variabel teramati pada prestasi.



**Gambar 1.** Model prediksi TBS terhadap prestasi belajar

**C. Teknik Analisis Data**

Analisis prediksi menggunakan model persamaan struktural dan proses benchmarking dengan menggunakan teknologi web server ASP.net dengan perangkat lunak lisrel 8 dan c#. Benchmarking diawali dengan menentukan kemampuan orang menggunakan pendekatan *Item Response Theory*. Proses *benchmarking* TBS seperti Gambar 2 berikut.



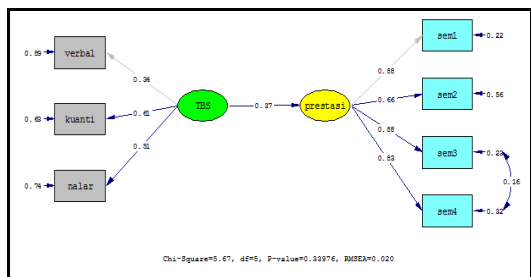
**Gambar 2.** Diagram Alur Proses Benchmark

**HASIL PENELITIAN DAN PEMBAHASAN**

**A. Hasil Analisis Dengan Model Persamaan Struktural**

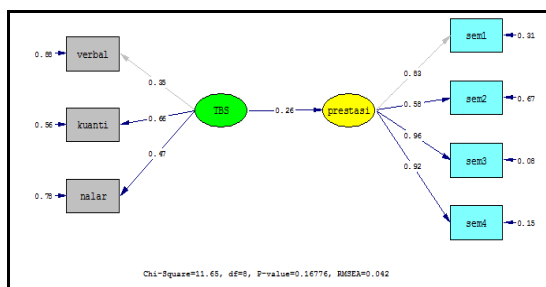
Analisis ini dimaksudkan untuk meneliti pengaruh potensi akademik terhadap prestasi belajar. Besarnya muatan faktor pada model jenjang SMA dapat dilihat pada Gambar 3. Berdasarkan hasil analisis ditemukan semua variabel menunjukkan muatan faktor yang signifikan ( $t > 1,96$ ). Muatan faktor yang terbesar pada variabel teramati dari variabel laten TBS adalah kuantitatif yaitu 0,61, sedangkan muatan

terkecil pada variabel verbal. Varian yang dijelaskan oleh potensi bakat skolastik pada prestasi belajar adalah 14%, sedangkan 86% dijelaskan oleh faktor lain. Varian yang dijelaskan oleh variabel prestasi terhadap prestasi siswa pada 4 semester menunjukkan prestasi pada semester 2 dijelaskan lebih kecil dibandingkan dengan semester lain yaitu sebesar 44%.



**Gambar 3.** Diagram Alur Hasil Analisis dengan Model Persamaan Struktural Jenjang SMA

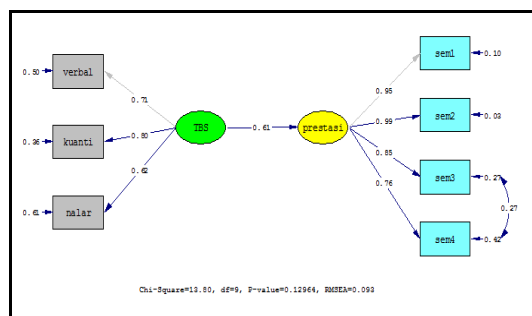
Analisis selanjutnya adalah melihat hubungan antara bakat skolastik dengan prestasi pada jurusan IPA dan jurusan IPS. Berdasarkan analisis dengan metode model persamaan struktural, variabel potensi akademik berpengaruh terhadap prestasi siswa. Besar muatan faktor pada model jurusan IPA dan IPS dapat dilihat pada Gambar 4 dan 5. Berdasarkan hasil analisis diperoleh semua muatan faktor pada model signifikan ( $t > 1,96$ ) artinya variabel teramati merupakan muatan faktor dari variabel latennya.



**Gambar 4.** Diagram Alur Hasil Analisis dengan Model Persamaan Struktural Jenjang SMA Jurusan IPA

Potensi bakat skolastik memberikan kontribusi yang paling besar pada sub kuantitatif artinya pada jurusan IPA, kuantitatif merupakan faktor terbesar yang dijelaskan oleh potensi skolastik. Varian yang dijelaskan bakat skolastik terhadap prestasi siswa pada jurusan IPA tidak terlalu besar yaitu hanya 7% dan sisanya dijelaskan oleh faktor-faktor lainnya. Pada variabel prestasi, kontribusi yang lebih besar pada semester 3 dan 4. Pada jurusan IPS, kontribusi bakat skolastik pada prestasi sebesar 37% dan dapat menjelaskan prestasi semester 1 dan 2 lebih

besar jika dibandingkan dengan semester 3 dan 4. Pada TBS, muatan faktor pada sub verbal dan kuantitatif lebih besar jika dibandingkan dengan penalaran.



**Gambar 5.** Diagram Alur Hasil Analisis dengan Model Persamaan Struktural Jenjang SMA Jurusan IPS

Berdasarkan hasil analisis, besarnya nilai prediksi tergantung pada jurusan, jurusan IPS cenderung lebih besar dibandingkan dengan jurusan IPA. Seperti halnya penelitian Lyren (2008), SweSAT mempunyai prediksi yang berbeda tergantung pada program studi yang dianalisis, program kedokteran dan hukum mempunyai prediksi yang lebih baik dibandingkan teknik sipil dan pendidikan guru. Dari analisis ternyata pengaruh TBS terhadap prestasi tidak terlalu besar, varian yang dapat dijelaskan oleh TBS cenderung kecil. Salah satu faktor yang menyebabkan hal tersebut adalah keterbatasan data dalam penelitian ini. Data yang diperoleh merupakan data siswa yang mempunyai potensi tinggi karena mereka adalah orang-orang yang lulus seleksi. Distribusi kemampuan orang pada jurusan IPA cenderung menceng kanan yang menunjukkan rata-rata kemampuan peserta cukup tinggi sehingga varian yang dijelaskan menjadi lebih kecil jika dibandingkan jurusan IPS yang mempunyai distribusi cenderung lebih simetris yang artinya peserta pada jurusan IPS mempunyai kemampuan yang lebih menyebar. Selain faktor di atas masih ada beberapa faktor atau kemampuan yang diperlukan untuk keberhasilan studi di jurusan IPA yang tidak dijadikan sebagai prediktor misalnya koordinasi motorik dan kemampuan spasial (Asrijanty, 2014: 530).

Selain itu, masih banyak faktor lain yang mempunyai kontribusi terhadap prestasi seseorang. Prediksi di perguruan tinggi akan lebih baik dengan menyertakan status sosial ekonomi sekolah dan nilai mereka sewaktu di sekolah menengah atas (Zwick & Himelfarb, 2011), kombinasi nilai sekolah di SMA, tes prestasi dan tes potensi merupakan prediktor terbaik pada nilai IP (Alnahdi, G. H., 2015: 4). Pendalaman materi juga berpengaruh pada skor

SAT, pada sub matematika akan menaikkan skor sebesar 10-20 sedangkan pada sub membaca kritis akan menaikkan skor sebesar 5-10 (College board, 2012). TBS akan menjadi prediktor yang lebih baik dengan menyertakan nilai di SMA dan tes prestasi (UN).

Muatan faktor TBS jurusan IPA cenderung lebih besar pada sub kuantitatif dan penalaran sedangkan pada jurusan IPS, subtes verbal dan kuantitatif cenderung lebih besar. Pada variabel prestasi, kontribusi yang cenderung lebih besar pada jurusan IPA adalah pada semester 3 dan 4, sedangkan pada jurusan IPS pada semester 1 dan 2. Penjelasan penulis dalam hal ini adalah: (1) seseorang pada jurusan IPA mempunyai potensi kuantitatif dan nalar yang lebih tinggi jika dibandingkan dengan seseorang pada jurusan IPS, (2) materi yang diajarkan pada semester 1 dan 2 masih bersifat umum jika dibandingkan dengan semester 3 dan 4.

Siswa SMA yang menjadi sampel dalam penelitian ini sebagian besar mempunyai latar belakang pendidikan pesantren, masih ada keterkaitan materi pelajaran pada semester 1 dan 2. Akibatnya, kontribusi prestasi pada semester 1 dan 2 lebih tinggi pada jurusan IPS. Materi pelajaran pada jurusan IPA semester 3 dan 4 lebih fokus pada pelajaran yang memerlukan potensi kuantitatif dan penalaran. Oleh karena itu, kontribusi variabel prestasi pada semester 3 dan 4 pada siswa IPA lebih besar. Pada beberapa model di atas, dilakukan modifikasi dengan mengkorelasikan kesalahan pengukuran pada variabel teramati. Misalnya, pada model persamaan struktural jenjang SMA jurusan IPS, dilakukan modifikasi pada semester 3 dan 4. Hal tersebut dilakukan dengan pertimbangan struktur atau cluster mata pelajaran saling terkait.

### **B. Performance-Level Descriptors (PLDs) pada setiap subtes TBS**

Klasifikasi TBS per subtes berdasarkan empat nilai persentil benchmark internasional (Kelly, 2002: 378) menghasilkan empat titik batas (bench) dan sekelompok butir soal pada setiap bench. Berdasarkan analisis ternyata soal-soal yang terjaring pada bench 3 dan 4 lebih banyak dibanding dengan bench 1 dan 2. Hal tersebut dapat dijelaskan karena paket tes yang dianalisis merupakan paket-paket yang digunakan untuk keperluan seleksi. Perakitan butir soal menjadi sebuah alat tes sebaiknya memperhatikan tujuan tes tersebut diselenggarakan dan dapat mengantisipasi distribusi kemampuan peserta tes. Tes yang ditujukan sebagai alat seleksi harus dapat menjaring orang-orang dengan kemampuan tinggi.

Berdasarkan klasifikasi hasil tes, langkah selanjutnya adalah penyusunan deskripsi potensi pada setiap level. Pada tahap ini peneliti dibantu oleh beberapa orang narasumber yang kompeten pada bidangnya. Sebelum pembuatan deskripsi didahului dengan langkah FGD antara peneliti dengan beberapa narasumber.

Berdasarkan deskripsi potensi setiap bench pada subtes verbal, kuantitatif dan penalaran ditemukan bahwa masing-masing bench mempunyai keunikan. Keunikan pada subtes verbal antara lain: (1) diferensiasi derajat keumuman pemakaian kata-kata dalam kehidupan sehari-hari dan kombinasi kategori pilihan jawaban, (2) pola hubungan yang terjadi pada soal analogi, (3) aktivitas kognitif mulai dari ingatan sampai dengan evaluasi pada soal wacana. Hayes (1989) menjabarkan aktifitas kognitif menjadi beberapa tahap: identifying the problem, representation of the problem, planning the solution, execute the plan, evaluate the plan, evaluate the solution. Deskriptor TOEFL IBT pada reading test antara level rendah dan tinggi mempunyai perbedaan dalam kemampuan memahami kalimat yang dinyatakan secara eksplisit atau implisit, fakta atau abstrak, dan kompleksitas suatu konsep (Gomez, et.al, 2007: 424-437). Semakin tinggi bench soal wacana pada TBS, diperlukan aktivitas kognitif tahap evaluasi terhadap solusi.

Deskripsi subtes kuantitatif dibedakan dalam sub deretan angka, aritmatika dan aljabar dan geometri. Setiap bench juga mempunyai ciri khas masing-masing, diantaranya: (1) kompleksitas pola yang membentuk deretan angka, makin tinggi bench maka pola yang ada pada deret makin kompleks, (2) operasi matematika, jenis angka, jumlah variabel dalam persamaan dan materi soal pada sub aritmatika dan aljabar, (3) bentuk bangun geometri baik dalam soal gambar atau cerita.

Namun demikian proses kognitif pada bench 3 dan 4 mempunyai kompleksitas yang hampir sama sehingga tidak terjadi peningkatan secara konsisten. Hasil penelitian yang dilakukan oleh Ferrara, et.al. (2011) juga menunjukkan bahwa deskripsi proses kognitif dan bahasa pada butir soal matematika tidak meningkat secara konsisten pada level 3, 4, dan 5. Deskriptor yang tidak dapat menggambarkan kemampuan yang harus dikuasai pada setiap level menyebabkan ketidakjelasan kemampuan apa yang harus dimiliki oleh seseorang pada setiap level.

Berdasarkan deskripsi potensi penalaran yang telah dijelaskan sebelumnya, ciri khas yang membedakan antar bench adalah: (1) sifat premis pada sub logis dan penarikan kesimpulan berdasarkan kedua premis yang diberikan, (2)

karakteristik hubungan subjek/konsep pada soal diagram, (3) proses kognitif pada sub analitis mulai dari menggunakan sampai dengan mengolah informasi untuk memperoleh solusi. Deskripsi potensi pada ketiga subtes dapat dilihat pada tabel 1 sampai.

**Tabel 1.** Deskripsi potensi subtes verbal jenjang SMA

Bench	Subtes Verbal			
	Sinonim	Antonim	Analogi	Wacana
1 270	Individu mampu mengenali padanan kata yang umum dipakai sehari-hari dari satu set pilihan kata yang juga umum dipakai sehari-hari dan satu sama lainnya cenderung tidak memiliki kemiripan arti satu sama lain.	Tidak terjaring soal	Individu mampu mengidentifikasi analogi terkait subjek/konsep-tempat, objek-produk, konsep-contoh, dan konsep-fungsi	Individu mampu mengetahui, menyebutkan informasi dari wacana yang disajikan
2 290	Individu mampu mengenali padanan kata yang tidak cukup umum dipakai sehari-hari dari satu set pilihan kata yang umum dipakai sehari-hari namun sebagiannya memiliki kemiripan arti.	Individu mampu mengidentifikasi lawan kata yang umum dipakai dari satu set pilihan kata yang di dalamnya terdapat arti kata yang nyata berbeda, umum dipakai, dan cenderung tidak ada padanan kata.	Individu mampu mengidentifikasi analogi terkait subjek/konsep-tempat (data hanya 1 soal, simpulan tidak bisa dilakukan)	Individu mampu menjelaskan kembali berbagai hal dari wacana yang disajikan
3 320	Individu mampu mengenali padanan kata yang cenderung jarang dipakai sehari-hari dari satu set pilihan kata yang umum dipakai sehari-hari dan sebagian besarnya memiliki kemiripan arti	Individu mampu mengidentifikasi lawan kata yang jarang dipakai dari satu set pilihan kata yang di dalamnya terdapat arti kata-kata yang nyata berbeda, umum dipakai, dan cenderung tidak ada padanan kata.	Individu mampu mengidentifikasi analogi yang bersifat refleksi atas suatu konsep	Individu mampu memahami, membedakan antara fakta dan opini, menginterpretasikan dan menyatakan dengan kata-kata/kalimat yang berbeda dari wacana yang disajikan
4 340	Individu mampu mengenali padanan kata yang cenderung jarang dipakai sehari-hari dari satu set pilihan kata yang jarang dipakai sehari-hari dan sebagian besar atau seluruhnya memiliki kemiripan arti.	Individu mampu mengidentifikasi lawan kata yang jarang dipakai dari satu set pilihan kata yang di dalamnya terdapat arti kata-kata yang memiliki kemiripan arti, cenderung jarang dipakai, dan tidak ada sebuah padanan kata.	Individu mampu mengidentifikasi analogi yang bersifat generik-spesifik	Individu mampu membaca cepat, mempunyai kemampuan mengingat dan menggunakan nalarnya untuk menentukan simpulan dari wacana yang disajikan.

**Tabel 2.** Deskripsi potensi subtes kuantitatif jenjang SMA

Bench	Subtes Kuantitatif		
	Deretan Angka	Aritmatika & Aljabar	Geometri
1 240	Tidak terjaring soal	Individu mampu menyelesaikan permasalahan aritmatika yang berhubungan dengan rugi/laba dan kecepatan.	Tidak terjaring butir soal

2 280	Individu mampu menentukan pola deretan angka dalam bentuk operasi penjumlahan dan pengurangan.	Individu mampu menyelesaikan permasalahan aritmatika yang berhubungan dengan himpunan.	Individu mampu menyelesaikan permasalahan geometri bentuk 2D dan 3D baik gambar atau cerita.
3 320	Individu mampu menyelesaikan deretan angka loncat 1 atau 2 angka dengan pola operasi penjumlahan, pengurangan dan perkalian.	a. Individu mampu menyelesaikan permasalahan aritmatika pada bilangan bulat dengan beberapa operasi bilangan. b. Individu mampu menyelesaikan permasalahan aritmatika yang berhubungan dengan perbandingan.	Individu mampu menyelesaikan permasalahan geometri yang merupakan gabungan bentuk 2D atau 3D baik gambar atau cerita.
4 360	Individu mampu menyelesaikan deretan angka loncat 1 atau 2 angka dengan pola operasi penjumlahan, pengurangan, perkalian dan pembagian.	a. Individu mampu menyelesaikan permasalahan aritmatika pada bilangan bulat atau berpangkat dengan beberapa operasi bilangan. b. Individu mampu menyelesaikan permasalahan aritmatika dengan membuat persamaan matematika yang terdiri dari 2 variabel.	Individu mampu menentukan luas permukaan suatu bangun geometri Individu mampu menentukan besarnya sudut bangun geometri.

**Tabel 3.** Deskripsi potensi subtes penalaran jenjang SMA

Bench	Subtes Penalaran		
	Logis	Diagram	Analitis
1 280	Individu mampu menarik kesimpulan dari dua premis yang bersifat faktual.	Individu mampu menentukan hubungan antar objek yang bersifat konkrit	Individu mampu memahami dan menggunakan informasi untuk menyelesaikan permasalahan.
2 310	Individu mampu menarik kesimpulan dari dua premis yang bersifat semi faktual.	Individu mampu menentukan hubungan antar objek yang bersifat makhluk hidup.	Individu mampu menganalisis dan menentukan struktur hubungan dari informasi untuk menyelesaikan permasalahan.
3 340	Individu mampu menarik kesimpulan dari dua premis yang bersifat semi faktual yang tidak berlaku secara umum.	Individu mampu menentukan hubungan bagian antar objek yang bersifat makhluk hidup atau konsepsi abstrak.	Individu mampu memahami, menganalisis dan mengolah informasi untuk menyelesaikan permasalahan.
4 370	Individu mampu menarik kesimpulan dari dua premis yang bersifat fiktif dan mengandung hipotesis.	Individu mampu menentukan hubungan antar objek yang bersifat makhluk hidup, konsep, profesi atau status secara simultan.	Individu mampu memahami, menganalisis dan menyelesaikan permasalahan berdasarkan struktur hubungan dari informasi yang kompleks dan memerlukan pengolahan terlebih dahulu.

**SIMPULAN DAN SARAN**

Dari hasil analisis prediksi dan gambaran potensi pada setiap subtes dapat diambil kesimpulan sebagai berikut:

Prediksi potensi bakat skolastik pada prestasi seseorang ketika menempuh studi pada jenjang

yang lebih tinggi masih cenderung rendah. Alasan yang dapat dikemukakan adalah masih ada faktor lain yang berpengaruh tapi tidak dijadikan prediktor. Dalam penelitian ini, faktor yang mungkin dijadikan prediktor juga adalah prestasi pada jenjang pendidikan sebelumnya dan kemampuan lain yang mendukung keberhasilan



seseorang. Besarnya kontribusi pada setiap subtes dapat menjadi pertimbangan dalam penjurusan siswa yang akan masuk SMA.

Deskripsi potensi pada setiap bench menunjukkan keunikan masing-masing sehingga dapat membedakan potensi pada bench rendah dan bench tinggi. Pada bench tinggi diperlukan daya nalar yang tinggi dalam menganalisis dan mengolah informasi sehingga dapat memecahkan masalah dengan solusi yang tepat.

Atas dasar simpulan di atas maka dapat dirumuskan rekomendasi sebagai berikut:

Deskripsi potensi yang telah disusun dalam penelitian ini dapat dimanfaatkan oleh Puspendik dalam melaporkan hasil tes sehingga menjadi lebih bermakna dan instansi yang menyelenggarakan seleksi dapat menggunakan deskripsi tersebut sebagai bahan pertimbangan dalam menentukan pembelajaran selanjutnya atau penempatan pada bidang-bidang sesuai dengan potensinya.

Temuan penelitian pada prediksi jurusan IPA atau IPS dapat digunakan oleh penyelenggara pendidikan di sekolah menengah atas dalam menentukan peminatan siswa pada mata pelajaran yang sesuai dengan potensi yang dimilikinya..

\*\*\*\*\*

## REFERENSI

- Alnahdi, G. H. (2015). Aptitude tests and successful college students: the predictive validity of the general aptitude test (GAT) in Saudi Arabia. *International Education Studies*, 8(4), 1-7.
- Asrijanty. (2014). Validitas prediktif bakat skolastik dan prestasi belajar sebagai kriteria seleksi masuk perguruan tinggi. *Jurnal Pendidikan dan Kebudayaan*, 20(4), 515-534.
- Anastasi, A. (1988). *Psychological testing 6<sup>th</sup> ed.* New York : Macmillan Publishing Co.
- Beaton, A.E. & Allen, N.L. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics*. Summer , 17( 2). 191-204
- Bejar, I.I. (2008). Standard setting: What is it? Why is it important?. *R&D Connections*, 7.
- Berk, L. (2000). *Child development 5<sup>th</sup> ed.* Massachusetts: Allyn and Bacon
- Camara, W.(2013). Defining and measuring college and career readiness: A validation framework. *Educational Measurement: Issues and Practice* ,00(00), 1-12.
- Cizek, G.J, Bunch, M.B , & Koons, H. (2004). Setting Performance Standards: Contemporary methods. *Educational Measurement , Issues and Practice* , 23 (4), 31-50.
- Cohen, R.J. & Swerdlik, M.E. (2002). *Psychological testing and assessment: an introduction to test and measurement 5<sup>th</sup> ed.* Boston: McGraw-Hill.
- Cohen, R.J. & Swerdlik, M.E. (2009). *Psychological testing and assessment: an introduction to test and measurement*
- College Board (2012). *The SAT report on college and career readiness.*
- Cronbach, L.J. (1984). *Essentials of psychological testing 4<sup>th</sup> ed* New York : Harper & Row.
- Ferrara, S., et al (2011). Test Development with performance standards and achievement growth in mind. *Educational Measurement : Issues and practice*, 30(4), 3 -14.
- Forsyth, R. A. (1991). Do NAEP scales yield valid criterion-referenced interpretations? *Educational Measurement: Issues and Practice*, 10(3), 3-9, 16.
- Frey, M.C & Detterman, D.K. Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability . Department of Psychology, Case Western Reserve
- Gregory, R.J. (2007). *Psychological testing (history, principles and applications) 5th ed., USA: Pearson education.*
- Gomez, P. G., et.al. (2007). Proficiency descriptors based on a scale-anchoring study of the new TOEFL iBT reading test. *Language Testing*, 24 (3): 417-444.
- Hayes, J.R. (1989). *The complete problem solver (2nd.Ed).* Hillsdale, NJ: Erlbaum.
- Horn, C., et.al. (1993). Paths to success in the college classroom. *Contemporary Educational Psychology*. 18:464 - 478.
- Jaeger, R.M. (1989). Certification of student competence. Dalam R.L. Linn (Eds), *Educational Measurement (3rd ed., pp. 485-514).* New York : American Council on Education/Macmillan.
- Kelly, D. L. (2002). Application of the scale anchoring method to interpret the TIMSS achievement scales. Dalam Robitaille, D.F., & Beaton, A.E (Eds). *Secondary analysis of the TIMSS data.* New York : Kluwer Academic Publishers.

- Lohman, D. F. 2004. Aptitude for College: The Importance of Reasoning Tests for Minority Admissions. In R. Zwick (Ed.). *Rethinking the SAT: The Future of Standardized Testing in University Admissions* (pp. 41-55). New York: RoutledgeFalmer
- Lyre'n, P.-E. (2008) Prediction of academic performance by means of the Swedish scholastic assessment test. *Scandinavian Journal of Educational Research* Vol. 52, No. 6, December 2008, pp. 565–581
- Martin, M. O., et.al. (1997). *Science achievement in the primary school years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- Mullis, I. V. S., et.al. (1998). *Mathematics and science achievement in the final year of secondary school: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- Olatoye, R.A., & Aderogba, A. A. (2011). Performance of senior secondary school science students in aptitude test: The role of student verbal and numerical abilities. *Journal of Emerging Trends in Educational Research and Policy Studies (JETERAPS)* 2 (6): 431-435, ISSN : 2141-6990
- Resnick, L.B, Nolan, K.J., & Resnick, D.P. (1995). Benchmarking education standards. *Educational Evaluation and Policy Analysis* : 17(438).
- Wyatt, J. et.al. (2011). SAT benchmarks : Development of the college readiness and its relationship to secondary and postsecondary school performance. College Board : Research Report , 5.
- Zwick, R., & Himelfarb, I. (2011) . The effect of high school socioeconomic status on the predictive validity of SAT scores and high school grade-point average. , Summer , 48(2), 101–121