

## KUALITAS SOAL UJIAN SEKOLAH MATEMATIKA PROGRAM IPA DAN KONTRIBUSINYA TERHADAP HASIL UJIAN NASIONAL

*C. Heri Sulistianan*

SMA Stella Duce 1 Yogyakarta

heri.cornel@yahoo.co.id

### Abstrak

Penelitian ini bertujuan untuk mendeskripsikan kualitas soal Ujian Sekolah Matematika program IPA dan kontribusi skor Ujian Sekolah terhadap nilai Ujian Nasional. Jenis penelitian ini deskriptif eksploratif dengan pendekatan survei. Populasi penelitian ini adalah enam SMA Swasta Yayasan Katolik di Kota Yogyakarta. Sampel adalah lima SMA yang representatif dalam menggambarkan ciri populasi. Naskah soal ditelaah tiga orang ahli untuk analisis kualitas soal secara kualitatif. Jawaban siswa dianalisis secara Teori Tes Klasik dan Teori Respons Butir untuk mengetahui kualitas soal secara kuantitatif. Skor Ujian Sekolah dikorelasikan terhadap nilai Ujian Nasional untuk mengukur kontribusinya. Hasil penelitian: (1) kualitas soal secara kualitatif adalah sangat baik (satu SMA), baik (dua SMA), cukup baik (satu SMA), dan kurang baik (satu SMA); (2) secara kuantitatif menurut Teori Tes Klasik adalah cukup baik (satu SMA), kurang baik (tiga SMA), dan tidak baik (satu SMA), dengan reliabilitas  $\alpha$  termasuk reliabel; (3) secara kuantitatif menurut Teori Respons Butir adalah baik (tiga SMA), cukup baik (satu SMA), dan kurang baik (satu SMA); (4) Indeks konsistensi analisis butir soal termasuk cukup konsisten; (5) korelasi skor Ujian Sekolah terhadap nilai Ujian Nasional termasuk kategori besar; (6) berdasarkan telaah validitas isi perangkat, lebih dari 90% soal valid dan skor Ujian Sekolah valid/akurat dalam memprediksi hasil Ujian Nasional.

**Kata kunci:** *kualitas soal, ujian sekolah, matematika, ujian nasional*

## THE QUALITY OF MATHEMATICS SCHOOL EXAMINATION OF SCIENCE MAJOR AND ITS CONTRIBUTION TO THE RESULT OF NATIONAL EXAMINATION

*C. Heri Sulistianan*

SMA Stella Duce 1 Yogyakarta

heri.cornel@yahoo.co.id

### Abstract

The purpose of this research is to describe the quality of Mathematics School Examination of Science Major and the contribution of School Examination scores to the result of National Examination. This research is a descriptive explorative research with survey approach. The population in this research are six Senior High Schools in Yogyakarta. The samples are five representative senior high schools to describe the characteristics of the population. The test was analyzed by three experts to find out the test quality qualitatively. The students' answer was analyzed with Classic Test Theory and Item Response Theory to find out the test quality quantitatively. The School Examination score was correlated with the National Examination score to measure their contribution. The research result: (1) the school exam quality is very good (one school), good (two schools), quite good (one school), and not so good (one school); (2) Quantitatively based on Classic Test Theory are quite good (one school), not so good (three schools), dan poor (one school), by Alpha reliability is reliable; (3) Quantitatively based on Item Response Theory are good (three schools), quite good (one school), and not so good (one school); (4) The test item index consistency are quite consistent; (5) The correlation of the School Examination scores with the result of National Examination is in large categories; (6) Based on content validity study, more than 90% items are valid and scores of School Examination is valid/accurate in predicting the result of the National Examination.

**Keywords:** *the test quality, school examination, mathematics, national examination*

## **Pendahuluan**

Pendidikan mempunyai peranan sangat strategis dalam pembangunan nasional untuk mencapai bangsa yang maju, mandiri dan beradab. Keberhasilan dalam pendidikan merupakan tanggung jawab bersama antara keluarga, masyarakat, dan pemerintah. Menyadari akan pentingnya pendidikan bagi seluruh anak bangsa, pemerintah terus berupaya memenuhi hak setiap warga negara dalam memperoleh layanan pendidikan untuk meningkatkan kualitas hidup bangsa Indonesia. Hal ini penting seperti telah ditegaskan dalam UUD 1945 bahwa setiap warga negara berhak mendapatkan pendidikan dan Pemerintah mempunyai kewajiban dan tanggung jawab dalam penyelenggaraan pendidikan untuk mencapai tujuan negara, yaitu mencerdaskan kehidupan bangsa dan memajukan kesejahteraan umum. Pendidikan bahkan merupakan syarat mutlak yang harus dipenuhi dalam memasuki era persaingan global yang sarat dengan persaingan antarbangsa yang sangat ketat.

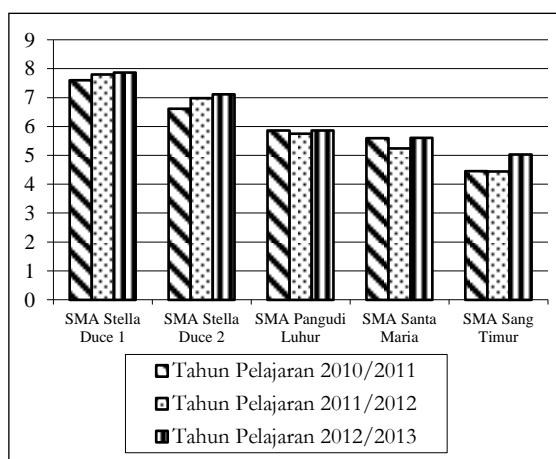
Untuk melakukan kontrol dan pengawasan terhadap penyelenggaraan pendidikan di Indonesia, pemerintah telah mengeluarkan delapan Peraturan Menteri. Salah satunya yaitu Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 tanggal 11 Juni 2007 tentang Standar Penilaian Pendidikan. Dalam peraturan tersebut disebutkan bahwa standar penilaian pendidikan adalah standar nasional pendidikan yang berkaitan dengan mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik, dan penilaian pendidikan adalah proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar peserta didik. Penilaian pendidikan melalui tiga cara yakni: penilaian oleh pendidik, penilaian oleh satuan pendidikan (sekolah), dan penilaian oleh pemerintah. Penilaian hasil belajar oleh pemerintah dilakukan dalam bentuk pelaksanaan Ujian Nasional (UN) yang bertujuan untuk menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu dalam kelompok mata pelajaran ilmu pengetahuan dan teknologi. Dalam Undang-Undang Republik

Indonesia Nomor 20 Tahun 2003 pasal 57 dinyatakan bahwa dalam rangka pengendalian mutu pendidikan secara nasional dilakukan evaluasi sebagai bentuk akuntabilitas penyelenggaraan pendidikan kepada pihak-pihak yang berkepentingan. UN digunakan sebagai standarisasi dari pemerintah untuk menguji kelayakan seorang siswa untuk dapat melanjutkan pendidikan ke jenjang yang lebih tinggi dan sebagai pemerataan pendidikan secara nasional.

Kriteria kelulusan dari satuan pendidikan yang terkait dengan hasil dari ujian yang diselenggarakan pemerintah mengalami perkembangan. Perubahan dimulai dari periode sebelum tahun 2003 yang diberi nama Evaluasi Tahap Akhir Nasional (EBTANAS) dengan kelulusan dihitung menggunakan formula PQR. Tahun 2003-2004 ujian berganti nama menjadi Ujian Akhir Nasional (UAN) dengan kelulusan berdasarkan nilai yang diperoleh setiap mata pelajaran yang diujikan yakni nilai minimal 3,00 (UN 2003) dan 4,00 (UN 2004). Mulai tahun 2005 digunakan istilah Ujian Nasional dengan kriteria kelulusan berdasarkan nilai minimal yang dicapai tiap mata pelajaran yang diujikan mengalami peningkatan dan juga diperhitungkan nilai rata-rata minimal dari semua mata pelajaran yang diujikan. Terkait dengan kontroversi penentuan kelulusan yang seolah-olah tidak memperhitungkan proses yang dijalani peserta didik selama belajar di sekolah, mulai tahun 2011 pemerintah menetapkan perubahan kriteria kelulusan Ujian Nasional untuk satuan pendidikan SMA yang memperhitungkan nilai proses yakni nilai rapor semester 3 (tiga) sampai dengan 5 (lima) dan juga adanya Ujian Sekolah untuk mata pelajaran yang diujikan dalam Ujian Nasional. Perhitungannya dimulai dari Nilai Sekolah (NS) yang dihitung dari 60% dari hasil Ujian Sekolah dan 40% dari rata-rata nilai rapor semester 3 sampai 5. Kemudian dihitung Nilai Akhir (NA) yaitu 60% dari hasil Ujian Nasional dan 40% dari NS. Untuk tahun 2013 kriteria kelulusan Ujian Nasional adalah nilai NA minimal untuk tiap mata pelajaran yang diujikan adalah 4,00 dan nilai rata-rata dari

semua mata pelajaran yang diujikan minimal 5,50 (Peraturan Kepala Dinas Dikpora DIY Nomor 0111, 2013, p.16).

Hasil Ujian Nasional mata pelajaran Matematika untuk program IPA yang dicapai lima SMA Swasta Yayasan Katolik Kota Yogyakarta tahun pelajaran 2010/2011 sampai tahun pelajaran 2012/2013 disajikan pada Gambar 1.



Gambar 1. Rerata Nilai Ujian Nasional Mata Pelajaran Matematika

Berdasarkan Gambar 1 tampak bahwa hasil Ujian Nasional mata pelajaran Matematika program IPA yang dicapai lima SMA Swasta Yayasan Katolik Kota Yogyakarta. SMA Stella Duce 1 selama kurun waktu tiga tahun terakhir selalu meningkat dan berada di kisaran 7,50 - 8,00. SMA Stella Duce 2 selalu meningkat dari 6,62 menjadi 6,98 dan kemudian 7,12. SMA Santa Maria berada di kisaran 5,50-6,00. SMA Pangudi Luhur mengalami fluktuasi dari 5,86 turun menjadi 5,24 kemudian naik menjadi 5,61. SMA Sang Timur mengalami fluktuasi yang cukup kecil di kisaran 4,45 – 5,03. Nilai Ujian Nasional dari tiga SMA Swasta Yayasan Katolik Kota Yogyakarta belum bisa dikatakan baik karena nilai rata-rata Ujian Nasional kurang dari 6,00.

Terkait dengan otonomi sekolah dalam penyusunan soal dan penyelenggaraan Ujian Sekolah, di SMA Swasta Yayasan Katolik Kota Yogyakarta naskah soal disusun oleh guru-guru yang ada di sekolah masing-masing. Kondisi guru di setiap sekolah sangat beragam. Termasuk juga dalam ke-

mampuan menyusun naskah soal setara Ujian Sekolah. Harus diakui bahwa tidak semua guru pernah mengikuti pelatihan dalam membuat soal yang baik dan tidak semua guru mampu menyusun naskah soal ujian yang sesuai dengan kaidah-kaidah penulisan soal yang baku. Apalagi jika hal itu dikaitkan dengan kesahihan dan keterandalan butir soal. Dalam Peraturan Kepala Dinas Pendidikan, Pemuda dan Olahraga Daerah Istimewa Yogyakarta Nomor 0111 Tahun 2013 tentang Petunjuk Teknis Penyelenggaraan Ujian Nasional SMP/MTs, SMPLB, SMA/MA, SMALB DAN SMK Daerah Istimewa Yogyakarta Tahun Pelajaran 2012/2013 tertanggal 4 Februari 2012 disebutkan bahwa tanggal 20 Maret 2013 Ujian Sekolah tingkat SMA/MA sudah harus selesai. Dengan demikian dalam rentang waktu mendesak guru harus membuat naskah soal Ujian Sekolah. Perlu dilakukan penelitian mengenai kualitas soal Ujian Sekolah yang dibuat oleh guru mata pelajaran di setiap sekolah tersebut.

Guru harus profesional dalam melaksanakan tugas dan kewajibannya. Ini sesuai dengan apa yang tercantum dalam Undang-Undang Guru dan Dosen (Undang-Undang RI Nomor 14, 2005) pasal 1 ayat 1 yang menyatakan bahwa:

*Guru adalah pendidik profesional dengan tugas utama mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi peserta didik pada pendidikan anak usia dini, jalur pendidikan formal, pendidikan dasar dan pendidikan menengah.*

Program pendidikan guru membantu guru untuk mendapatkan pengetahuan dan keterampilan, mengembangkan sikap profesional, dan dampak keyakinan mereka tentang pembelajaran yang baik yang dapat mempengaruhi hasil belajar (Bahous & Nabhani, 2011, p. 22). Perwujudan unjuk kerja profesional guru ditunjang dengan jiwa profesionalisme, yaitu sikap mental yang senantiasa mendorong untuk mewujudkan diri sebagai guru profesional. Kualitas profesionalisme ditunjukkan oleh lima unjuk kerja, antara lain yaitu keinginan untuk senantiasa mengejar kesempatan pengembangan profesional yang dapat meningkatkan

dan memperbaiki kualitas pengetahuan dan keterampilan, serta melaksanakan pekerjaan dengan mengedepankan kualitas dan cita-cita dalam profesi (Aqib & Rohmanto, 2008, pp. 145-146). Dalam melaksanakan evaluasi dan penilaian, guru dituntut untuk dapat membuat soal tes yang berkualitas baik dan hasil tes tersebut benar-benar dapat memberikan informasi yang tepat tentang siswa mana yang telah/belum menguasai materi yang telah diajarkan.

Nilai Sekolah yang diperoleh dari hasil Ujian Sekolah dan nilai rapor, kemudian digabungkan dengan hasil Ujian Nasional menjadi Nilai Akhir yang merupakan syarat kelulusan dari Ujian Nasional. Penyusunan soal Ujian Sekolah seperti halnya Ujian Nasional mengacu pada Standar Kompetensi dan Kompetensi Dasar yang tertuang pada Standar Isi dalam Permendiknas Nomor 22 Tahun 2006. Ujian Sekolah dilaksanakan sekitar sebulan sebelum Ujian Nasional, sehingga mestinya hasil Ujian Sekolah memiliki kontribusi yang positif terhadap hasil Ujian Nasional.

Tujuan dari penelitian ini adalah mendeskripsikan kualitas soal Ujian Sekolah mata pelajaran Matematika program IPA di SMA Swasta Yayasan Katolik Kota Yogyakarta tahun pelajaran 2012/2013 secara kualitatif, kuantitatif menurut Teori Tes Klasik, kuantitatif menurut Teori Respons Butir, mengukur seberapa besar kontribusi hasil Ujian Sekolah terhadap hasil Ujian Nasional, dan mendeskripsikan validitas dan reliabilitas soal Ujian Sekolah.

Manfaat dari penelitian ini adalah sebagai masukan bagi tim pembuat soal Ujian Sekolah mata pelajaran matematika dalam rangka penyempurnaan soal berdasarkan temuan hasil penelitian dan sebagai langkah awal pengembangan bank soal matematika dari sekumpulan butir-butir soal yang baik berdasarkan hasil penelitian, sebagai bahan masukan sekolah untuk menentukan pendampingan atau pelatihan lebih lanjut kepada guru terkait perencanaan, penyusunan, dan pengelolaan soal Ujian Sekolah yang berkualitas baik, dan sebagai bahan masukan bagi Dinas Pendidikan Kota Yogyakarta

bagaimana kualitas soal Ujian Sekolah mata pelajaran matematika yang dibuat oleh guru mata pelajaran di sekolah masing-masing.

## **Metode Penelitian**

Penelitian ini merupakan penelitian deskriptif eksploratif kuantitatif. Dengan bantuan ahli, peneliti menganalisis soal Ujian Sekolah secara kualitatif untuk mendeskripsikan kualitas soal berdasarkan aspek materi, konstruksi, dan bahasa. Respons/jawaban dari peserta didik dianalisis secara kuantitatif untuk mengetahui kualitas butir soal secara empiris berdasarkan teori tes klasik, meliputi taraf kesukaran, daya beda, dan efektivitas pengecoh, kemudian juga dianalisis secara empiris berdasarkan Teori Respons Butir 2-parameter untuk mengetahui taraf kesukaran dan daya beda, kecocokan butir dengan model logistik, fungsi informasi, dan kesalahan baku pengukuran (*SEM*). Validitas soal dibuktikan dengan menelaah kesesuaian butir-butir soal dengan kisi-kisi. Reliabilitas soal diperoleh dari hasil *output* analisis menggunakan program *Itean* pada statistik perangkat tes bagian Alpha.

Pendekatan yang digunakan dalam penelitian ini adalah pendekatan survei dengan tidak memberi perlakuan terhadap variabel yang akan diteliti. Data pendukung terhadap kualitas soal di setiap sekolah dikumpulkan untuk dianalisis guna mendeskripsikan kualitas soal Ujian Sekolah.

Penelitian dilakukan dengan mengambil dokumen naskah soal Ujian Sekolah beserta respons/jawaban peserta didik di setiap SMA Swasta Katolik Kota Yogyakarta. Waktu penelitian Maret 2014 sampai dengan Agustus 2014.

Populasi dalam penelitian ini adalah enam SMA Swasta Yayasan Katolik di Kota Yogyakarta (269 peserta didik). Sampel diambil secara *purposive sampling* yakni lima sekolah yaitu SMA Stella Duce 1, SMA Stella Duce 2, SMA Pangudi Luhur, SMA Santa Maria, dan SMA Sang Timur dengan total 264 peserta didik, yang sekiranya sudah cukup representatif dalam menggambarkan ciri populasi dalam penelitian ini. Satu seko-

lah tidak diikutsertakan sebagai objek penelitian karena soal Ujian Sekolah hanya diujikan kepada lima peserta didik, sehingga data yang ada kurang representatif untuk melihat kualitas soal secara empirik.

Analisis butir soal secara kualitatif menggunakan teknik panel yakni teknik menelaah butir soal berdasarkan kaidah penulisan butir soal, yakni dari aspek materi, konstruksi, bahasa, kebenaran kunci atau pedoman penskoran (Kusaeri & Suprananto, 2012, p. 165). Penelaah terdiri atas tiga orang guru/dosen matematika yang ahli di bidangnya (*professional judgement*) yang memiliki pengalaman mengajar lebih dari 10 tahun dengan kualifikasi akademik minimal S2. Caranya, ketiga penelaah diberikan kisi-kisi soal, butir-butir soal yang akan ditelaah, format penelaahan, dan pedoman penelaahan/penilaian. Ketiga penelaah disamakan persepsinya dahulu sebelum melakukan kegiatan penelaahan, kemudian mereka bekerja sendiri-sendiri di tempat yang berbeda. Kemudian ditetapkan kriteria butir soal yang baik apabila memenuhi semua kriteria, kurang baik apabila ada satu kriteria yang tidak terpenuhi, dan tidak baik apabila lebih dari satu kriteria yang tidak terpenuhi. Untuk masing-masing nomor kriteria dikatakan tidak terpenuhi apabila lebih dari satu penelaah menyatakan kriteria penelaahan tidak terpenuhi.

Dari hasil analisis butir soal secara kualitatif tersebut, kualitas soal Ujian Sekolah untuk setiap sekolah ditentukan berdasarkan persentase jumlah butir soal yang kurang baik atau tidak baik. Semakin kecil persentase butir soal yang kurang baik atau tidak baik mengindikasikan bahwa soal Ujian Sekolah tersebut semakin berkualitas.

Analisis butir soal pilihan ganda menurut Teori Tes Klasik dihitung dengan bantuan program *IteMan*. Kualitas empirik butir soal pilihan ganda dilihat dari taraf kesukaran, daya beda, dan efektivitas pengecoh.

Berdasarkan hasil analisis butir soal tersebut, kualitas butir soal secara empiris dikategorikan baik atau direvisi atau tidak baik. Butir soal baik, jika ketiga syarat be-

rikt terpenuhi yakni: tingkat kesukarannya sedang yakni berada pada interval  $0,33 \leq TK \leq 0,66$ , daya beda:  $DB \geq 0,30$ , dan pilihan jawaban dipilih oleh minimal 2% dari peserta tes. Butir soal yang mudah atau sukar dapat dikategorikan baik apabila memiliki daya beda  $\geq 0,30$  dan pilihan jawaban dipilih oleh minimal 2% dari peserta tes (Kartowagiran, 2005, p. 49). Butir soal direvisi, apabila memiliki daya beda:  $DB \geq 0,30$  tetapi ada pilihan jawaban yang dipilih oleh kurang dari 2% peserta tes. Butir soal tidak baik, apabila memiliki daya beda:  $DB < 0,30$ .

Analisis butir soal pilihan ganda menurut Teori Respons Butir dihitung dengan bantuan program *Bilog MG*. Analisis ini dilakukan dengan menggunakan model logistik untuk 2 parameter butir, yaitu tingkat kesukaran ( $b_i$ ) dan daya beda ( $a_i$ ).

Berdasarkan hasil analisis butir soal tersebut, kualitas butir soal secara empiris dikategorikan baik atau tidak baik. Butir soal baik, jika tingkat kesukaran ( $b_i$ ) berada pada interval  $-2 \leq b_i \leq 2$  dan daya beda ( $a_i$ ) berada pada interval  $0 \leq a_i \leq 2$ . Butir soal tidak baik, apabila salah satu kriteria di atas tidak dipenuhi, yakni: tingkat kesukaran:  $b_i < -2$  atau  $b_i > 2$ , daya beda:  $a_i < 0$  atau  $a_i > 2$ .

Reliabilitas atau konsistensi hasil tes dilihat dari hasil analisis program *IteMan*. *Output* hasil analisis program *IteMan* didapat reliabilitas *Alpha*.

Analisis konsistensi butir soal menggunakan tabel koefisien  $\phi$  ( $\Phi$ ), yang dimaksudkan untuk melihat sejauh mana analisis dengan dua cara berbeda itu konsisten, yakni butir-butir soal yang tidak baik menurut Teori Tes Klasik juga tidak baik menurut Teori Respons Butir. Hasil analisis konsistensi diinterpretasikan berdasarkan nilai korelasinya.

Besarnya kontribusi skor Ujian Sekolah terhadap hasil Ujian Nasional, dihitung dengan rumus korelasi *product moment*.

## Hasil Penelitian dan Pembahasan

Naskah soal Ujian Sekolah mata pelajaran Matematika yang dianalisis dalam

penelitian ini dibuat oleh guru Matematika yang mengajar di kelas XII program IPA di sekolah masing-masing berdasarkan kisi-kisi yang dibuat oleh guru tersebut. Naskah soal yang telah dibuat umumnya tanpa melalui penelaahan soal langsung dipakai untuk Ujian Sekolah. Guru hanya memastikan bahwa setiap butir soal pilihan ganda yang telah dibuat memiliki kunci jawaban pada salah satu pilihan jawaban yang tersedia.

#### Kualitas Soal secara Kualitatif

Penelaahan dilakukan oleh seorang dosen bidang Matematika dan dua orang guru mata pelajaran Matematika tingkat SMA yang telah memiliki pengalaman mengajar lebih dari 10 tahun.

Hasil penelaahan butir soal dikelompokkan menjadi tiga kategori yakni baik, kurang baik, dan tidak baik. Butir soal yang baik berdasarkan hasil telaah butir soal adalah butir yang telah memenuhi semua kriteria telaah. Butir soal yang kurang baik adalah butir soal yang tidak memenuhi salah satu kriteria telaah. Butir soal yang tidak baik adalah butir soal yang lebih dari satu kriteria telaah tidak terpenuhi ditinjau dari aspek materi, konstruksi, dan bahasa. Butir soal dikatakan memenuhi kriteria yang ditentukan apabila minimal memperoleh dua dukungan dari tiga penelaah yang telah ditetapkan.

Persentase butir soal yang kurang baik/tidak baik berdasarkan hasil penelaahan para ahli adalah 6,7% (SMA Stella Duce 1), 17,1% (SMA Stella Duce 2), 12,5% (SMA Pangudi Luhur), 37,5% (SMA Santa Maria), dan 21,4% (SMA Sang Timur). Kategori kualitas soal secara kualitatif disajikan pada Tabel 1.

Hasil telaah butir soal Ujian Sekolah dari lima SMA Swasta Yayasan Katolik di Kota Yogyakarta pada tahun pelajaran 2012/2013 mata pelajaran Matematika program IPA menunjukkan ada satu sekolah yang soal Ujian Sekolahnya termasuk dalam kategori sangat baik, dua sekolah soal Ujian Sekolahnya termasuk dalam kategori baik, satu sekolah soal Ujian Sekolahnya termasuk dalam kategori cukup baik, dan satu

sekolah lainnya soal Ujian Sekolahnya termasuk dalam kategori kurang baik.

Tabel 1. Kualitas Soal secara Kualitatif

Nama Sekolah	Kualitas Soal
SMA Stella Duce 1	Sangat baik
SMA Stella Duce 2	Baik
SMA Pangudi Luhur	Baik
SMA Santa Maria	Kurang baik
SMA Sang Timur	Cukup baik

Penyebab butir soal kurang baik/tidak baik karena tidak memenuhi salah satu kriteria dari aspek materi atau konstruksi atau bahasa. Kriteria yang tidak terpenuhi paling banyak pada aspek konstruksi yakni pilihan jawaban yang tidak diurutkan berdasarkan besar/kecilnya angka atau urutan kronologisnya, sebanyak 20 butir soal yang berasal dari kelima sekolah yang diteliti. Kemudian diikuti oleh kriteria pada aspek materi yakni hanya ada satu kunci jawaban, sebanyak 3 butir soal yang berasal dari tiga sekolah yang diteliti. Ini menunjukkan bahwa belum semua guru memahami pedoman dalam membuat soal pilihan ganda yang benar dan guru tidak melakukan penelaahan soal sebelum soal tersebut diujikan.

#### Kualitas Soal secara Kuantitatif menurut Teori Tes Klasik

Hasil analisis butir soal menggunakan bantuan program *ITEMAN* versi 3.00 ditinjau dari tingkat kesukarannya, soal Ujian Sekolah mata pelajaran Matematika program IPA tahun pelajaran 2012/2013 di SMA Stella Duce 1 Yogyakarta terdiri atas 36,7% soal mudah dan 63,3% soal sedang, di SMA Stella Duce 2 Yogyakarta terdiri atas 34,3% soal mudah, 31,4% soal sedang, dan 34,3% soal sukar, di SMA Pangudi Luhur Yogyakarta terdiri atas 7,5% soal mudah, 55% soal sedang, dan 37,5% soal sukar, di SMA Santa Maria Yogyakarta terdiri atas 6,3% soal mudah, 65,6% soal sedang, dan 21,8% soal sukar, serta di SMA Sang Timur Yogyakarta terdiri atas 3,6% soal mudah, 71,4% soal sedang, dan 25,0% soal sukar. Secara umum persentase soal dengan

kategori sedang di empat sekolah sangat dominan, kecuali di satu sekolah yang persentase soal hampir merata untuk ketiga kategori tingkat kesukaran. Butir soal yang mudah berasal dari materi yang secara umum mudah dipahami siswa dan pengecoh yang kurang berfungsi. Butir soal yang sulit berasal dari materi yang tingkat kompleksitasnya tinggi dan perlu ketelitian dalam perhitungan. Rerata tingkat kesukaran masing-masing adalah 0,604, 0,513, 0,404, 0,413, dan 0,440 yang semuanya berada pada kategori sedang.

Ditinjau dari daya bedanya persentase butir soal yang diterima dan yang tidak diterima untuk masing-masing sekolah adalah 76,7% dan 23,3%, 62,9% dan 37,1%, 60,0% dan 40,0%, 53,1% dan 46,9%, 67,9% dan 32,1%. Secara umum untuk tiap-tiap sekolah lebih dari 50% butir soal Ujian Sekolahnya mampu membedakan antara siswa yang pandai dan siswa yang kurang pandai. Rerata daya beda untuk masing-masing sekolah adalah 0,340, 0,309, 0,34, 0,274, dan 0,346 yang berarti bahwa hanya satu sekolah yang rerata daya bedanya tidak diterima (kurang dari 0,3). Kemungkinan penyebab butir soal memiliki daya beda yang tidak diterima adalah siswa kurang memahami materi pada soal tersebut dan secara kebetulan faktor tebakan jawaban yang benar cukup tinggi.

Selanjutnya, dilihat dari keberfungsian pengecoh, hampir semua sekolah soal Ujian Sekolahnya memiliki cukup banyak pengecoh yang tidak berfungsi dengan baik dan hanya satu sekolah saja yang persentase butir soal dengan pengecoh tidak berfungsi cukup kecil. Persentase soal yang pengecohnya tidak berfungsi untuk masing-masing sekolah adalah 16,7%, 45,7%, 5,0%, 37,5%, dan 32,1%. Penyebab pengecoh tidak berfungsi dengan baik karena mencolok kesalahannya dan ada pengecoh yang dibuat secara asal artinya tidak memperhatikan besarnya peluang pengecoh itu dipilih oleh siswa yang tidak menguasai materi atau tidak teliti dalam teknis perhitungan matematis.

Berdasarkan persentase butir soal yang tidak baik, maka kualitas soal Ujian Sekolah secara kuantitatif menurut Teori

Tes Klasik di lima sekolah disajikan pada Tabel 2 berikut.

Tabel 2. Kualitas Soal secara Kuantitatif menurut Teori Tes Klasik

Nama Sekolah	Kualitas Soal
SMA Stella Duce 1	Cukup baik
SMA Stella Duce 2	Kurang baik
SMA Pangudi Luhur	Kurang baik
SMA Santa Maria	Tidak baik
SMA Sang Timur	Kurang baik

Lima SMA Swasta Yayasan Katolik di Kota Yogyakarta pada tahun pelajaran 2012/2013, hanya satu sekolah yang hasil analisis butir soal Ujian Sekolah mata pelajaran Matematika program IPA menurut Teori Tes Klasik termasuk dalam kategori cukup baik yakni SMA Stella Duce 1, tiga sekolah soal Ujian Sekolahnya termasuk dalam kategori kurang baik yakni SMA SMA Stella Duce 2, SMA Pangudi Luhur, dan SMA Sang Timur, serta satu sekolah lainnya soal Ujian Sekolahnya termasuk dalam kategori tidak baik yakni SMA Santa Maria. Tidak ada sekolah yang kualitas soal Ujian Sekolahnya menurut Teori Tes Klasik termasuk dalam kategori baik ataupun amat baik.

#### Kualitas Soal secara Kuantitatif menurut Teori Respons Butir

Hasil analisis butir soal menggunakan bantuan program *Bilog-MG* ditinjau dari tingkat kesukarannya, soal Ujian Sekolah mata pelajaran Matematika program IPA tahun pelajaran 2012/2013 di SMA Stella Duce 1 Yogyakarta terdiri atas 16,7% soal mudah dan 83,3% soal sedang, di SMA Stella Duce 2 Yogyakarta terdiri atas 11,4% soal mudah, 65,7% soal sedang, 8,6% soal sukar, dan 14,3% soal tidak valid, di SMA Pangudi Luhur Yogyakarta terdiri atas 85% soal sedang dan 15% soal sukar, di SMA Santa Maria Yogyakarta terdiri atas 78,1% soal sedang, 6,3% soal sukar, dan 15,6% soal tidak valid, serta di SMA Sang Timur Yogyakarta terdiri atas 82,1% soal sedang, 3,6% soal sukar, dan 14,3% soal tidak valid.

Secara umum untuk kelima sekolah, butir soal dengan tingkat kesukaran kategori sedang sangat dominan (lebih dari 60%) dan untuk tiga sekolah terdapat butir soal yang tidak valid karena pada bagian *Correlation Biser* di *output Phase 1* bernilai negatif karena butir soal tersebut tidak dapat membedakan siswa yang pandai dan yang kurang pandai. Rerata tingkat kesukaran butir soal untuk masing-masing sekolah adalah -0,73, -0,20, 0,73, 0,45, dan 0,18 yang semuanya berada pada kategori sedang.

Ditinjau dari daya bedanya persentase butir soal yang diterima dan yang tidak diterima untuk masing-masing sekolah adalah 100% dan 0%, 85,7% dan 14,3%, 100% dan 0%, 84,4% dan 15,6%, 85,7% dan 14,3%. Ini berarti untuk tiap-tiap sekolah yang diteliti lebih dari 80% soal daya bedanya diterima. Rerata daya beda untuk masing-masing sekolah yang diteliti adalah 0,56, 0,75, 0,54, 0,60, dan 0,73 yang semuanya berada pada kategori diterima dan artinya secara rata-rata soal tersebut dapat membedakan antara siswa yang pandai dan siswa yang kurang pandai.

Ditinjau dari kecocokan butir soal dengan model logistik dua parameter yang digunakan, soal Ujian Sekolah dari SMA Stella Duce 1 Yogyakarta dan SMA Pangudi Luhur Yogyakarta yang sangat cocok, karena jumlah butir yang cocok pada kedua sekolah tersebut lebih dari 90%. Soal Ujian Sekolah SMA Santa Maria Yogyakarta kurang cocok karena jumlah butir soal yang cocok hanya 53,1%. Sedangkan soal Ujian Sekolah dari SMA Stella Duce 2 Yogyakarta dan SMA Sang Timur Yogyakarta tidak cocok karena jumlah butir yang cocok sangat sedikit yakni sama-sama 28,6%.

Hasil perhitungan fungsi informasi tes dan *SEM* dari lima sekolah yang diteliti, soal yang paling baik berasal dari SMA Sang Timur Yogyakarta karena mampu memberikan fungsi informasi maksimum yakni 8,107 (terbesar dibandingkan sekolah yang lain) dengan *SEM* terkecil 0,351 (terkecil dibandingkan sekolah yang lain) untuk skala kemampuan yang sama. Urutan kedua adalah soal Ujian Sekolah dari SMA Pangudi Luhur

Yogyakarta dengan fungsi informasi maksimum sebesar 0,801 dan *SEM* terkecil 0,353. Urutan ketiga adalah soal Ujian Sekolah dari SMA Stella Duce 2 Yogyakarta dengan fungsi informasi maksimum sebesar 7,825 dan *SEM* terkecil 0,357.

Berdasarkan persentase butir soal yang tidak baik, maka kualitas soal Ujian Sekolah di lima sekolah yang diteliti menurut Teori Respons Butir disajikan pada Tabel 3.

Tabel 3. Kualitas Soal secara Kuantitatif menurut Teori Respons Butir

Nama Sekolah	Kualitas Soal
SMA Stella Duce 1	Baik
SMA Stella Duce 2	Kurang baik
SMA Pangudi Luhur	Baik
SMA Santa Maria	Cukup baik
SMA Sang Timur	Baik

Hasil analisis butir soal Ujian Sekolah dari lima SMA Swasta Yayasan Katolik di Kota Yogyakarta pada tahun pelajaran 2012/2013 mata pelajaran Matematika program IPA menurut Teori Respons Butir menunjukkan ada tiga sekolah yang soal Ujian Sekolahnya termasuk dalam kategori baik, satu sekolah soal Ujian Sekolahnya termasuk dalam kategori cukup baik dan satu sekolah lainnya soal Ujian Sekolahnya termasuk dalam kategori kurang baik.

#### Analisis Konsistensi Butir Soal

Hasil perhitungan dengan koefisien *p<sub>bi</sub>* masing-masing sekolah adalah 0,811, 0,691, 0,515, 0,563, dan 0,478. Hasil tersebut menunjukkan tingkat konsistensi yang sangat tinggi (SMA Stella Duce 1), satu sekolah dengan tingkat konsistensi tinggi (SMA Stella Duce 2), dan tiga sekolah lainnya (SMA Pangudi Luhur, SMA Santa Maria, dan SMA Sang Timur) memiliki tingkat konsistensi sedang. Secara umum dapat disimpulkan bahwa hasil penentuan kualitas soal pada penelitian ini dengan Teori Tes Klasik maupun dengan Teori Respons Butir adalah cukup konsisten.

## Kontribusi Ujian Sekolah terhadap Ujian Nasional

Rata-rata tingkat korelasi antara skor Ujian Sekolah terhadap hasil Ujian Nasional dari lima sekolah sebesar 0,674 menunjukkan bahwa tingkat korelasi termasuk dalam kriteria besar. Tingkat korelasi antara skor Ujian Sekolah terhadap hasil Ujian Nasional dari lima sekolah disajikan dalam Tabel 4.

Tabel 4. Korelasi Ujian Sekolah terhadap Ujian Nasional

Nama Sekolah	Korelasi	Kriteria
SMA Stella Duce 1	0,656	Besar
SMA Stella Duce 2	0,664	Besar
SMA Pangudi Luhur	0,802	Sangat besar
SMA Santa Maria	0,665	Besar
SMA Sang Timur	0,582	Sedang

## Validitas Soal

Validitas soal dalam penelitian ini adalah validitas isi dan validitas kriteria. Validitas isi dibuktikan dengan hasil penelaahan oleh tiga orang ahli di bidang Matematika mengenai kesesuaian butir soal dengan indikator pada kisi-kisi soal dan isi materi sesuai dengan jenjang dan tingkatan kelas. Hasil penelaahan oleh tiga penelaah, dari lima soal Ujian Sekolah yang diteliti terdapat empat SMA yang 100% butir soal Ujian Sekolah valid. Soal Ujian Sekolah dari SMA Santa Maria Yogyakarta terdapat 3 butir soal (9,4%) soal yang tidak valid dan 29 butir soal (90,6%) soal valid.

Validitas kriteria dibuktikan dengan tingkat korelasi antara skor soal Ujian Sekolah dengan nilai Ujian Nasional. Hasil perhitungan untuk lima sekolah yang diteliti menunjukkan koefisien validitas (kriteria) sebesar 0,656, 0,664, 0,802, 0,665, dan 0,582. Koefisien korelasi yang semuanya positif mendekati satu dan berada pada kategori besar (satu sekolah berada pada kategori cukup besar) menunjukkan bahwa skor Ujian Sekolah tersebut akurat dalam memprediksi hasil Ujian Nasional. Ini berarti bahwa derajat validitas soal Ujian Sekolah mata pelajaran Matematika program IPA di

SMA Swasta Yayasan Katolik Kota Yogyakarta pada tahun pelajaran 2012/2013 adalah besar, artinya mampu memprediksi dengan akurat hasil yang diperoleh pada Ujian Nasional.

## Reliabilitas Soal

Tingkat ketepatan dan ketetapan hasil tes pada Ujian Sekolah di lima sekolah yang diteliti ditunjukkan oleh reliabilitas *Alpha* untuk masing-masing sekolah sebesar 0,753, 0,744, 0,811, 0,645, dan 0,747. Artinya soal Ujian Sekolah mata pelajaran Matematika program IPA di SMA Swasta Yayasan Katolik Kota Yogyakarta pada tahun pelajaran 2012/2013 adalah reliabel atau skor Ujian Sekolah konsisten dalam mengukur kemampuan peserta didik yang sebenarnya.

## Simpulan

Berdasarkan hasil analisis data diperoleh kesimpulan: kualitas soal Ujian Sekolah secara kualitatif berdasarkan telaah butir soal dikategorikan sangat baik pada satu sekolah, baik pada dua sekolah, cukup baik pada satu sekolah, dan kurang baik pada satu sekolah. Secara kuantitatif menurut analisis butir soal berdasarkan Teori Tes Klasik, kualitas soal Ujian Sekolah dikategorikan cukup baik pada satu sekolah, kurang baik pada tiga sekolah, dan tidak baik pada satu sekolah. Secara kuantitatif menurut analisis butir soal berdasarkan Teori Respons Butir, kualitas soal Ujian Sekolah dikategorikan baik pada tiga sekolah, cukup baik pada satu sekolah, dan kurang baik pada satu sekolah. Kontribusi skor Ujian Sekolah terhadap nilai Ujian Nasional secara rata-rata termasuk dalam kategori besar, yakni kategori sangat besar pada satu sekolah, kategori besar pada tiga sekolah, dan kategori cukup besar pada satu sekolah. Ditinjau dari validitas isi berdasarkan telaah perangkat soal, 100% butir soal sesuai dengan kisi-kisi pada empat sekolah dan 90,6% butir soal sesuai dengan kisi-kisi pada satu sekolah. Ditinjau dari validitas kriteria, skor Ujian Sekolah terhadap hasil Ujian Nasional, memiliki koefisien validitas (kriteria) yang besar pada empat

sekolah dan cukup besar pada satu sekolah. Hal ini berarti skor Ujian Sekolah tersebut valid/akurat dalam memprediksi hasil Ujian Nasional. Soal Ujian Sekolah adalah reliabel yang ditunjukkan oleh besarnya koefisien reliabilitas *Alpha* lebih dari 0,700 pada empat sekolah dan 0,645 pada satu sekolah.

#### Saran

Saran yang dapat dikemukakan berdasarkan kesimpulan yang diperoleh dari hasil penelitian ini, yaitu perlu adanya pelatihan penulisan soal agar guru memahami pedoman dalam membuat soal yang baik dan sebagai sarana untuk meningkatkan kemampuan dan keterampilan dalam menulis soal, khususnya pada aspek materi dan bahasa. Guru hendaknya melakukan administrasi terhadap butir-butir soal yang telah teruji kualitasnya sebagai langkah awal pembuatan bank soal, sehingga nantinya guru dapat merakit soal Ujian Sekolah yang berasal dari butir-butir soal yang baik. Kerja sama antarguru bidang studi antarsekolah dalam forum-forum ilmiah perlu ditingkatkan, misalnya dalam MGMP, agar guru dapat berbagi pengalaman dan mengetahui perubahan-perubahan terkini terkait dengan situasi pendidikan yang dinamis sehingga guru dapat semakin profesional dalam menjalankan tugas profesinya, salah satunya mengenai pengukuran hasil belajar peserta didik.

#### Daftar Pustaka

Kartowagiran, B. (2005). *Perbandingan berbagai metode untuk mendeteksi bias butir.*

Disertasi doktor, tidak diterbitkan, Universitas Gadjah Mada, Yogyakarta.

Bahous, R. & Nabhani, M. (2011). Assessing education program learning outcomes. *Educational Assessment, Evaluation and Accountability* 23, 21-29.

Depdiknas. (2003). *Undang-Undang RI Nomor 20, Tahun 2003, tentang Sistem Pendidikan Nasional.*

Depdiknas. (2005). *Undang-Undang RI Nomor 14, Tahun 2005, tentang Guru dan Dosen.*

Depdiknas. (2006). *Peraturan Menteri Pendidikan Nasional Nomor 22 Tahun 2006, tentang Standar Isi untuk Satuan Pendidikan Dasar dan Menengah.*

Depdiknas. (2007). *Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007, tentang Standar Penilaian Pendidikan.*

Dinas Dikpora DIY. (2013). *Peraturan Kepala Dinas Dikpora DIY Nomor 0111 Tahun 2013, tentang Petunjuk Teknis Penyelenggaraan Ujian Nasional SMP/MTs, SMP-LB, SMA/MA, SMALB DAN SMK Daerah Istimewa Yogyakarta Tahun Pelajaran 2012/2013.*

Kusaeri & Suprananto. (2012). *Pengukuran dan penilaian pendidikan.* Yogyakarta: Graha Ilmu.

Aqib, Z. & Rohmanto, E. (2007). *Membangun profesionalisme guru dan pengawas sekolah.* Bandung: Yrama Widya.

## MODEL EVALUASI PEMBELAJARAN AKIDAH DAN AKHLAK DI MADRASAH TSANAWIYAH (MTs)

<sup>1)</sup>Darodjat, <sup>2)</sup>Darmiyati Zuchdi, <sup>3)</sup>Zamroni

<sup>1)</sup>Universitas Muhammadiyah Purwokerto, <sup>2, 3)</sup>Universitas Negeri Yogyakarta  
<sup>1)</sup>darodjatjt@gmail.com, <sup>2)</sup>darmiyatiz@yahoo.com, <sup>3)</sup>zamronihardjowirono@yahoo.com

### Abstrak

Penelitian ini bertujuan: (1) menghasilkan model evaluasi pembelajaran Akidah dan Akhlak di MTs, (2) menghasilkan instrumen evaluasi yang memiliki reliabilitas dan validitas, (3) mengetahui kelayakan Model Logik Evaluasi Pembelajaran Akidah dan Akhlak (MLEPAA). Penelitian ini menggunakan model Borg dan Gall. Teknik penentuan subjek penelitian untuk unit analisis siswa dengan *simple random sampling* sejumlah 330 siswa dari tiga MTsN di Kabupaten Banyumas, 3 orang Guru Akidah dan Akhlak, dan 3 orang Kepala MTsN. Teknik pengumpulan data yang digunakan adalah: kuesioner, tes, dan penilaian antarteman. Validitas instrumen meliputi *face validity* dan *content validity*, dianalisis dengan *expert judgment*, validasi konstruk dianalisis dengan konfirmatori faktor analisis (CFA), sedangkan instrumen tes dianalisis dengan bantuan program ITEMAN. Estimasi reliabilitas instrumen menggunakan formula Alpha Cronbach dan reliabilitas interrater menggunakan Cohens' Kappa dengan program SPSS for Window 16.0. Simpulan penelitian: *pertama*, model evaluasi yang dihasilkan (MLEPAA) mencakup empat komponen utama, yaitu: *input, activities, output, dan outcomes*. *Kedua*, hasil analisis dengan CFA diperoleh indeks: (1) *Q-value* > 0,05; (2) *Root Mean Square Error of Approximation (RMSEA)* < 0,8; dan (3) *Goodness of Fit Index (GFI)* < 0,90. *Ketiga*, instrumen yang dikembangkan memenuhi validitas dan reliabilitas yang baik. *Keempat*, berdasarkan penilaian *user*, MLEPAA layak digunakan.

**Kata kunci:** *model logik, evaluasi pembelajaran Akidah dan Akhlak, MTs*

## AKIDAH AND AKHLAK LEARNING EVALUATION MODEL IN THE MADRASAH TSANAWIYAH

<sup>1)</sup>Darodjat, <sup>2)</sup>Darmiyati Zuchdi, <sup>3)</sup>Zamroni

<sup>1)</sup>Universitas Muhammadiyah Purwokerto, <sup>2, 3)</sup>Universitas Negeri Yogyakarta  
<sup>1)</sup>darodjatjt@gmail.com, <sup>2)</sup>darmiyatiz@yahoo.com, <sup>3)</sup>zamronihardjowirono@yahoo.com

### Abstract

This study aimed to: (1) develop an evaluation model on Akidah Akhlak subject of Madrasah Tsanawiyah (MTs), (2) generate some instruments of evaluation which have validity and reliability, (3) know a worthiness model evaluation of Akidah Akhlak subject in implementation. This study was a research and development of Borg and Gall model. The subjects were determined by simple random sampling which consisted of 330 students, 3 teachers, and 3 principles in three public Madrasah Tsanawiyah in Banyumas regencies. The data gathering techniques used were questionnaires, tests, and peer assessment. The instrument validation process was carried out in terms of the face validity and content validity through expert judgment, the construct validity through the confirmatory factor analysis (CFA), and test by using ITEMAN program. The instrument reliability was analyzed using the Cronbach's Alpha technique, and inter-rater reliability was assessed using SPSS for Window 16.0. with Cohens' Kappa technique. The results of the research were as follows: firstly, the components of MLEPAA evaluation model consisted of input, activities, output, and outcomes. Secondly, the results of CFA analysis were as follows: (a) Chi-Square,  $p > 0.05$ , the acquired value is = 553.72, (b) Goodness of Fit Indices (GFI) must be 0 to 1, the acquired value is = 0.892, and (c) Root Mean Square Error of Approximation (RMSEA) < 0.05, the acquired value is = 0.000. Thirdly, instrument utilized in research had fulfilled reliability and validity. Fourth, based on assessment from user, the MLEPAA was suitable for using in MTs.

**Keywords:** *logic model, Akidah Akhlak learning evaluation, MTs*

## Pendahuluan

Madrasah Tsanawiyah (MTs) merupakan salah satu lembaga pendidikan dasar dalam Sistem Pendidikan Nasional. Oleh karena itu, MTs harus melaksanakan ketentuan dalam standar nasional pendidikan. Ketentuan tersebut telah diatur dalam undang-undang, salah satunya ada di dalam Pasal 35 Ayat 1 Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional, yang mengatur tentang: standar isi, proses, kompetensi lulusan, tenaga kependidikan, sarana prasarana, pengelolaan, pembiayaan, dan penilaian pendidikan.

Standar nasional pendidikan merupakan dasar untuk penjaminan dan pengendalian mutu pendidikan. Pada proses pelaksanaan ketentuan standar nasional, masih banyak ditemukan kelemahan pada institusi madrasah, sehingga untuk mencapai tingkat kualitas berstandar nasional masih mengalami kendala. Menurut Fadjar (1998, pp.7-8), ada empat kelemahan dalam sistem pendidikan di madrasah, yaitu: (1) kurang menerapkan manajemen berbasis mutu, (2) sumber daya manusia yang kurang, (3) sistem pembelajaran dan evaluasi yang tidak tepat, dan (4) sarana-prasarana yang kurang mendukung.

Sejalan dengan temuan Fadjar tersebut, menurut Mastuhu (1999, p. 59), pembelajaran di madrasah juga masih banyak kelemahan, berupa: (1) lebih mementingkan materi (2) mementingkan memori di atas analisis dan dialog, (3) mementingkan penguatan pada “otak kiri” di atas “otak kanan”, (4) materi yang diberikan masih bersifat tradisional, belum menyentuh aspek rasional, (5) penekanan yang terlalu berlebihan pada ilmu sebagai produk final, bukan pada proses metodologinya, dan (6) mementingkan orientasi “memiliki” di atas “menjadi”.

Pada aspek metode, pembelajaran yang aktif, inovatif, kreatif, dan menarik belum digunakan dalam pembelajaran Akidah dan Akhlak. Materi ini masih bertumpu pada metode yang bersifat indoktrinatif, dan muatan materi berhenti pada tataran *normative-theocentric*. Kondisi tersebut dapat mele-

mahkan pembentukan sikap dan akhlak mulia peserta didik. Secara psikologis, usia peserta didik di Madrasah Tsanawiyah (MTs) berada pada masa remaja. Pada masa ini, peserta didik berada pada masa kritis dan kegoncangan jiwa, karena berada pada masa peralihan, dari masa anak-anak ke masa dewasa. Pada masa peralihan, terjadi percepatan pertumbuhan dalam segi fisik maupun psikis, baik ditinjau dari bentuk badan, sikap, cara berpikir dan bertindak. Mereka bukan lagi anak-anak. Mereka juga belum dikatakan manusia dewasa yang memiliki kematangan berpikir.

Masa kritis yang dialami oleh para remaja berkaitan dengan beralihnya keterikatan pada pola lama (orang tua) kepada pola baru/teman sebaya (*peer group*), sehingga mereka cenderung menghabiskan waktu mereka bersama teman sebaya dengan segenap aturan yang berbeda dengan aturan sebelumnya. Kondisi ini berpotensi memunculkan konflik batin antara diri remaja dengan orang tua dan diri remaja dengan teman sebaya. Keadaan ini secara simultan mengondisikan remaja banyak melakukan tindak pelanggaran (LeBlanc, et al., 2008; Kerr, Stattin, & Burk, 2010). Orang tua berperan besar dalam mengarahkan dan menjaga agar para remaja tidak terjebak pada tindakan yang merusak diri dan masa depan mereka.

Kondisi kritis lainnya yang dialami remaja adalah adanya tuntutan dalam diri remaja akibat perkembangan fisik dan psikis, dan tuntutan dari lingkungan di luar dirinya (*peer group*) yang harus dipenuhinya (*social adjustment*). Dalam proses penyesuaian diri pada diri remaja muncul kegelisahan batin, kurang percaya diri, kelesuan/tidak bersemangat, banyak berangan-angan, sehingga memunculkan stres, bahkan depresi dan tindak perilaku negatif lainnya seperti: kenakalan remaja, penyalahgunaan zat adiktif, dan bunuh diri (Gould, et.al., 2003; Barber & Olsen, 2004; Newman, et.al., 2007). Oleh karena itu, dibutuhkan pendekatan pembelajaran yang humanis, kondusif, dan memperhatikan kecenderungan remaja tersebut.

Pembelajaran Akidah dan Akhlak sangat berperan bagi proses penyelamatan masa kritis yang dialami remaja, dan sekaligus berfungsi mengembangkan potensi yang ada pada dirinya. Akidah yang berintikan kepercayaan terhadap eksistensi Allah dapat berfungsi sebagai faktor pendorong dan pengarah agar semua aktivitas dan ibadah dalam makna yang luas dilakukan hanya mencari keridaan Allah (surat *al-An'ām*: 162). Peserta didik yang telah memiliki akidah yang mantap dan benar, maka akan selalu merasakan bahwa ke mana pun dirinya berada akan selalu diawasi oleh Allah. Akidah mengajarkan bahwa Allah itu Mahamendengar dan Mahamelihat setiap gerak dan keadaan makhluk-makhluk-Nya (surat *Yunus*: 61, *al-Baqā-rab*: 129), Mahasempurna dari segala kekurangan dan kelemahan (surat *ar-Rahmān*: 27). Sedangkan, akhlak sebagai perwujudan dari keimanan atau kepercayaan yang mantap dalam dirinya akan menjadi motivator dan landasan untuk selalu mengaktualisasikan keimanan tersebut dalam cara berpikir, bersikap, dan bertindak dalam kerangka mencari keridaan-Nya.

Pembelajaran Akidah dan Akhlak adalah upaya sadar dan terencana dalam menyiapkan peserta didik untuk mengenal, memahami, menghayati dan mengimani Allah, dan merealisasikannya dalam kehidupan sehari-hari melalui kegiatan bimbingan, pengajaran, latihan, penggunaan pengalaman, dan kebiasaan. Dalam kehidupan masyarakat yang majemuk dalam bidang keagamaan, pembelajaran ini juga diarahkan pada peneguhan akidah di satu sisi, dan peningkatan toleransi serta saling menghormati dengan penganut agama lain dalam rangka mewujudkan kesatuan dan persatuan bangsa (Peraturan Menteri Agama RI Nomor 2 tahun 2008 tentang Standar Kompetensi Lulusan dan Standar Isi Pendidikan Agama Islam dan Bahasa Arab di Madrasah).

Implementasi pembelajaran Akidah dan Akhlak di madrasah (MTs) masih belum kondusif sesuai dengan Peraturan Menteri Agama RI tersebut. Pembelajaran Akidah dan Akhlak masih menekankan pada sisi

*normative-theocentric*, dan banyak menggunakan strategi indoktrinatif. Kondisi ini dapat mengurangi efektivitas pencapaian tujuan dan fungsi pendidikan yang sangat strategis bagi pengembangan potensi peserta didik, yaitu: menjadi manusia yang bertakwa, berakhlak mulia, sehat, berilmu, cakap kreatif, mandiri dan bertanggung jawab (Undang-Undang RI Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional Pasal 3).

Penerapan metode indoktrinasi dalam pembelajaran mengandung banyak kelemahan, khususnya dalam pengembangan kepribadian siswa. Menurut Chazan dan Soltis (1975, pp. 39-40) kelemahan tersebut karena: (1) indoktrinasi memanifestasikan proses belajar mengajar yang menekankan pada *teacher centered*, dan siswa sebagai objek yang pasif. (2) indoktrinasi dapat mengondisikan proses *depersonalization*, sebab indoktrinasi lebih mengakui eksistensi masa lampau, sehingga cenderung membatasi potensi masing-masing individu peserta didik. Sebaliknya, melalui indoktrinasi menempatkan manusia, ruang, dan ideologi tertentu sebagai model pengaruh yang dianggap paling signifikan dan dibutuhkan.

Metode yang bersifat indoktrinatif juga dapat mengondisikan pembelajaran Akidah Akhlak menjadi tidak menarik, kelas menjadi pasif, peserta didik kurang termotivasi untuk mengikuti pembelajaran, serta menimbulkan sikap yang kurang positif di kalangan peserta didik terhadap pembelajaran. Sarana-prasarana yang kurang, kinerja guru yang tidak optimal, secara simultan menambah problem yang ada dalam pendidikan Akidah dan Akhlak di MTs. Pendidikan Akidah dan Akhlak yang cenderung indoktrinatif juga dapat menumbuhkan kekeerasan atas nama agama.

Berkaitan dengan upaya mengatasi kelemahan penggunaan metode indoktrinasi yang masih diterapkan pada pembelajaran Akidah dan Akhlak, Zuchdi (2010, p. 6) menjelaskan bahwa sekiranya metode indoktrinasi tidak dapat dihindari, maka harus ada usaha mengatasi kelemahan tersebut, di antaranya yaitu: *pertama*, sekolah bersama dengan seluruh komponennya perlu men-

ciptakan *setting* sosial yang memungkinkan implementasi pengetahuan yang telah diperolehnya untuk memecahkan masalah atau problem yang sedang dihadapi masyarakat. *Kedua*, peserta didik dirangsang atau difasilitasi agar mereka menemukan alasan-alasan yang mendasari keputusan moral. Dengan cara demikian, pendidikan yang diselenggarakan di tengah-tengah masyarakatnya, termasuk madrasah akan lebih efektif dalam membina peserta didik memiliki akhlak mulia, memiliki kepribadian yang utuh dan bersikap humanis. Menurut al-Abrasyi (1984, p. 5) tujuan yang sebenarnya dari pendidikan Islam adalah terbentuknya akhlak mulia pada diri siswa.

Adapun fungsi pembelajaran Akidah dan Akhlak sebagaimana tercantum dalam Peraturan Menteri Agama RI Nomor 2 Tahun 2008 adalah: (1) penanaman nilai ajaran Islam untuk mencapai kebahagiaan hidup, baik dunia maupun akhirat (2) pengembangan keimanan, ketakwaan, dan akhlak mulia peserta didik seoptimal mungkin, (3) penyesuaian mental peserta didik terhadap lingkungan fisik dan sosial, (4) perbaikan kesalahan, kelemahan, keyakinan, dan pengamalan ajaran agama Islam (5) pencegahan dari hal-hal negatif dari lingkungannya atau dari budaya asing, (6) pengajaran tentang informasi, pengetahuan keimanan dan akhlak, serta sistem dan fungsionalnya, (7) penyaluran peserta didik untuk mendalami Akidah dan Akhlak pada jenjang pendidikan yang lebih tinggi.

Implementasi pembelajaran Akidah dan Akhlak di MTs untuk mencapai tujuan tersebut harus dievaluasi dan dikembangkan sesuai dengan kebutuhan dan perkembangan masyarakat. Evaluasi dan pengembangan ini tidak hanya menyangkut materi, tetapi juga pengembangan pada aspek kinerja guru, motivasi dan sikap siswa, fasilitas pembelajaran, iklim kelas, metode pembelajaran, serta sarana pembelajaran. Dengan kata lain, evaluasi dalam konteks model logik harus menyeluruh baik menyangkut *input*, proses, *output* maupun *outcome* pembelajarannya.

Berdasarkan studi pendahuluan terhadap pelaksanaan pembelajaran Akidah dan Akhlak MTs di Kabupaten Banyumas pada Juli 2009 ditemukan bahwa: (1) terdapat kelemahan dalam pembelajaran Akidah dan Akhlak di MTs, kelemahan ini menyangkut: (1) metode yang cenderung indoktrinatif, (2) isi/materi yang kurang kontekstual dan tidak diarahkan pada *setting* sosial sebagai upaya untuk mengamalkan ajaran Islam dalam kehidupan sehari-hari, (3) keteladanan dari semua guru dan tenaga kependidikan dalam penerapan nilai akidah dan akhlak belum kondusif, sehingga implementasi nilai-nilai tersebut menjadi kurang efektif, (4) penciptaan lingkungan madrasah dalam penerapan nilai-nilai akidah dan akhlak belum terpadu dan terintegrasi, sehingga pembiasaan akhlak mulia peserta didik belum teraktualisasikan dalam kehidupan keseharian di lingkungan madrasah, (5) sikap dan motivasi belajar peserta didik yang rendah dalam mengikuti pembelajaran, (6) evaluasi pembelajaran belum menggunakan model yang komprehensif.

Sepanjang hasil penelitian pendahuluan, belum ada model evaluasi pembelajaran Akidah dan Akhlak di tingkat MTs yang secara komprehensif (meliputi unsur *input*, *proses*, *output* dan *outcomes*) dapat memberikan informasi secara tepat dan akurat bagi guru dan kepala madrasah terhadap pembelajaran Akidah dan Akhlak, baik dari segi isi, cakupan, format, metode, maupun waktu penyampaian. Oleh karena itu, penelitian ini dilakukan dalam rangka pengembangan model evaluasi tersebut.

Tujuan penelitian ini adalah: (1) menghasilkan produk berupa model evaluasi, (2) menghasilkan instrumen evaluasi pembelajaran yang valid dan reliabel untuk mengevaluasi pembelajaran Akidah dan Akhlak di MTsN, dan (3) mengetahui kelayakan model logik dan instrumennya dalam penerapan di lapangan.

Model evaluasi dikembangkan dari model logik. Secara keseluruhan model logik ini terdiri dari tujuh komponen, yaitu: *goal*, *rationale*, *assumption*, *input*, *activities*, *output*, dan *outcome* (Kellogg Foundation, 2004, p.

4., Powell & Heneret, 2008, p. 10, dan Porteous, 2002, pp. 16-117). Fokus penelitian dan pengembangan ini pada empat komponen yang terakhir, yaitu: *input*, *activities*, *output*, dan *outcomes*. Tiga komponen pertama yaitu: *goal*, *rationale*, dan *assumption* merupakan komponen yang menjadi dasar penyusunan model logik atau sebagai *need assessment*, sehingga tersusun empat komponen yang terakhir dengan prinsip *if-then relationship*.

Pada komponen *input*, ada 5 variabel yaitu: (1) kinerja guru akidah akhlak, (2) materi dan metode, (3) sarana pembelajaran, (4) kultur madrasah dan kelas, dan (5) kepemimpinan kepala madrasah. Berikutnya, komponen *activities*, meliputi tiga aktivitas pembelajaran yaitu: kegiatan pendahuluan, kegiatan inti, dan kegiatan penutup pembelajaran. Dua komponen terakhir dari model logik adalah *output* dan *outcomes*. Untuk komponen *output*, perubahan hasil pembelajaran yang akan dievaluasi adalah: perubahan motivasi belajar, perubahan pengetahuan, dan perubahan sikap berakidah akhlak siswa MTs. Sedangkan evaluasi pada komponen *outcomes* difokuskan pada perubahan perilaku berakidah akhlak siswa MTs di luar kelas dalam area atau wilayah madrasah.

### Metode Penelitian

Jenis penelitian ini adalah penelitian pengembangan (*research and development, R&D*), menggunakan model Borg & Gall (1983, p. 775) yang terdiri dari 10 langkah. Namun, dalam penelitian ini disederhanakan menjadi tiga langkah sebagaimana yang direkomendasikan oleh Cennamo & Kalk (2005, p. 6), sehingga tahapannya menjadi: (1) tahap prapengembangan model, (2) tahap pengembangan model, dan (3) implementasi model. Esensi tahapan penelitian dan pengembangan model Borg & Gall tetap dipakai.

### Desain Uji Coba

Uji coba dimaksudkan untuk memperoleh data secara lengkap yang dapat digunakan sebagai bahan revisi produk yang

dihasilkan. Uji coba dilakukan tiga tahap yaitu: uji coba dengan sampel kecil (30 siswa), sedang (50 siswa) dan sampel besar atau *operational field testing* (255 siswa). Setelah dilakukan uji coba pada tiap tahap, maka dilakukan validasi, sehingga didapatkan instrumen yang valid dan reliabel. Bagian lain yang diujicobakan adalah: (1) model evaluasi, (2) panduan model evaluasi, dan (3) keterbacaan instrumen model evaluasi.

### Teknik Analisis Data

Instrumen pengumpul data dengan responden peserta didik dianalisis dengan Confirmatory Factor Analysis (CFA). Tujuan dari CFA ini adalah untuk: pertama, menganalisis validitas instrumen pengumpul data. Jika indeks muatan faktor ( $\lambda$ ) nilainya  $> 0,3$ , maka dinyatakan valid, karena dapat menjelaskan konstruk atau variabel yang ada. Sebaliknya, jika butir memiliki muatan faktor/*factor loading* lebih kecil dari  $0,3 (< 0,3)$  maka butir tersebut harus dibuang atau gugur (Fernandes, 1984, p. 28). Berkaitan dengan validitas, McMillan & Scumacher (2010, p. 116) menyarankan agar didapatkan validitas eksternal yang tinggi pada desain penelitian kuantitatif, maka harus diperhatikan faktor berikut, yaitu: (1) subjek penelitian (populasi) harus memiliki karakteristik yang sama, (2) kondisi penelitian dan hasil yang akan digeneralisasi memiliki kondisi yang sama. Oleh karena itu, faktor-faktor berikut harus dianulir yaitu: (1) pengaruh interaksi seleksi yang bias, (2) pengaruh interaksi pretest, sebab subjek yang diberi pretest akan memberikan respon yang berbeda dengan subjek yang tidak diberi pretest, (3) pengaruh reaktif dari prosedur eksperimental, pengaruh yang muncul dari setting eksperimental yang tidak akan terjadi pada setting noneksperimental, (4) pengaruh perlakuan yang berulang-ulang terhadap subjek yang sama, sebab perlakuan tersebut akan berpengaruh terhadap perlakuan berikutnya karena pengaruh yang terdahulu tidak dapat dihilangkan.

Kedua, tujuan dari CFA adalah mengidentifikasi dimensi instrumen, menguji apakah dimensi tersebut dapat dikonfirmasi,

serta cocok dengan data empirik yang sesungguhnya. Sedangkan, model logik akan dianalisis secara deskriptif berdasarkan penilaian dari guru Akidah dan Akhlak, kepala madrasah, dan Pengawas PAI. Disamping itu, data yang bersumber dari siswa dianalisis dengan menggunakan *Structural Equation Modeling* untuk menguji kecocokan antara model teoritis dengan data empiris sehingga didapatkan model yang baik. Menurut Ghazali (2005, pp. 31-32), kriteria model fit bila memenuhi kriteria: (1) Chi-Square yang diperoleh dari hasil pengujian memiliki probabilitas lebih besar dari 0,05 ( $p > 0,05$ ); (2) *Root Mean Square Error of Approximation* (RMSEA)  $\leq 0,08$ ; (3) *Goodness of Fit Index* (GFI)  $> 0,90$ .

Data yang telah terkumpul kemudian dianalisis dengan menggunakan teknik analisis deskriptif, yaitu mendeskripsikan dan memaknai tiap komponen data evaluasi, kemudian membandingkan dengan kriteria yang telah ditentukan sebelumnya. Berdasarkan hasil deskripsi tersebut dapat dijadikan dasar penilaian terhadap model evaluasi pembelajaran Akidah dan Akhlak yang dikembangkan. Kriteria yang digunakan sebagai pedoman untuk menganalisis dari data kuantitatif ke data kualitatif merujuk Sudijono (2006, p. 125), sebagaimana tampak pada Tabel 1.

Tabel 1. Kriteria Penilaian

Skor (X)	Kriteria	Simpulan
$X > 4,2$	Sangat baik	Dapat dijadikan contoh
$X > 3,4-4,2$	Baik	Dapat digunakan tanpa perbaikan
$X > 2,6-3,4$	Cukup	Dapat digunakan, sedikit perbaikan
$X > 1,6-2,8$	Kurang	Dapat digunakan, banyak perbaikan
$X < 1,8$	Tidak baik	Belum dapat digunakan

### Hasil Penelitian dan Pembahasan

Produk yang dihasilkan dari penelitian dan pengembangan ini adalah dihasilkan sebuah model evaluasi dan perangkatnya yang dapat digunakan untuk mengevaluasi pembelajaran Akidah dan Akhlak di MTs. Tahap yang dilakukan meliputi: prapengembangan, pengembangan, dan penerapan.

*Pertama*, tahap prapengembangan dilakukan: (a) observasi ke MTsN di Kabupaten Banyumas dari bulan Juni 2012 sampai bulan September 2012, (b) *interview* kepada guru Akidah dan Akhlak, dan (3) pemberian angket kepada *user* (guru Akidah dan Akhlak, kepala MTsN, dan pengawas Pendidikan Agama Islam). Hasil observasi, *interview*, dan pemberian angket tersebut digunakan sebagai data awal bagi *need assessment*, yaitu kajian tentang kebutuhan pengembangan model evaluasi pembelajaran Akidah dan Akhlak di kelas VIII MTsN. Hasil *need assessment* tersebut kemudian diwujudkan dalam bentuk prototipe model awal.

*Kedua*, tahap pengembangan, pada tahap ini dilakukan penentuan dan pembuatan desain evaluasi pembelajaran, kemudian divalidasi oleh para ahli. Ahli yang terlibat dalam validasi ini meliputi bidang: evaluasi pendidikan, metodologi penelitian, pengukuran/konstruksi instrumen, pendidikan Islam, pendidikan karakter, dan Bahasa Indonesia. Mereka dimintai pendapat tentang model logik evaluasi pembelajaran Akidah dan Akhlak di MTsN. Pendapat mereka diberikan dalam bentuk skala Likert, dengan skor tertinggi 5 dan skor terendah. Penskoran amat baik= 5, baik= 4, cukup= 3, kurang = 2, dan sangat kurang= 1. Selain itu, validator juga diberi ruang untuk memberikan pendapat, usulan, dan saran berkaitan dengan model tersebut dan perangkat-perangkatnya.

Hasil validasi menunjukkan bahwa evaluasi pembelajaran menggunakan model logik dan perangkatnya mendapatkan rerata kategori baik, sehingga dapat digunakan untuk tahapan penelitian lebih lanjut. Setelah mendapat penilaian dan masukan dari para ahli, maka dilakukan penyempurnaan model dan perangkatnya. Hasil penyempurnaan tersebut selanjutnya dikonsultasikan kembali untuk mendapatkan pengesahan, dan dikonsultasikan ke *Promotor* dan *Copromotor*. Selanjutnya, draf model dan perangkat yang telah dinilai oleh ahli diberikan kepada *user* untuk dinilai dan diberi masukan. Unsur *user* yang terlibat adalah: guru Akidah dan Akhlak MTs, Kepala MTsN,

dan Pengawas Pendidikan Agama Islam (PPAI).

*Ketiga*, tahap penerapan, setelah model beserta perangkatnya mendapat penilaian dan pengesahan dari *expert*, *user*, dan teman sejawat, maka instrumen tersebut diujicobakan pada kelompok kecil, sedang, dan uji coba utama/operasional dengan melibatkan peserta didik kelas VIII di MTsN se-Kabupaten Banyumas. Berikut ditampilkan data hasil pengembangan yang dicapai berkaitan dengan model logik evaluasi pembelajaran Akidah dan Akhlak di MTsN.

#### Hasil Expert Judgment dan Validasi User

Penilaian terhadap model dan perangkatnya oleh ahli, *user*, dan teman sejawat dapat dideskripsikan sebagai berikut: *pertama*, jumlah keseluruhan penilai ada 17 orang, terdiri dari: 5 orang ahli (*expert*), 8 orang *user*, dan 4 orang teman sejawat. Perincian unsur ahli terdiri dari: 1 orang pakar bidang metodologi evaluasi pendidikan, 1 orang pakar pengukuran pendidikan, 1 orang pakar bidang metodologi penelitian, 1 orang pakar bidang Pendidikan Islam, dan 1 orang pakar bidang bahasa Indonesia, dilaksanakan pada bulan Juni 2013 sampai bulan September 2013.

*Kedua*, perincian 8 orang *user* terdiri dari unsur: 3 orang guru Akidah dan Akhlak di tiga MTsN, 3 orang kepala madrasah yang berasal dari tiga MTsN, dan 2 orang pengawas Pendidikan Agama Islam pada Kementerian Agama Kabupaten Banyumas. Penilaian dilakukan pada tanggal 16-21 September 2013. *Ketiga*, validasi/penilaian oleh teman sejawat dilakukan oleh 4 mahasiswa program doktor prodi PEP UNY yang dilaksanakan pada tanggal 24 Juni sampai 8 Juli 2013.

Perincian hasil penilaian model dan perangkatnya oleh ahli (*expert judgment*), *user*, dan teman sejawat adalah sebagai berikut: (a) rerata skor total hasil penilaian keterbacaan instrumen model evaluasi oleh 17 orang sebesar 3.8. Nilai tersebut jika dikonversikan dengan kriteria penilaian data kuantitatif ke data kualitatif dengan skala 5 masuk dalam kategori baik; (b) rerata total

skor hasil penilaian Panduan Model Evaluasi oleh 17 orang sebagaimana diperinci di atas sebesar 3.8. Nilai tersebut jika dikonversikan dengan kriteria penilaian data kuantitatif ke data kualitatif dengan skala 5 masuk dalam kategori baik; (c) rerata total skor hasil penilaian keefektifan model evaluasi oleh 17 orang sebagaimana diperinci di atas sebesar 3.6. Nilai tersebut jika dikonversikan dengan kriteria penilaian data kuantitatif ke data kualitatif dengan skala 5 masuk dalam kategori baik; (d) validasi terhadap panduan model evaluasi yang dilakukan oleh ahli (*expert*), *user*, dan teman sejawat didapatkan rerata total skor sebesar 3.5. Nilai tersebut jika dikonversikan dengan kriteria penilaian data kuantitatif ke data kualitatif dengan skala 5 masuk dalam kategori baik.

#### Hasil Uji Coba Produk

Uji coba tahap pertama dilaksanakan tanggal 10-19 November 2013, disebut juga dengan uji coba pendahuluan atau *preliminary field testing* menurut Borg & Gall. Uji coba ini dilakukan dengan membagikan instrumen evaluasi pembelajaran Akidah dan Akhlak model logik yang meliputi komponen *input*, proses, produk, dan *outcomes* kepada 30 peserta didik di tiga MTsN di Kabupaten Banyumas, yaitu: MTsN Tambak, MTsN Sumbang, dan MTsN Model. Rerata keseluruhan skor hasil uji coba tahap pertama kepada 30 siswa di tiga MTsN terhadap keterbacaan instrumen model evaluasi pembelajaran Akidah dan Akhlak sebesar 3,7. Nilai rerata total skor ini jika dikonversikan dengan kriteria penilaian data kuantitatif ke data kualitatif dengan skala 5 sebagaimana telah ditetapkan sebelumnya masuk dalam kategori baik.

Pada uji coba tahap II, ada dua karakteristik yang menonjol yaitu: (a) subjek penelitian difokuskan pada siswa, karena kata kunci dalam komponen model logik yang kedua yaitu *activities* berkenaan dengan siswa, yaitu *what we do in MTsN (learning)*, dan *who we reach in learning (siswa)*; (b) pada tahap ini jumlah siswa yang dilibatkan lebih besar dibandingkan pada uji coba tahap pertama,

yaitu 50 siswa kelas VIII pada tiga MTsN di Kabupaten Banyumas. Uji coba tahap kedua dilaksanakan tanggal 20-27 Maret 2014; (c) uji coba ini dimaksudkan untuk mengestimasi reliabilitas instrumen dan membuktikan validitas instrumen yang digunakan dalam penelitian. Berdasarkan hasil analisis dengan bantuan program *SPSS for Windows* 16,0 dapat diketahui indeksinya. Parameter yang digunakan untuk mengetahui reliabilitas adalah dengan melihat nilai *Alpha Cronbach* pada setiap tabel *output* yang diperoleh, menurut Nunally (1981, p. 230) jika indeks Alphanya  $> 0,7$ , maka instrumen tersebut termasuk reliabel. Melalui teknik CFA (*Confirmatory Factor Analysis*) juga dapat diketahui apakah setiap indikator yang diestimasi secara valid mengukur dimensi dari konsep yang diujinya, yaitu dengan cara melihat nilai *factor loading*.

Berikut ini ditampilkan hasil uji coba instrumen komponen model evaluasi disertai dengan estimasi reliabilitas dan validitasnya: *pertama*, pada uji coba tahap II, instrumen untuk mengukur kinerja guru Akidah dan Akhlak ada 2 macam, *pertama* menggunakan kuesioner dalam bentuk *self assessment* (guru Akidah dan Akhlak yang menilai dirinya sendiri), *kedua* menggunakan kuesioner dengan responden para siswa. Berdasarkan *print out* SPSS dapat dijelaskan bahwa hasil estimasi reliabilitas dan validitas kuesioner kinerja guru Akidah dan Akhlak dengan responden 50 siswa MTsN di Kabupaten Banyumas menunjukkan bahwa instrumen kinerja guru Akidah dan Akhlak yang masuk dalam komponen *input* memiliki indeks reliabilitas yang baik, karena koefisien Alphanya  $> 0,7$  yaitu 0,905. Demikian halnya dengan indeks yang diperoleh pada semua indikator yang diestimasi dengan teknik CFA menunjukkan bahwa keseluruhan instrumen kinerja guru Akidah dan Akhlak yang berjumlah 25 mempunyai nilai faktor *loading*  $> 0,5$ . Dengan demikian, semua indikator variabel yang digunakan dalam kuesioner ini adalah valid, kecuali nomor 1 yang mempunyai nilai faktor *loading*  $< 0,5$  sehingga dinyatakan gugur.

*Kedua*, pada uji coba tahap II, instrumen materi dan metode pembelajaran Akidah dan Akhlak respondennya adalah 50 siswa MTsN di Kabupaten Banyumas. Estimasi reliabilitas dan validitas instrumen tersebut dapat disimpulkan bahwa instrumen materi dan metode pembelajaran Akidah dan Akhlak yang masuk dalam komponen *input* memiliki reliabilitas yang baik, karena koefisien Alphanya lebih dari 0,7 yaitu sebesar 0,914. Selanjutnya, terdapat 3 item kuesioner yang mempunyai nilai faktor *loading*  $< 0,5$ , yaitu pada item: MM1\_6, MM1\_14, dan MM1\_26. Dengan demikian, ketiga item instrumen tersebut termasuk yang tidak baik/gugur dan tidak dimasukkan dalam instrumen untuk pengambilan data pada tahap III (*main field testing*). Selain dari tiga butir instrumen tersebut dinyatakan sebagai indikator yang baik karena nilai faktor *loading* lebih besar dari 0,5.

*Ketiga*, pada uji coba tahap II terhadap instrumen sarana/fasilitas pembelajaran dalam ruang kelas didapatkan koefisien *Alpha Cronbach* sebesar 0,843. Instrumen dinyatakan reliabel jika nilai *Alpha Cronbach*  $> 0,70$ . Dengan demikian, instrumen sarana dalam ruang kelas (SP-1) adalah reliabel. Selanjutnya, secara keseluruhan instrumen ini memiliki muatan faktor yang baik, karena nilai faktor *loading*-nya lebih besar dari 0,5.

*Keempat*, pada uji coba tahap II, instrumen kultur madrasah dan kelas respondennya adalah 50 siswa di tiga MTsN, yaitu: MTsN Model, MTsN Sumbang, dan MTsN Tambak. Hasil uji coba instrumen tersebut didapatkan koefisien reliabilitas *Alpha Cronbach*  $> 0,70$ . Namun, ada dua butir instrumen yang mempunyai nilai faktor *loading*  $< 0,5$  yaitu pada item: KM1\_5 dan KM1\_9. Dengan demikian, kedua butir instrumen tersebut dinyatakan gugur, sehingga tidak dimasukkan lagi dalam instrumen untuk pengambilan data pada tahap III (*main field testing*).

*Kelima*, instrumen untuk mengetahui kultur madrasah dan kelas digunakan dua instrumen, yaitu: instrumen angket dengan kode KM-1 dan angket dengan kode KM-3 dalam bentuk *semantic differential*. Hasil esti-

masi reliabilitas instrumen kultur madrasah dan kelas dalam bentuk *semantic differential* didapatkan koefisien *Alpha Cronbach* sebesar 0,831. Dengan demikian, dapat disimpulkan bahwa instrumen kultur madrasah dan kelas menggunakan *semantic differential* adalah reliabel. Selanjutnya, perihal validitas instrumen kultur madrasah dan kelas, dari KM3\_1 sampai KM3\_15 memiliki nilai faktor *loading* di atas 0,5. Dengan demikian, instrumen kultur madrasah dan kelas dalam bentuk *semantic differential* memiliki validitas yang baik.

*Keenam*, responden uji coba tahap II pada pembelajaran Akidah Akhlak adalah siswa kelas VIII di 3 MTsN, yaitu: MTsN Model Purwokerto, MTsN Sumbang, dan MTsN Tambak, keseluruhan respondennya berjumlah 50 siswa. Estimasi terhadap reliabilitas instrumen ini didapatkan koefisien *Alpha Cronbach* sebesar 0,856, jadi nilai ini > 0,70. Dengan demikian dapat disimpulkan bahwa instrumen Pembelajaran Akidah dan Akhlak adalah reliabel. Selanjutnya, estimasi terhadap validitas instrumen didapatkan bahwa tidak ada satu pun butir item (dari PAA1\_1 sampai dengan butir PAA1\_20) yang nilai faktor *loading*-nya < 0,5. Dengan demikian, dapat disimpulkan bahwa instrumen Pembelajaran Akidah dan Akhlak (PAA-1) memiliki validitas yang baik.

*Ketujuh*, uji coba instrumen perubahan motivasi belajar siswa dalam pembelajaran Akidah dan Akhlak pada tahap II ini melibatkan 50 siswa kelas VIII MTsN di Kabupaten Banyumas, terdiri dari MTsN Model Purwokerto, MTsN Sumbang, dan MTsN Tambak. Estimasi reliabilitas instrumen tersebut didapatkan *Alpha Cronbach* sebesar 0,96. Instrumen dinyatakan reliabel jika nilai *Alpha Cronbach* > 0,70. Selanjutnya, estimasi terhadap validitas instrumen didapatkan nilai faktor *loading* pada masing-masing item > 0,5. Dengan demikian dapat disimpulkan bahwa instrumen/kuesioner perubahan motivasi belajar siswa adalah valid.

*Kedelapan*, instrumen yang digunakan untuk mengetahui perubahan pengetahuan siswa (komponen *output*) sebagai hasil proses pembelajaran Akidah dan Akhlak adalah

dengan instrumen tes. Instrumen ini diberikan kepada 51 siswa kelas VIII MTsN di Kabupaten Banyumas. Berdasarkan hasil analisis dengan program ITEMAN, maka dapat diketahui bahwa dari ke 37 item butir instrumen, ada 8 item butir yang dikategorikan tidak baik/gugur. Butir yang memiliki tingkat kesukaran soal antara 0,25-0,75 dikategorikan sebagai butir yang baik. Tingkat kesukaran soal dalam *print out* ITEMAN adalah pada kolom *point biserial* dan *biserial*. Menurut Croker dan Algina (1986, p. 135) indeks daya beda soal yang intervalnya berkisar antara 0,19-0,00 termasuk dalam kategori soal yang tidak boleh dipakai/harus dibuang, sedangkan jika indeks daya bedanya berkisar antara 0,20-0,29 maka soal harus diperbaiki. Butir soal yang tidak baik ada 8 butir, yaitu: butir soal nomor 6, 10, 14, 20, 21, 30, 33, dan 34. Kedelapan butir soal yang gugur tersebut tidak dimasukkan dalam instrumen pada pengambilan data tahap III (*main field testing*). Selanjutnya, instrumen tes tersebut dinyatakan reliabel karena nilai *Alpha Cronbach* yang didapat (0,880) dan nilai tersebut lebih besar dari 0,70.

*Kesembilan*, uji coba tahap II terhadap instrumen perubahan sikap berakidah dan berakhlak melibatkan siswa kelas VIII pada tiga MTsN di Kabupaten Banyumas, dengan jumlah sampel sebanyak 50 siswa. Berdasarkan kriteria, instrumen dinyatakan reliabel jika koefisien *Alpha Cronbach* > 0,70. Nilai *Alpha Cronbach* yang didapat sebesar 0,960. Dengan demikian, dapat disimpulkan bahwa instrumen perubahan sikap berakidah dan akhlak (OP-3) adalah reliabel.

Perihal estimasi validitas instrumen dapat dijelaskan sebagai berikut: dua item butir instrumen yang tidak baik, yaitu butir OP3\_11 dan OP3\_14, karena kedua item butir tersebut mempunyai nilai *r*-hitung (*corrected item total correlation*) < 0,3, sehingga harus dibuang dan tidak dipergunakan dalam instrumen pada tahap III. Sedangkan, 38 item termasuk baik, sehingga dapat dipergunakan lagi pada tahap III.

*Kesepuluh*, untuk mengetahui perubahan perilaku berakidah akhlak peserta didik dengan subjek penelitian 50 siswa kelas VIII

MTsN di Kabupaten Banyumas, yaitu dengan penilaian antarteman. Estimasi reliabilitas instrumen tersebut didapatkan *Alpha Cronbach* (0,844) > 0,70. Dengan demikian, instrumen tersebut adalah reliabel. Selanjutnya, estimasi *Intraclass Correlation Coefficient/ICC* sebesar 0,731. Koefisien sebesar 0,731 menunjukkan bahwa reliabilitas interrater-nya baik (Eye & Mun, 2005, p. 6). Hal ini menunjukkan pula bahwa dua rater atau penilai terhadap perubahan perilaku berakidah yang berasal dari penilai yang berbeda menyetujui tentang bagaimana mereka memberi skor kepada seorang peserta didik.

Uji coba tahap ketiga atau disebut juga uji coba operasional (*main field testing*), melibatkan subjek lebih luas yaitu 255 siswa kelas VIII MTsN di Kabupaten Banyumas. Uji coba operasional ini dilaksanakan pada tanggal 21-24 Juni 2014 di tiga MTsN di Kabupaten Banyumas, yaitu: MTsN Model Purwokerto, MTsN Sumbang, dan MTsN Tambak pada kelas VIII. Tujuan uji coba operasional (*main field testing*) dari instrumen angket adalah untuk memperoleh gambaran model evaluasi pembelajaran Akidah Akhlak dalam skala yang lebih luas. Uji coba terhadap instrumen tes dimaksudkan untuk menghasilkan instrumen tes yang baik yang dapat mengukur kemajuan/perubahan pengetahuan Akidah dan Akhlak siswa MTs kelas VIII. Jumlah responden 140 siswa dengan perincian: 47 siswa dari MTsN Model, 47 siswa dari MTsN Sumbang, dan 47 siswa dari MTs Tambak.

Berdasarkan hasil analisis menggunakan program ITEMAN dapat disimpulkan bahwa instrumen tes tersebut adalah reliabel, dengan koefisien Alpha (0,836) > 0,70. Sedangkan validitas tes secara keseluruhan adalah baik, namun ada empat butir yang gugur, yaitu nomor: 7, 13, 21, dan 29.

Pada tahapan selanjutnya adalah dilakukan uji *fit model*. Suatu model dikatakan *fit* jika memenuhi persyaratan di antaranya adalah: *pertama*, *Chi-Square* yang diperoleh dari hasil pengujian memiliki probabilitas ( $p$ ) lebih besar dari 0,05 ( $p > 0,05$ ). *Kedua*, *GFI* harus berkisar antara 0 dan 1. Nilai

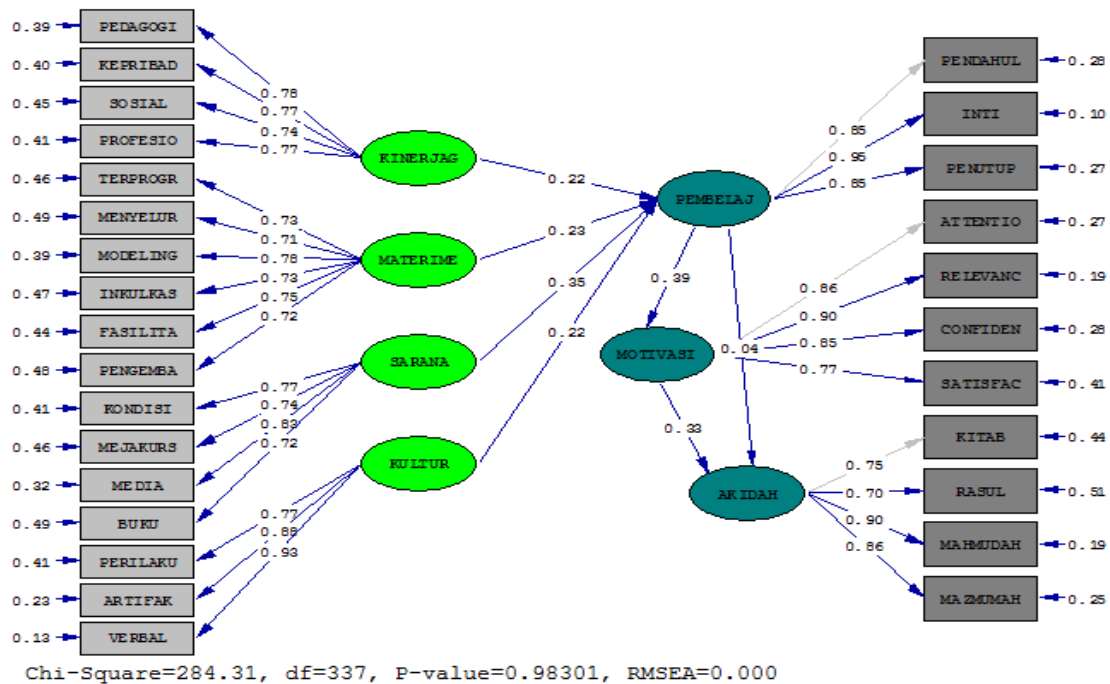
*GFI* > 0,9 menunjukkan *fit* suatu model yang baik. *GFI* merupakan suatu ukuran mengenai ketepatan model dalam menghasilkan *observed* matriks kovarians. Hasil perhitungan dengan program LISREL 8,50 diperoleh koefisien *GFI* = 0,926. *Ketiga*, *Root Mean Square Error of Approximation (RMSEA)* < 0,05. *RMSEA* > 0,1 mengindikasikan model yang tidak *fit*. Koefisien *RMSEA* yang berkisar antara 0,08 sampai dengan 0,1 merupakan model yang cukup *fit*, sedangkan *RMSEA* mengukur penyimpangan nilai parameter pada suatu model dengan matriks kovarians populasinya. Dari gambar model terlihat bahwa: (1) nilai *RMSEA* = 0,000; (2) *GFI* = 0,892; (3) *RMSEA* = 0,000. Dengan demikian, ketiga unsur tersebut terpenuhi sehingga model tersebut dikategorikan sebagai model yang *fit*.

Gambar 1 merupakan model empiris berdasarkan hasil analisis SEM.

#### Model Pengukuran Kinerja Guru

Uji model pengukuran kinerja sebagai variabel laten dan masing-masing aspek kinerja yaitu: pedagogi, kepribadian, sosial, dan profesional sebagai variabel tampak. Oleh karena itu, skor variabel tampak dalam model pengukuran kinerja guru Akidah dan Akhlak merupakan skor komposit dari model pengukuran masing-masing aspek kinerja guru. Untuk menguji kesesuaian model hipotetik dengan data empirik didasarkan pada empat indikator, yaitu: (1) semua variabel tampak memiliki nilai muatan faktor ( $\lambda$ ) > 0,3; (2) *Chi-square* = 284,31; (3) *GFI* = 0,892; (4) *RMSEA* = 0,000; (5) *P-value* = 0,98301.

Berdasarkan hasil analisis dengan CFA (*Confirmatory Factor Analysis*) dengan Lisrel terhadap instrumen Kinerja Guru Akidah dan Akhlak diperoleh hasil semua variabel tampak memiliki nilai muatan faktor ( $\lambda$ ) > 0,30. Dengan demikian, model pengukuran tersebut merupakan model yang cocok untuk mengumpulkan data tentang Kinerja guru Akidah dan Akhlak.



Gambar 1. Model Empiris Evaluasi Model Logik Pembelajaran Akidah dan Akhlak (MLEPAA)

#### Model Pengukuran Materi dan Metode

Uji model pengukuran materi & metode sebagai variabel laten, dan masing-masing terdiri dari variabel tampak yaitu: terprogram, menyeluruh, modeling, inkul-kasi, fasilitasi, dan pengembangan keterampilan akademik dan sosial. Oleh karena itu, skor variabel tampak dalam model peng-ukuran Materi dan Metode Pembelajaran Akidah dan Akhlak merupakan skor kom-posit dari model pengukuran masing-masing aspek materi dan metode.

Ada empat indikator yang dapat di-jadikan pedoman untuk menguji kesesuaian model hipotetik dengan data empirik: (1) semua variabel tampak memiliki nilai muatan faktor ( $\lambda$ ) > 0,3; (2) *Chi-square* = 23,98; (3) *GFI* = 0,969; (4) *RMSEA* = 0,081, (5) *P-value* = 0,00434. Berdasarkan hasil analisis dengan Lisrel terhadap instrumen Materi dan Metode Pembelajaran Akidah dan Akhlak diperoleh nilai muatan faktor ( $\lambda$ ) > 0,30, yaitu: terprogram= 0,72, menyeluruh= 0,72, *modeling*= 0,78, inkulkasi= 0,72, fa-silitasi= 0,75, pengembangan keterampilan akademik dan sosial= 0,72. Dengan dem-ikian, model pengukuran tersebut merupa-

kan model yang cocok untuk mengumpul-kan data tentang Materi dan Metode Pembelajaran Akidah dan Akhlak.

#### Model Pengukuran Sarana Pembelajaran

Uji model pengukuran sarana pem-belajaran Akidah dan Akhlak sebagai varia-bel laten, dan masing-masing terdiri dari variabel tampak yaitu: kondisi ruang kelas, ketersediaan meja-kursi, media pembelaj-aran, ketersediaan buku pembelajaran. Oleh karena itu, skor variabel tampak dalam mo-del pengukuran ini merupakan skor kompo-sit. Untuk menguji kesesuaian model den-gan data empirik, ada empat indikator yang dapat dijadikan pedoman, yaitu: (1) semua variabel tampak memiliki nilai muatan fak-tor ( $\lambda$ ) > 0,3; (2) *Chi-square* = 1,93; (3) *GFI* = 0,996; (4) *RMSEA* = 0,00; (5) *P-value* = 0,381. Dengan demikian, model pengukur-an tersebut merupakan model yang cocok untuk mengumpulkan data tentang Kinerja guru Akidah dan Akhlak. Dengan demikian, model pengukuran tersebut merupakan mo-del yang cocok untuk mengumpulkan data tentang sarana Pembelajaran

## Model Pengukuran Kultur Madrasah

Uji model pengukuran Kultur Madrasah sebagai variabel laten dan masing-masing terdiri dari variabel tampak yaitu: kultur perilaku, artifak, dan pesan verbal. Oleh karena itu, skor variabel tampak dalam model pengukuran Kultur Madrasah merupakan skor komposit dari model pengukuran masing-masing Kultur Madrasah.

### Parameter Gamma ( $\gamma$ )

Parameter gamma merupakan pengaruh langsung variabel eksogen terhadap variabel endogenus. Berdasarkan *output* diketahui bahwa nilai *standardized* estimasi parameter gama satu = 0,228 sebagai efek langsung kinerja guru Akidah dan Akhlak terhadap pembelajaran Akidah dan Akhlak, gama dua = 0,232 sebagai efek langsung materi dan metode pembelajaran Akidah dan Akhlak terhadap pembelajaran Akidah dan Akhlak, gama tiga = 0,349 sebagai efek langsung sarana terhadap pembelajaran, dan gama empat = 0,216 sebagai efek langsung kultur terhadap pembelajaran.

### Parameter Beta ( $\beta$ )

Parameter beta menjelaskan besarnya pengaruh antara sesama variabel endogenus. Bagian kolom adalah variabel endogen independen, dan bagian baris adalah variabel endogen dependen. Dari *output* tersebut diketahui nilai *standardized* pengaruh antara variabel pembelajaran Akidah dan Akhlak terhadap perubahan motivasi sebesar 0,388, pengaruh antara variabel pembelajaran Akidah Akhlak terhadap perilaku berakidah akhlak sebesar 0,039, dan pengaruh perubahan motivasi dalam mengikuti pembelajaran Akidah Akhlak terhadap perubahan perilaku berakidah akhlak siswa sebesar 0,335. Pengaruh pembelajaran terhadap perilaku berakidah dan akhlak termasuk tidak signifikan, tetapi pengaruh pembelajaran terhadap perilaku berakidah dan akhlak melalui perubahan motivasi adalah signifikan.

### Parameter Zeta ( $\zeta$ )

Parameter zeta merupakan estimasi kesalahan pengukuran yang distandarisasi

pada variabel laten endogen. Secara berturut-turut *error variance* pada variabel endogen tersebut adalah sebagai berikut: pembelajaran Akidah dan Akhlak memiliki *error variance* sebesar 0,495, variabel perubahan motivasi 0,850, dan variabel perubahan perilaku berakidah akhlak sebesar 0,876.

### Total & Indirect Effect

Koefisien pengaruh kinerja guru terhadap pembelajaran Akidah dan Akhlak sebesar 0,224, sedangkan nilai errornya sebesar 0,060. Jika nilai 0,224 dibagi dengan error-nya (0,060) maka diperoleh nilai t-hitung sebesar 3,732. Dengan demikian, terdapat pengaruh positif signifikan antara kinerja guru Akidah dan Akhlak terhadap pembelajaran Akidah dan Akhlak, karena nilai t yang diperoleh (3,732) jauh lebih besar dibandingkan dengan nilai t-tabel (1,960) pada taraf signifikansi 5 % ( $3,732 > 1,960$ ). Berturut-turut nilai t-hit. yang diperoleh untuk: materi dan metode (3,975), sarana pembelajaran (5,609), serta kultur kelas dan madrasah adalah dan sebesar (3,927), dan semuanya berpengaruh secara positif signifikan terhadap pembelajaran Akidah dan Akhlak, karena t-yang diperoleh lebih besar nilainya dari t-tabel (1,960).

Variabel kinerja guru Akidah dan Akhlak, materi dan metode pembelajaran, sarana dan prasarana, serta kultur kelas dan madrasah berpengaruh secara total terhadap perubahan motivasi siswa secara positif signifikan, karena koefisien nilai t-yang diperoleh lebih besar dari nilai t-tabelnya (3,201; 3,350; 4,166; 3,321  $>$  1,960). Selanjutnya, kinerja guru Akidah dan Akhlak, materi dan metode pembelajaran, sarana dan prasarana, serta kultur kelas dan madrasah berpengaruh secara total terhadap perubahan perilaku berakidah dan berakhlak siswa dalam lingkup madrasah secara positif signifikan, karena koefisien nilai t-yang diperoleh lebih besar dari nilai t-tabelnya (2,076; 2,115; 2,281; 2,107  $>$  1,960).

Variabel kinerja guru Akidah dan Akhlak, materi dan metode pembelajaran, sarana dan prasarana, serta kultur kelas dan madrasah berpengaruh secara tidak lang-

sung terhadap perubahan motivasi siswa secara positif signifikan, karena koefisien nilai t-yang diperoleh lebih besar dari nilai t-tabelnya (3,201; 3,350; 4,166; 3,321 > 1,960). Demikian halnya dengan variabel kinerja guru Akidah dan Akhlak, materi dan metode pembelajaran, sarana dan prasarana, serta kultur kelas dan madrasah berpengaruh secara tidak langsung terhadap perubahan perilaku berakidah akhlak siswa dalam lingkungan madrasah secara positif signifikan, karena koefisien nilai t-yang diperoleh lebih besar dari nilai t-tabelnya (2,076; 2,115; 3,281; 2,107 > 1,960).

Koefisien pengaruh total variabel pembelajaran Akidah dan Akhlak terhadap variabel perubahan motivasi sebesar 0,388, dan terhadap perubahan perilaku berakidah akhlak sebesar 0,169. Koefisien pengaruh total variabel perubahan motivasi terhadap perubahan perilaku berakidah akhlak siswa dalam lingkup madrasah sebesar 0,335. Keseluruhan pengaruh total pembelajaran Akidah dan Akhlak terhadap perubahan motivasi dan perubahan perilaku berakidah akhlak adalah positif signifikan, karena harga t-hitungnya lebih besar dari harga t-tabelnya (5,852; 2,471) > 1,960. Demikian halnya dengan pengaruh total perubahan motivasi terhadap perubahan perilaku berakidah akhlak adalah positif signifikan, karena koefisien t-hitungnya lebih besar dari koefisien t-tabelnya (4,461 > 1,960)

Koefisien pengaruh tidak langsung variabel pembelajaran Akidah dan Akhlak terhadap perubahan perilaku berakidah akhlak siswa MTsN sebesar 0,135, dan pengaruh ini positif signifikan, hal ini ditunjukkan dengan koefisien t-hitung (3,609) lebih besar dari t-tabelnya (1,960).

Deskripsi hasil penelitian untuk menjawab pertanyaan nomor tiga, yaitu bagaimanakah kelayakan evaluasi model logik pembelajaran Akidah dan Akhlak di MTsN di Kabupaten Banyumas dapat dijelaskan sebagai berikut. *Pertama*, hasil penilaian *user* terhadap panduan model diperoleh rerata skor total 3,8, jika nilai ini dikonversikan ke dalam pedoman penilaian, maka masuk dalam kategori “baik” sehingga layak diguna-

kan. *Kedua*, penilaian user terhadap keefektifan model diperoleh rerata skor total 3,6, jika nilai ini dikonversikan ke dalam pedoman penilaian, maka masuk dalam kategori “baik” sehingga layak digunakan dan dapat digunakan di lapangan.

## Simpulan

Berdasarkan pada analisis data yang telah dipaparkan, maka dapat disimpulkan hasil penelitian sebagai berikut. Pertama, Model logik yang diperoleh telah diuji secara empirik sehingga didapatkan model yang *fit*, dilihat dari indikator-indikator berikut: (a) *Chi-Square*,  $p > 0,05$ , nilai yang diperoleh sebesar = 553,72, (b) *Goodness of Fit Indices* (GFI) harus berkisar antara 0 dan 1, nilai yang diperoleh = 0,892; (c) *Root Mean Square Error of Approximation* (RMSEA) < 0,05; nilai yang diperoleh = 0,000

Kedua, pengaruh variabel pembelajaran terhadap perubahan motivasi adalah signifikan, sedangkan pengaruh pembelajaran terhadap perubahan perilaku berakidah akhlak tidak signifikan. Namun, jika perubahan pembelajaran akidah akhlak melalui perubahan motivasi adalah signifikan. Ketiga, validitas dan reliabilitas instrumen pada komponen model logik telah memenuhi persyaratan yang ditentukan. Keempat, Model logik secara empirik telah dinyatakan layak dalam uji implementasi oleh *user*, yaitu guru Akidah dan Akhlak, kepala MTs, dan PPAI, dengan rerata skor total 3,6 > 3,4 - 4,2 sehingga dinyatakan baik dan layak digunakan.

Kelima, kelebihan model logik untuk mengevaluasi pembelajaran Akidah dan Akhlak kelas VIII MTsN yaitu: (a) dapat memberikan pemahaman terhadap rencana kerja/peta evaluasi dan *outcomes* yang diharapkan; (b) dapat memberikan panduan yang jelas tentang evaluasi pembelajaran Akidah Akhlak; (c) berfungsi sebagai alat pengawas evaluasi dan membantu mengidentifikasi pertanyaan kunci tentang apakah kunci/inti komponen evaluasi pembelajaran Akidah Akhlak telah dilaksanakan kelemahan/kekurangan jika diterapkan.

Keenam, kelemahan model logik antara lain: (a) kurang efektif jika waktu yang

tersedia untuk menyelesaikan penelitian terbatas; (b) untuk mengungkap proses pembelajaran dan kemajuan siswa yang dicapai dari hasil pembelajaran dibutuhkan instrumen yang komprehensif; (c) model logik ini banyak melibatkan siswa, jika siswa banyak diberi instrumen maka akan terjadi kebosanan dalam proses pengisian hal ini dapat berakibat pada pengisian data yang asal-asalan. Boleh jadi, data yang diperoleh tidak sesuai dengan yang diharapkan atau kenyataan yang sesungguhnya.

#### Saran

Berdasarkan pada simpulan yang telah disampaikan, maka dapat disampaikan beberapa saran sebagai berikut. Pertama, pembelajaran akidah akhlak harus melibatkan afeksi siswa, sehingga perubahan perilaku berakidah akhlak siswa dapat ditingkatkan. Kedua, sebelum digunakan untuk mengevaluasi pembelajaran, model logik ini harus dipahami oleh para pengguna, karena model digambarkan dalam bentuk *flow chart* dan memiliki sifat hubungan yang rasional dan terpadu.

Ketiga, untuk menghindari kebosanan dalam pengisian instrumen, maka instrumen yang diberikan kepada siswa diberikan sesuai dengan tahapan evaluasi yang sedang berjalan, dan waktu yang tepat. Keempat, penggunaan model ini harus disesuaikan dengan tujuan, kegunaan dan level program yang diinginkan. Kelima, model ini dapat digunakan oleh kepala madrasah, pengawas, guru MTsN/S. Agar proses evaluasi lebih efektif, maka model ini dapat dikembangkan lagi dengan pembuatan *software*, sehingga memudahkan evaluator dalam proses tabulasi, mengoreksi dan mengetahui hasil evaluasinya, pada komponen mana yang sudah baik, dan komponen mana yang belum baik. Dengan adanya *software* tersebut, bisa diketahui hasil secara cepat, tepat dan dapat dilakukan perbaikan pada bagian yang kurang. Pembuatan *software* yang solutif tersebut membutuhkan tahapan pengembangan lebih lanjut. Keenam, diseminasi sebaiknya dilaksanakan dalam skala yang lebih

luas, sehingga menghasilkan kualitas model dan perangkat yang lebih baik.

#### Daftar Pustaka

- Al-Abrasyi, M.A. (1993). *Dasar-dasar pokok pendidikan Islam*. (Terjemahan Bustami A.Gani & Johar Bahri). Jakarta: Bulan Bintang.
- Assegaf, A. (2001). Memahami sumber konflik antariman. Dalam TH. Sumartana (Eds.), *Pluralisme, konflik dan pendidikan agama di Indonesia* Yogyakarta: Dian Interfidei.
- Azwar, S. (2005). *Sikap manusia: Teori dan pengukurannya*. Yogyakarta: Pustaka Pelajar.
- Barber, B. K., & Olsen, J. A. (2004). Assessing the transitions to middle and high school. *Journal of Adolescent Research*, 19, 3–30.
- Bickman, L. (Ed.). (1987). The functions of program theory. *New Directions in Program Evaluation: Using Program Theory in Evaluation*.
- Borg, W.R., & Gall, M.D. (1983). *Educational Research: An Introduction*. (4<sup>th</sup> ed.) New York: Longman.
- Cennamo, K., & Kalk, D. (2005). *Real world instructional design*. Canada: Thomson Learning, Inc.
- Chazan, B., & Soltis, J.F. (1975). *Moral Education*. New York: Teacher College Press.
- Eye, A. V. & Mun, E. Y. (2005). *Analyzing rater agreement (Manifest variable methods)*. Marwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Fadjar, A. M. (1998). *Madrasah dan tantangan modernitas*. Bandung: Mizan.
- Fernandez, H.J.X (1984). *Testing and measurement*. Jakarta: National Educational and Curriculum Development.
- Ghozali, I. (2005). *Structural equation modeling: Teori, konsep dan aplikasinya dengan program Lisrel 8.80* (Edisi II). Semarang:

- Badan Penerbit Universitas Diponegoro.
- Griffin, P., & Nix, P. (1991). *Educational assessment and reporting*. Sydney: Harcourt Brace Javanovich, Publisher.
- Gould, M. S., Greenberg, T., Velting, D., & Shaffer, D. (2003). Youth suicide risk and preventive interventions: A review of the past 10 years. *Journal of the American Academy of Child and Adolescent Psychiatry*, 42, 386–405.
- Heneret, E., & Powel, E.T. (2008). *Developing a logic model: Teaching and training guide*. Madison: University of Wisconsin.
- Keller, J. M. (1979). Motivation and instructional design: A theoretical perspective. *Journal of Instructional Development*, 2 (4), 26-34.
- Keller, J. M. (1983). Motivation design of instruction. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status* (pp. 383-434). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Keller, J. M. (1987a). Development and use of the ARCS model of instructional design. *Journal of Instructional Development*, 10 (3), 2-10.
- Kellogg, W.K. Foundation. (2004). *Using logic models to bring planning, evaluation, and action*. Michigan: WK Kellogg Foundation. Diambil pada tanggal 10 Juli 2012, dari <http://www.wkkf.org>.
- Kerr, M., et.al. (2010). A reinterpretation of parental monitoring in longitudinal perspective. *Journal of Research on Adolescence*. 20 (1), 39–64, 2010.
- La Greca, A. M., & and Hannah Moore Harrison, H. M. (2005). Adolescent peer relations, friendships, and romantic relationships: Do they predict social anxiety and depression?. *Journal of Clinical Child and Adolescent Psychology*, 34 (1), 49–61.
- LeBlanc, L., et.al. (2008) High school social climate and antisocial behavior: A 10 year longitudinal and multilevel study. *Journal of Research on Adolescence*. 18 (3), 395–419.
- Mardapi, D. (2007). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendikia.
- Mastuhu. (1999). *Memberdayakan sistem pendidikan Islam*. Jakarta: Logos.
- McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your programs Performance Story. *Evaluation and Program Planning*, 22, 65-72.
- McMillan, J.H., & Schumacher, S. (2010). *Research in education (7<sup>th</sup> ed.)*. New York: Pearson.
- Miskawaih, I. (1999). *Menuju kesempurnaan akhlak (Cetakan V)*. (Terjemahan Helmi Hidayat). Libanon: Darul Kutub al-‘Ilmiyyah. (Buku asli diterbitkan tahun 1985).
- Newman, B. M., et.al. (2007). The relationship of social support to depressive symptoms during the transition to high school. *Journal of Adolescent Research*, 42, 441–459.
- Nunally, J.C. (1981). *Psychometric Theory (2<sup>nd</sup> ed)*. New Delhi: Tata McGraw-Hill Publishing Company Limited
- Peraturan Menteri Agama RI No. 2 tahun 2008 tentang Standar Kompetensi Lulusan dan Standar Isi. Diambil pada tanggal 12 Juli 2011, dari <http://www.scribd.com/doc/57115839/Peraturan-Menteri-Agama-Republik-Indonesia>.
- Porteous, N.L., et al. (2002). Introducing program team to logic models: Facilitating the learning process. *Canadian Journal of Program Evaluation*. Vol. 17, No. 3, 113-141.
- Powell, T., & Heneret, E. (2008). Enhancing program performance with logic models. Diakses tanggal 15 November 2011 dari Wisconsin Extension Website:

<http://www.UWEX/edu.ces/pdande/>

- Rahman, F. (1985). Approach to Islam in religious studies review essays. Dalam Richard C. Martin (Ed.). *Approaches to Islam in religious studies*. Tuscon: The University of Arizona.
- Schunk, D.H., Pintrich, P.R., & Meece, J.L. (2010). *Motivation in education: Theory, research, and application (3<sup>rd</sup> ed.)*. New Jersey: Pearson Education, Inc., Upper Saddle River.
- Slavin, R. E. (2006). *Educational psychology theory and practice (8<sup>th</sup> ed.)*. Boston: Pearson Education, Inc.

Sudijono, A. (2003). *Pengantar evaluasi pendidikan*. Jakarta: PT Raja Grafindo Persada.

Zuchdi, D. (2010). *Humanisasi pendidikan: Menemukan kembali pendidikan yang manusiawi*. Jakarta: Bumi Aksara.

Zuchdi, D. et al. (2011). *Laporan Penelitian Hibah Penelitian Tim Pascasarjana-HPTP (Hibah Pasca)*. Pengembangan model pendidikan karakter dengan pendekatan komprehensif terpadu dalam pembelajaran bahasa indonesia, IPA, dan IPS di Sekolah Dasar. Yogyakarta: Universitas Negeri Yogyakarta.

## EVALUASI DAN STRATEGI PENGEMBANGAN SMA INDONESISCH NERDERLANDSCHE SCHOOL (INS) KAYUTANAM

*David, Badrun Kartowagiran, Slamet PH*

IAIN Batusangkar, Universitas Negeri Yogyakarta, Universitas Negeri Yogyakarta  
davidazis@gmail.com, badrunkw@yahoo.com, slametph@yahoo.com

### Abstrak

Penelitian ini bertujuan untuk: (1) mengevaluasi SMA INS Kayutanam dan (2) menyusun strategi pengembangan SMA INS Kayutanam berdasarkan hasil evaluasi tersebut. Pendekatan yang digunakan dalam penelitian ini adalah kuantitatif dan kualitatif. Sumber data dalam penelitian ini adalah 50 siswa kelas X dan XI, satu kepala sekolah, empat wakil kepala sekolah, 30 guru, dua pembina asrama dan dua orang dari pemerintahan. Hasil penelitian dibagi dalam dua kelompok yaitu hasil evaluasi dan strategi pengembangan. Berdasarkan hasil evaluasi ditemukan bahwa ada beberapa komponen yang perlu perbaikan yaitu kurikulum, kompetensi pendidik dan tenaga kependidikan, fasilitas, peran serta masyarakat, dana dan organisasi. Strategi pengembangan terdiri dari: (1) perumusan kurikulum terintegrasi yang memuat tujuan dan sasaran pendidikan yang diprakarsai oleh Sjafei; (2) perumusan metode mengajar yang sesuai dengan tujuan dan sasaran pendidikan yang diprakarsai Sjafei; (3) pengelolaan aset produktif dengan pemanfaatan modal sosial; (4) pemanfaatan basis alumni; (5) membangun kerja sama dengan instansi terkait seperti sekolah lain, pemerintah daerah, dan pihak swasta.

**Kata kunci:** *evaluasi, strategi pengembangan*

## THE EVALUATION AND THE STRATEGY FOR DEVELOPING INDONESISCH NEDERLANDSCHE SCHOOL (INS) KAYUTANAM SENIOR HIGH SCHOOL

*David, Badrun Kartowagiran, Slamet P.H.*

Batusangkar State Islamic Academy, Yogyakarta State University, Yogyakarta State University  
davidazis@gmail.com, badrunkw@yahoo.com, slametph@yahoo.com

### Abstract

The study was to: (1) evaluate the development of INS Kayutanam Senior High School; and (2) design the strategy to develop INS Kayutanam Senior High School based on the results of the evaluation. The approach used of in the study was quantitative and qualitative. The data source in the study was 50 students from the X grade and the XI grade, one principal, four vice principals, 30 teachers, two dormitory advisors and two government representatives. The results of the study were divided into two groups namely the evaluation results and the development strategies. Based on the results of the study, the researcher found that there were several components that demanded revision namely the curriculum, the educator and educational staff competence, the facility, the society participation, the fund and the organization. Then, the development strategies consisted of: (1) formulating an integrated curriculum that contained the educational objectives and targets that had been provided by Sjafei; (2) formulating the teaching methods that had to be in accordance with the educational objectives and targets that had been provided by Sjafei; (3) managing the productive assets by utilizing the social capital; (4) utilizing the alumni basis; and (5) establishing partnership with the related institutions such as other schools, regional governments and private sectors.

**Keywords:** *evaluation, development strategies*

## **Pendahuluan**

Evaluasi merupakan suatu rangkaian kegiatan dalam meningkatkan kualitas kinerja, atau produktivitas suatu lembaga dalam melaksanakan program (Mardapi, 2008, p.8). Kemudian Kartowagiran (2013, p.19) mengemukakan bahwa evaluasi merupakan kegiatan yang sistematis untuk mengidentifikasi, mengklarifikasi dan mengaplikasikan suatu kriteria untuk menentukan keberhasilan suatu program. Rossi (1999, p.4) mendefinisikan evaluasi mencakup semua upaya untuk menempatkan nilai pada peristiwa, proses atau orang-orang. Sejalan dengan yang dikemukakan sebelumnya Mertens (2004, pp.17-18) menyatakan bahwa evaluasi adalah penyelidikan sistematis dan objektif tentang kelayakan nilai dari suatu objek dengan tujuan mengurangi ketidakpastian dalam pengambilan keputusan. Berdasarkan beberapa pendapat yang dikemukakan, diketahui bahwa evaluasi merupakan penilaian terhadap suatu persoalan yang umumnya menunjuk baik buruknya persoalan tersebut atau tercapai atau tidaknya tujuan dari suatu program.

Pelaksanaan evaluasi pada suatu lembaga pendidikan bisa untuk mengetahui komponen-komponen yang sudah berjalan menurut semestinya atau yang belum berjalan, sehingga dengan pelaksanaan evaluasi ini diketahui komponen-komponen yang sudah baik dan yang masih perlu diperbaiki. Apabila sekolah mengetahui kondisi real saat ini, sekolah bisa melakukan strategi pengembangan ke depan. Slamet PH (2013, p.2) menjelaskan bahwa rencana pengembangan sekolah penting dimiliki untuk memberikan arah dan bimbingan para pelaku sekolah dalam rangka menuju perubahan atau tujuan sekolah yang lebih baik, dengan resiko yang kecil dan untuk mengurangi ketidakpastian masa depan. Lebih lanjut, Slamet PH (2013, p.3) mengemukakan bahwa arti perencanaan adalah suatu proses untuk menentukan tindakan masa depan sekolah yang tepat, melalui urutan pilihan, dengan mempertimbangkan sumberdaya yang tersedia. Kemudian rencana pengembangan sekolah adalah dokumen tentang

gambaran kegiatan sekolah di masa depan dalam rangka untuk mencapai perubahan/tujuan sekolah yang telah ditetapkan. Hasil evaluasi terhadap suatu lembaga bisa dijadikan pertimbangan dalam menentukan perencanaan suatu sekolah, termasuk untuk menentukan strategi pengembangan Ruang Pendidik INS Kayutanam.

Ruang Pendidik INS Kayutanam didirikan pada tanggal 31 Oktober 1926 oleh Engku M. Sjafei di Kabupaten Padang-pariaman Sumatera Barat. Kata ruang di sini bukan bermakna ruangan, tetapi bermakna suatu tempat yang luasnya tiada terbatas, sedangkan pendidik artinya belajar dan mengajar. Jadi Ruang Pendidik adalah satu tempat luas yang digunakan untuk belajar mengajar, bukan hanya terbatas pada adanya guru dan siswa, tetapi belajar dari pengalaman dan kehadiran alam sekitar (Navis, 1996, p.60), yang dalam falsafah Minangkabau dikenal dengan alam takambang jadikan guru. Penggunaan filosofi Minangkabau dalam pendidikan oleh Sjafei juga sesuai dengan yang dikemukakan oleh John Dewey bahwa pengalaman manusia adalah hubungan antara manusia dengan masyarakat serta alam (Tilaar, 2008, p.105). Lingkungan sebagai tempat belajar ini juga sejalan dengan Firman Allah swt, di antaranya terdapat dalam surat Az-Zumar ayat 21 dan surat An-Nahl ayat 66.

Ruang Pendidik INS Kayutanam didirikan Sjafei untuk mendidik dan menanamkan tradisi semangat kerja dan kemandirian. Penanaman sifat kemandirian dituangkan dalam tujuan sekolah ini yaitu: (1) mendidik anak-anak kearah hidup yang merdeka melalui hidup mandiri; (2) menanamkan kepercayaan kepada diri sendiri, membina kemauan keras, dan membiasakan berani bertanggung jawab; (3) membiayai diri sendiri dengan semboyan cari sendiri dan kerjakan sendiri; (4) mengembangkan anak secara harmonis, yang mencakup aspek perasaan, kecerdasan, dan keterampilan; (5) mengembangkan sikap sosial, agar dapat hidup bermasyarakat dengan baik; (6) menyesuaikan pendidikan sesuai dengan bakat anak. Relevansi konsep pendidikan Sjafei terletak pada

mengubah peserta diri dari budaya santai kepada tradisi etos kerja, semangat juang untuk menegakkan diri sendiri secara mandiri (INS, t.th, p.5).

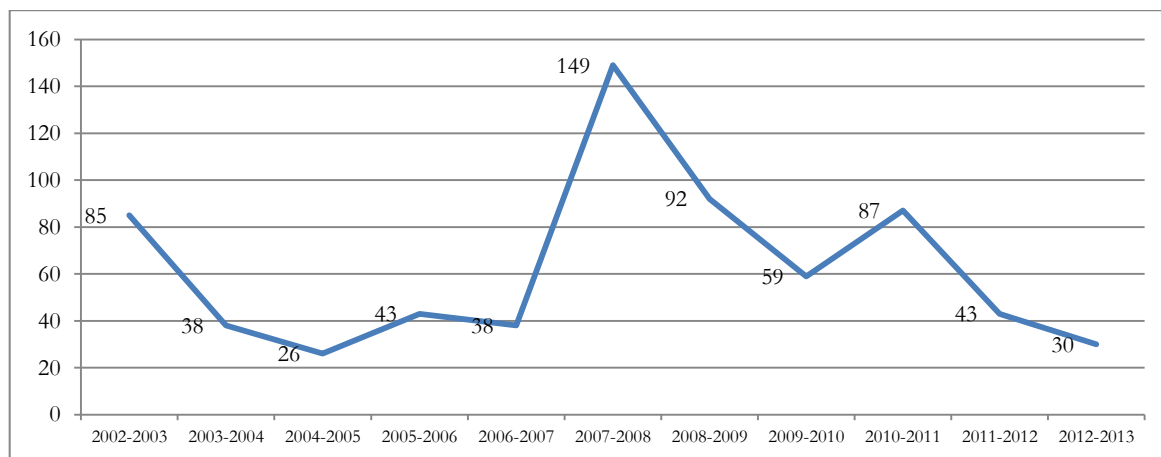
Rancangan pendidikan Sjafei menurut Said dan D. Mansoer (Raharja, 2008, p.13) mendekati rancangan John Dewey di Amerika Serikat dan Kerschensteiner di Jerman. Raharja (2008, p.17) menambahkan bahwa “dasar pendidikan di Ruang Pendidik INS Kayutanam menggambarkan bahwa pendidikan yang diselenggarakannya dapat dikatakan sebagai pendidikan yang humanis dan religius, baik secara eksplisit maupun implisit”. Program yang ada di Ruang Pendidik INS Kayutanam adalah mengembangkan komponen penting peserta didik yaitu otak (program akademik), hati (program akhlak mulia), dan tangan (program keterampilan) agar seimbang dan sesuai dengan kodratnya.

Pada awal berdiri sekolah ini tidak mengenal jenjang pendidikan seperti SMP ataupun SMA, tetapi terdiri dari Ruang Rendah terdiri dari 7 tahun setara sekolah dasar dan Ruang Dewasa terdiri dari 4 tahun setara dengan sekolah menengah (Tirtahardja, 2012, p.210). Seiring dengan perjalanannya sekolah ini mengikuti kebijakan yang ada yaitu harus menentukan pilihan SMP atau SMA. Akhirnya, lembaga ini dibagi menjadi dua yang terdiri dari SMP Plus dan SMA Plus. Namun, seiring dengan perjalanan waktu peminat siswa yang masuk SMP menjadi berkurang sehingga yang ada saat ini hanya siswa SMA saja. Meskipun sudah menjadi SMA, namun keunikannya masih dipertahankan.

Di antara keunikan SMA INS Kayutanam ini adalah sebagai berikut. Pertama, SMA yang memiliki program keterampilan (kriya meubel, keramik, klose anyam, teknik las, elektronik), program pembinaan rohani/jiwa melalui pendidikan agama (teori dan aplikasi), seni (lukis, musik, silat, tari, dan teater), dan olahraga (sepak bola, bulu tangkis, voli, basket, dan takraw), dan pendidikan otak melalui program akademis (kurikulum 2013) serta didukung oleh pendidikan asrama. Kedua, sasaran pendidikan di SMA INS Kayutanam juga berbeda dengan

SMA lain, yaitu: (a) ilmu eksak/pasti/alam, sasarannya agar siswa berpikir logis, matematis dan dialektis. Sasaran pendidikan Ilmu Sosial adalah untuk menguatkan ingatan dan memperluas wawasan siswa. Mata pelajaran Budaya dan Bahasa memiliki sasaran untuk memahirkan komunikasi, etika dan meluaskan wawasan siswa. Mata Pelajaran PPKn/Pancasila menanamkan kesadaran dan kepedulian nasional; (b) sasaran pendidikan keterampilan membangun etos kerja, rajin, dan terampil. Hemat, cerdas dan produktif. Berwatak pengusaha dan mandiri; (c) sasaran pendidikan agama adalah takwa dan berakhlak mulia sebagai khalifah Tuhan di muka bumi serta toleran pada beda pendapat; (d) sasaran pendidikan seni adalah aktif, kreatif dan estetis, berjiwa peka dan kritis, serta komunikatif. Toleran pada beda pendapat; (e) Pendidikan Olahraga Jasmani sehat dan kuat. Berani dan percaya diri. Tabah ulet, disiplin dan koperatif serta sportif. Program di sekolah ini di dukung oleh beberapa fasilitas pembelajaran dan pelatihan diantaranya gedung belajar, laboratorium, bengkel, perkantoran, aula, asrama, tempat ibadah, sarana dan prasarana olahraga, lahan pertanian, peternakan, perkebunan, perikanan dan masih ada fasilitas yang lain.

Apabila dilihat dari jumlah siswa yang mendaftar ke SMA INS Kayutanam dari tahun 2002-2013 menimbulkan kegelisahan dan kekhawatiran. Berikut gambaran siswa yang mendaftar di SMA INS Kayutanam. Berdasarkan Gambar 1 terlihat SMA INS Kayutanam kekurangan siswa dan dikhawatirkan tahun demi tahun akan berkurang. Sementara jumlah seluruh siswa dari kelas X sampai kelas XII yang terdaftar saat ini 152 orang. Kalau INS Kayutanam mengalami penurunan yang rugi bukan masyarakat Kayutanam atau masyarakat Sumatera Barat saja, akan tetapi yang rugi adalah bangsa Indonesia. Agar bangsa Indonesia tidak kehilangan aset yang besar yaitu salah satu pilar dari dua pilar pemikiran pendidikan di Indonesia yaitu INS Kayutanam dan Tamansiswa, sehingga INS Kayutanam perlu dievaluasi yang menyeluruh atau kompre-



Gambar 1. Perkembangan Jumlah siswa yang Mendaftar di SMA INS Kayutanam dari tahun 2002-2013

hensif. Hal tersebut perlu, agar diketahui komponen-komponen mana yang belum berjalan menurut semestinya agar bisa ditemukan strategi pengembangan SMA INS Kayutanam ke depan.

### Metode Penelitian

Metode penelitian yang digunakan untuk evaluasi SMA INS Kayutanam adalah kuantitatif, kemudian untuk strategi pengembangan menggunakan metode kualitatif. Pengumpulan data kuantitatif dan kualitatif dilakukan secara bersama-sama termasuk dalam analisis.

Penelitian dilaksanakan pada bulan Mei sampai dengan Agustus 2014 di SMA INS Kayutanam Kabupaten Padang Pariaman Provinsi Sumatera Barat. Lokasi sekolah ini berada di Jl. Raya Padang-Bukittinggi KM. 53 Kayutanam, Padang Pariman 25585. Sumber data dalam penelitian ini 50 orang siswa kelas X dan XI, satu orang kepala sekolah, empat orang wakil kepala sekolah, 30 orang guru, dua orang pembina asrama dan dua orang dari pemerintahan yaitu Sekretaris Camat Kecamatan 2x11 Kayutanam dan Kepala Dinas Pendapatan, Pengelolaan Keuangan dan Aset Kabupaten Padang Pariaman. Penelitian ini secara garis besar terdiri dari tiga tahap yaitu tahap perencanaan, tahap evaluasi, dan tahap strategi pengembangan. Tahap perencanaan terdiri dari di-

agnosis permasalahan, perencanaan instrumen. Tahap kedua, pelaksanaan evaluasi, analisis data dan penyajian hasil evaluasi kepada pemangku kepentingan utama melalui FGD\_1. Tahap ketiga terdiri dari, FGD\_1 (melahirkan draf strategi), telaah oleh promotor/ahli, FGD\_2 (untuk validasi produk), dan produk penelitian. Instrumen penelitian terdiri dari kuesioner, pedoman observasi, pedoman wawancara. Data kuantitatif dianalisis dengan statistik deskriptif. Validasi produk penelitian didasarkan kepada telaah 7 orang pakar kemudian dianalisis dengan Aiken. Setelah produk divalidasi oleh pakar dilakukan uji keterimaan produk oleh praktisi yang terdiri dari kepala sekolah, wakil kepala sekolah dan guru SMA INS Kayutanam. Data dari kepala sekolah, wakil kepala sekolah, dan guru tersebut dianalisis menggunakan Aiken.

### Hasil Penelitian dan Pembahasan

Berdasarkan hasil evaluasi ditemukan beberapa kekurangan dan kekuatan yang ada di SMA INS Kayutanam. Secara rinci akan diuraikan sebagai berikut.

#### Kekurangan

Kekurangan yang dikemukakan ini berdasarkan hasil penilaian siswa, guru, pimpinan SMA INS Kayutanam dan juga berdasarkan pengamatan langsung, diantara-

nya adalah sebagai berikut. Pertama, kekurangan pada bidang kurikulum adalah kelengkapan pembelajaran untuk paket keterampilan yang terdiri dari silabus, RPP, dan bahan ajar. Kedua, kekurangan pada tenaga pendidik. Rasio guru dan siswa perlu direvisi karena kurang efisien sebagai contoh saat ini rasio guru dan siswa untuk praktek keterampilan 1:2, 1:3, sampai 1:5. Rasio ini bukan bukan karena standar yang ditentukan tetapi karena jumlah siswa yang kurang.

Ketiga, kekurangan pada fasilitas SMA INS Kayutanam adalah sebagai berikut: (1) ragam keterampilan di SMA INS Kayutanam terlalu banyak sehingga peralatan jadi kurang mendukung; (2) fasilitas yang kurang layak di SMA INS Kayutanam yang terdiri dari ruang teknik mesin, ruangan otomotif, ruang teknik elektronika, ruang desain grafis, ruang sanggar tari, ruang teater, dan sarana pendukung asrama; (3) peralatan pendukung yang termasuk kurang dari segi kecukupan terdiri dari peralatan kriya ukir, peralatan kriya keramik, peralatan desain grafis, peralatan sanggar tari, peralatan sanggar teater, peralatan musikalisasi puisi, peralatan pertanian, peralatan perkebunan, peralatan peternakan, peralatan pendukung kegiatan asrama, peralatan otomotif, dan peralatan elektronika; (4) peralatan latihan yang kurang relevan dengan kebutuhan yang terdiri dari peralatan teknik elektronika, peralatan kriya ukir, peralatan kriya keramik, peralatan kriya anyam, peralatan desain grafis, peralatan sanggar tari, peralatan sanggar teater, peralatan sanggar musikalisasi puisi, peralatan pertanian, peralatan perkebunan, peralatan peternakan, peralatan pendidikan asrama, dan peralatan otomotif; (5) peralatan latihan yang kurang mutakhir terdiri dari kriya anyam, sanggar teater, sanggar musikalisasi puisi, peralatan pertanian, peralatan kriya ukir, peralatan kriya keramik, peralatan desain grafis, peralatan perkebunan, dan peralatan peternakan.

Keempat, pelibatan masyarakat dalam pendidikan yang termasuk kurang adalah sebagai berikut: (1) peranan komite sekolah belum berfungsi dengan baik; (2) belum ada

kerja sama antara SMA INS Kayutanam dengan lembaga lain seperti pondok pesantren, SMP/MTs. Kerja sama baru sebatas untuk sosialisasi melalui penempelan brosur di sekolah-sekolah tersebut; (3) pemberdayaan alumni untuk kemajuan SMA INS Kayutanam belum berjalan dengan baik, selain itu juga belum memiliki data alumni yang lengkap. Kekurangan dari segi pembiayaan adalah sebagai berikut: (1) biaya peningkatan kompetensi guru; (2) biaya operasional untuk praktek keterampilan; (3) biaya operasional untuk sanggar; (4) kemampuan sekolah membayar gaji pendidik dan tenaga kependidikan; (5) kemampuan sekolah membayar transportasi pendidik dan tenaga kependidikan; (6) kemampuan sekolah membayar tunjangan pendidik dan tenaga kependidikan.

Kelima, kekurangan dari segi pengelolaan adalah sebagai berikut: (1) keikutsertaan warga sekolah merumuskan visi dan misi sekolah; (2) keikutsertaan warga sekolah dalam menyusun rencana menengah, keikutsertaan masyarakat dalam pembiayaan masih terbatas; (3) belum membangun kemitraan dengan lembaga relevan. Kekurangan dalam proses pendidikan di kelas, bengkel, dan asrama adalah sebagai berikut: (1) Guru belum melaksanakan pretes sebelum pembelajaran dimulai; (2) pembelajaran yang kurang menantang; (3) alat peraga yang ada di ruangan kelas kurang. (4) alat peraga kurang sesuai dengan materi. (5) alat peraga yang diambilkan dari alam masih kurang. (6) Guru tidak membuat rangkuman sebelum pembelajaran ditutup; (7) Guru hampir tidak pernah memberikan tugas yang dikerjakan diasrama. Hal tersebut berdasarkan informasi dari siswa karena hampir tidak ada waktu untuk mengerjakan tugas karena kegiatan sangat padat. (8) proses pelatihan teknik kayu kurang sesuai dengan kompetensi yang ditetapkan; (9) proses pelatihan elektronika kurang sesuai dengan kompetensi yang ditetapkan; (10) pelatihan siswa untuk selalu bermusyawarah sebelum melakukan kegiatan masih kurang. (11) ketercapaian peningkatan keterampilan siswa melalui pelatihan keterampilan masih kurang; (12) keter-

capaian peningkatan kinerja siswa masih kurang. (13) ketercapaian pembentukan watak produktif masih kurang; (14) ketercapaian pembentukan jiwa kewirausahaan masih kurang; (14) pelaksanaan salat ashar berjamaah belum berjalan menurut semestinya.

Kekurangan yang terdapat dalam proses penilaian adalah sebagai berikut: (1) penentuan nilai akhir kelompok mata pelajaran oleh sekolah tidak berjalan menurut semestinya; (2) penentuan KKM tidak melibatkan melibatkan orang tua siswa. Apabila dilihat hasil ujian nasional selama tiga tahun berturut-turut kelihatan bahwa 100% siswa lulus Ujian Nasional, namun apabila dilihat hasil Ujian Nasional sebelum masuk nilai sekolah tingkatketidak lulusan cukup tinggi yaitu antara 11,11%-77,27%.

#### Kekuatan

Kekuatan/potensi yang dimiliki oleh SMA INS Kayutanam antara lain memiliki akreditasi A, memiliki silabus, RPP dan bahan ajar yang lengkap. Pendidik dan tenaga kependidikan; (1) latar belakang pendidikan guru sudah sesuai dengan mata pelajaran yang diajarkan; (2) umumnya tenaga pendidik memiliki kualifikasi pendidikan S1, sangat sedikit sekali yang memiliki kualifikasi DIII; (3) sebanyak 34,29% guru telah memiliki sertifikat sertifikasi pendidik; (4) unsur pimpinan SMA INS Kayutanam memiliki semangat perubahan ke arah kemajuan serta ditopang oleh kinerja yang tinggi; (5) memiliki struktur badan wakaf/yayasan yang terdiri dari orang-orang penting pada pemerintahan, pendidikan, politik, swasta pada tingkat daerah sampai nasional.

Fasilitas yang dimiliki SMA INS Kayutanam adalah sebagai berikut: (1) memiliki lahan yang luas yaitu 18 hektar dan lokasi yang strategis; (2) memiliki gedung belajar, aula, masjid, asrama, dan sarana olahraga, serta ruang kantor yang memadai; (3) memiliki lahan perkebunan, pertanian, dan peternakan, serta memiliki kolam ikan yang luas; (4) meskipun ruang kurang memadai namun peralatan bengkel teknik kayu, teknik mesin, dan musik sudah memadai dari segi jumlah; (5) demikian juga dari segi tem-

pat/ruang kurang memadai namun peralatan teknik kayu dan teknik mesin sudah mutakhir; (6) memiliki laboratorium kimia yang lengkap. Mendapat dukungan yang baik dari pemerintah daerah. Memiliki alumni yang telah sukses diberbagai bidang seperti pemerintahan, pendidikan, politik, dan swasta (pengusaha) pada tingkat daerah sampai nasional. Memiliki visi dan misi yang baik, struktur organisasi yang lengkap. Memiliki sistem pendidikan yang menyeimbangkan antara kemampuan akademik (intelektual), tangan (keterampilan hidup), dan hati (akhlak mulia). Memiliki program akhlak mulia, aqidah, dan ibadah. Memiliki program yang melatih kedisiplinan, kejujuran, kesetiakawanan, kepemimpinan, dan pengabdian masyarakat. Berdasarkan hasil Ujian Nasional untuk sekolah swasta di Kabupaten Padang Pariaman memiliki peringkat 1 sampai 2 sepanjang tahun 2010-2013. Apabila digabungkan negeri dan swasta peringkatnya antara 3 sampai 15. Memiliki prestasi yang baik dibidang keterampilan yaitu dengan melihat hasil karya siswa dibengkel dan sanggar serta prestasi perlombaan pada tingkat kabupaten dan provinsi.

Hasil evaluasi yang menggambarkan kekurangan dan kelebihan SMA INS Kayutanam dipresentasikan waktu FGD dengan pimpinan dan guru di SMA Kayutanam. Selain itu FGD juga bertujuan untuk merumuskan strategi pengembangan SMA INS Kayutanam kedepan. Berdasarkan FGD dapat dikelompokkan kondisi sekolah saat ini menjadi 16 dan dikembangkan strategi sebanyak 16 juga. Kondisi sekolah saat ini dan strategi pengembangannya adalah sebagai berikut.

*Pertama*, Kurikulum akademik tidak berbeda dengan SMA biasa, sedangkan harapan guru berserta pimpinan SMA INS Kayutanam adalah tercapai sasaran pendidikan SMA INS Kayutanam dan tujuan pendidikan. Strategi pengembangan yang dirumuskan adalah pembentukan tim perumus kurikulum akademik dan perangkatnya yang dapat memfasilitasi filosofi serta sasaran pendidikan SMA INS Kayutanam dan tujuan pendidikan nasional. *Kedua*, pada prog-

ram keterampilan ditemukan kesesuaian kompetensi dengan proses pelatihan masih kurang, perangkat kurikulum keterampilan belum lengkap. Pimpinan beserta guru memiliki harapan terdapat kesesuaian kompetensi yang diinginkan dengan proses pelatihan serta tercapainya sasaran pendidikan keterampilan. Strategi pengembangan yang dirumuskan adalah Pembentukan tim perumus kurikulum keterampilan yang bisa memfasilitasi filosofi dan sasaran pendidikan keterampilan. *Ketiga*, program akhlak mulia merupakan program tersendiri. Harapan guru dan pimpinan sekolah tercapai sasaran pendidikan rohani sesuai kurikulum INS Kayutanam. Strategi pengembangan yang dirumuskan adalah sebagai berikut: (a) pembentukan tim perumus untuk mengintegrasikan program akhlak mulia dengan program akademik dan keterampilan. (b) pembentukan kurikulum yang bisa memfasilitasi sasaran pendidikan rohani. (c) pengembangan budaya sekolah yang kondusif untuk pembentukan akhlak mulia.

*Keempat*, kompetensi guru masih perlu ditingkatkan. Harapan guru dan pimpinan adalah guru yang kompeten pada bidangnya serta memahami filosofi serta sasaran pendidikan Ruang Pendidik INS Kayutanam. Strategi pengembangan yang dirumuskan adalah peningkatan kompetensi guru dalam bidang Teknologi Informasi, prosedur pengembangan instrumen tes/nontes dan bidang kosep serta sasaran pendidikan di INS Kayutanam melalui: pelatihan/workshop, pertemuan MGMP 1 kali dalam seminggu, melanjutkan pendidikan ke jenjang yang lebih tinggi, dan pembentukan program *mentoring* bagi guru-guru muda. *Kelima*, fasilitas berupa gedung perkantoran, pembelajaran, asrama, rumah guru dan kepala sekolah serta laboratorium ada yang rusak ringan. Harapan pimpinan dan guru adalah gedung perkantoran, pembelajaran, asrama, dan laboratorium serta rumah guru dan kepala sekolah yang kondusif. Strategi pengembangan yang dirumuskan adalah sarana prasarana harus memenuhi standar minimal, maka harus dilakukan langkah-langkah sebagai berikut: (a) melengkapi sarana pendi-

dikan terutama alat peraga dan *infocus* termasuk laboratorium fisika serta biologi; (b) perbaikan sarana dan prasarana asrama siswa dan diklat, rumah guru dan kepala sekolah; (c) penataan taman; (d) penataan tata ruang kampus. Pelaksanaan strategi ini oleh yayasan dan didukung oleh pemerintah daerah serta bantuan masyarakat.

*Keenam*, rasio guru dengan siswa program keterampilan belum efisien. Peralatan untuk latihan keterampilan ada yang kurang lengkap dan ada yang lengkap. Harapan pimpinan dan guru adalah (a) rasio guru dan siswa yang ideal pada program keterampilan; (b) peralatan latihan keterampilan yang lengkap, relevan, dan canggih. Strategi pengembangan yang dirumuskan adalah; (a) pembahasan variasi program keterampilan dengan memilih program yang *marketable* dan mengarah kepada kewirausahaan; (b) mengembangkan seni tradisi yang bisa meminta partisipasi partisipasi tukang ahli yang ada di masyarakat. Di antara keterampilan yang mungkin dilanjutkan adalah kriya meubel, teknik las, keramik dan desain grafis.

*Ketujuh*, peranan masyarakat dan orang tua siswa dalam hal pendanaan masih kurang. Harapan pimpinan dan guru adalah optimalnya peranan masyarakat dan orang tua/wali siswa. Strategi pengembangan yang dirumuskan sekolah harus bisa memanfaatkan modal sosial, salah satunya dengan memfungsikan komite sekolah menurut tugas pokok dan fungsi. *Kedelapan*, belum ada kerja sama dengan instansi yang relevan. Harapan pimpinan dan guru adalah adanya kerja sama dengan sekolah, perguruan tinggi dan perusahaan seperti PT. Semen Padang. PT. DOCK dan lain-lain. Strategi pengembangan yang dirumuskan adalah dibentuknya tim yang bertugas untuk penelusuran kerja sama dengan instansi lain yang dikoordinatori oleh wakil kepala sekolah bidang humas. MOU bisa dibangun dengan perusahaan yang punya *corporate social responsibility* (CSR), toko pengrajin, serta perusahaan yang sesuai dengan program keterampilan di sekolah. *Kesembilan*, kemampuan pembiayaan operasional masih kurang. Harapan pimpinan dan guru adalah meningkatnya kemam-

puan sekolah untuk biaya operasional termasuk gaji pendidik dan tenaga kependidikan. Strategi pengembangan yang dirumuskan adalah sebagai berikut: (a) menggali potensi masyarakat dengan modal sosial. (b) mencari perusahaan yang punya CSR untuk penyaluran karya siswa agar bisa dipasarkan; (c) meningkatkan kualitas hasil program keterampilan dengan seni yang tinggi agar bisa/laku dipasaran dan kerja sama dengan perusahaan; (d) pengelolaan sumber daya yang dimiliki Ruang Pendidik INS Kayutanam secara profesional dengan cara memberdayakan karyawan dan masyarakat (dalam hal pengelolaan pertanian, peternakan sapi, pengelolaan biogas, perikanan, waserda, restoran diklat, bumi perkemahan, dan lain-lain); (e) perlu pemberdayaan dengan konsep *entrepreneurship*. *Kesepuluh*, masih terdapat kecanduan kewenangan sekolah dalam pengelolaan. Harapan pimpinan dan guru adalah kepala sekolah memiliki kewenangan dalam pengelolaan pendidikan dan sumber daya yang ada di sekolah. Strategi pengembangan yang dirumuskan adalah pembagian kewenangan yang jelas antara yayasan (Badan Wakaf) dengan sekolah serta pemberian otonomi kepada kepala sekolah dalam mengelola sumber daya yang dimiliki dengan berpedoman pada UU No. 16 Tahun 2001 tentang Yayasan dan UU No. 28 Tahun 2004 tentang Yayasan serta Anggaran Dasar Dewan Sekolah dan Anggaran Rumah Tangga Dewan Sekolah Ruang Pendidik INS Kayutanam.

*Kesebelas*, proses pembelajaran akademik dari segi jumlah pelajaran masih kurang. Harapan pimpinan dan guru jumlah jam pelajaran akademik yang ideal. Strategi pengembangan yang dirumuskan pembagian jam pelajaran disesuaikan dengan bobot mata pelajaran dan intensitas penggunaan waktu dalam hal *time on tasks and time on learning*. *Kedua belas*, Belum memiliki instrumen yang terstandar untuk menilai proses dan hasil kegiatan keterampilan. Harapan pimpinan dan guru adalah sekolah memiliki instrumen yang terstandar untuk menilai proses dan hasil program keterampilan. Strategi pengembangan yang dirumuskan

adalah sebagai berikut: (a) pembentukan tim pengembangan instrumen penilaian untuk penilaian proses dan hasil pada program keterampilan; (b) pelaksanaan pelatihan penyusunan instrumen penilaian agar guru mampu menyusun instrumen standar. *Ketiga belas*, Belum memiliki instrumen yang terstandar untuk menilai proses dan hasil program akhlak mulia. Harapan pimpinan dan guru adalah sekolah memiliki instrumen yang terstandar untuk menilai proses dan hasil akhlak mulia. Strategi pengembangan yang dirumuskan adalah sebagai berikut: (a) pembentukan tim pengembangan instrumen penilaian afektif; (b) pelaksanaan pelatihan penyusunan instrumen afektif.

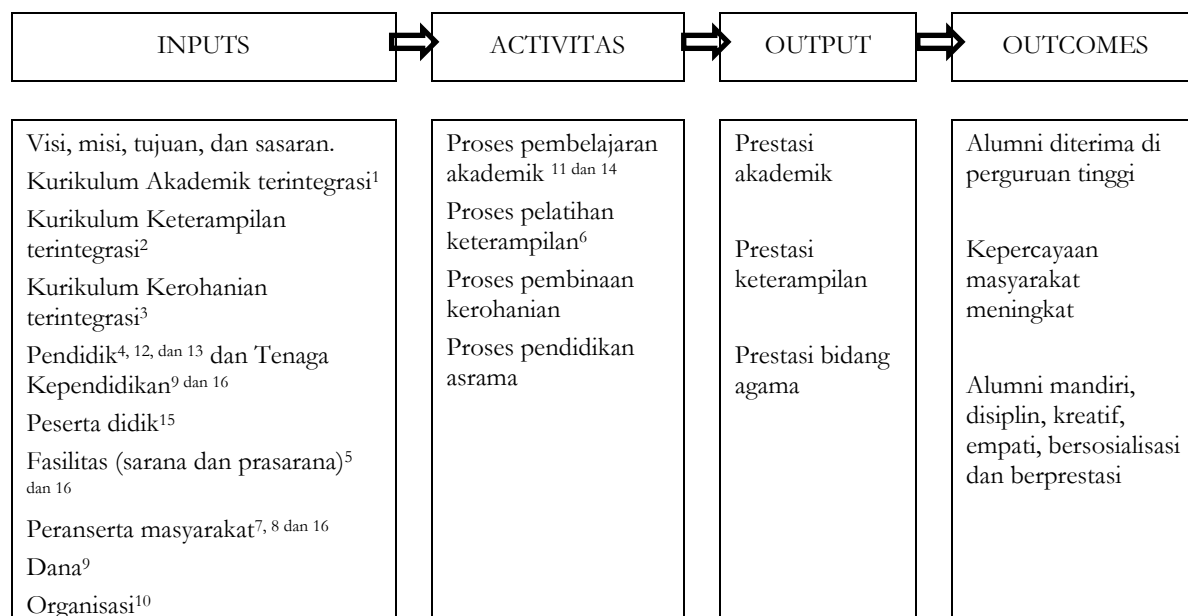
*Keempat belas*, berdasarkan ujian nasional daya serap siswa masih termasuk kategori kurang. Harapan pimpinan dan guru adalah meningkatnya daya serap siswa apabila dilihat dari hasil ujian nasional. Strategi pengembangan yang dirumuskan adalah sebagai berikut: (a) penggunaan model pembelajaran *inquire, scientific, higher order critical thinking*; (b) beberapa hal untuk meningkatkan kemampuan siswa adalah (i) meningkatkan kedisiplinan guru dan siswa, (ii) meningkatkan motivasi melalui metode yang menarik, (iii) latihan mengerjakan soal yang cepat dan tepat. *Kelima belas*, tiga tahun terakhir alumni yang diterima di perguruan tinggi kurang dari 50%. Harapan pimpinan dan guru, minimal 60% alumni diterima diperguruan tinggi. Strategi pengembangan yang dirumuskan adalah sebagai berikut: (a) mengundang perguruan tinggi untuk sosialisasi; (b) indentifikasi potensi siswa dari tingkat 1 melalui tes bakat; (c) bimbingan karir berdasarkan bakat siswa; (d) bimbingan belajar berdasarkan bakat siswa; (e) pengenalan jurusan yang cocok berdasarkan bakat siswa.

*Keenam belas*, SMA INS Kayutanam kurang diminati oleh calon siswa. Harapan pimpinan dan guru adalah SMA INS Kayutanam menjadi sekolah yang diminati oleh calon siswa. Strategi pengembangan yang dirumuskan adalah sebagai berikut: (a) penggunaan basis modal sosial; (b) sosialisasi profil sekolah agar diketahui masyarakat.

kat; (c) membangun jaringan dengan tokoh masyarakat, perguruan tinggi, dan pemerintah daerah; (d) meningkatkan kepercayaan masyarakat melalui prestasi belajar siswa; (e) penataan ruang kampus menjadi sekolah yang nyaman dan indah; (f) jaminan mutu

pendidikan, keamanan dan kenyamanan sekolah.

Berdasarkan strategi pengembangan yang dikemukakan sebelumnya dapat dikemukakan strukturnya seperti Gambar 2.



Keterangan: Angka 1-16 adalah strategi yang bisa digunakan berdasarkan uraian sebelumnya.

Gambar 2. Struktur Strategi Pengembangan SMA INS Kayutanam

Tabel 1. Hasil Analisis Aiken

Nomor Strategi	Nilai V			
	Pakar		Pimpinan/guru SMA INS Kayutanam	
	V <sub>o</sub>	V <sub>t</sub>	V <sub>o</sub>	V <sub>t</sub>
1	0,857	0,75	1,000	0,79
2	0,821	0,75	0,917	0,79
3	0,821	0,75	1,000	0,79
4	0,821	0,75	0,875	0,79
5	0,821	0,75	0,833	0,79
6	0,750	0,75	0,958	0,79
7	0,786	0,75	0,958	0,79
8	0,893	0,75	0,917	0,79
9	0,821	0,75	0,917	0,79
10	0,821	0,75	0,917	0,79
11	0,857	0,75	0,917	0,79
12	0,857	0,75	0,958	0,79
13	0,857	0,75	1,000	0,79
14	0,857	0,75	0,917	0,79
15	0,821	0,75	1,000	0,79
16	0,893	0,75	0,958	0,79

Setelah enam belas strategi dirumuskan langkah selanjutnya divalidasi oleh tujuh orang pakar dan Universitas Negeri Yogyakarta. Strategi pengembangan yang sudah divalidasi selanjutnya dilakukan uji penerimaan oleh praktisi yaitu pengguna yang dalam hal ini adalah pimpinan dan guru SMA INS Kayutanam. Hasil validasi oleh pakar dan hasil uji penerimaan di analisis dengan Aiken yang hasilnya adalah sebagai disajikan pada Tabel 1.

Berdasarkan penilaian produk penelitian oleh 7 pakar, kemudian dianalisis menggunakan validitas Aiken ditemukan bahwa indeks V yang paling rendah sebesar 0,75 dan yang paling tinggi adalah 0,893. Dengan rater 7 orang dan kategori rating 5 diperoleh  $V_t = 0,75$  dengan  $p = 0,041$ . Dengan demikian seluruh  $V_o \geq V_t$ , sehingga dapat disimpulkan seluruh strategi sudah valid. Apabila dilihat rata-rata penilaian rater terhadap setiap strategi pengembangan di-

temukan rata-rata yang paling rendah 4,00 dan yang paling tinggi 4,571, sehingga dapat disimpulkan bahwa seluruh strategi pengembangan SMA INS Kayutanam termasuk pada kategori baik dan dapat digunakan.

Meskipun sudah baik dan dapat digunakan menurut tujuh orang pakar, masih perlu uji keberterimaan dari pihak sekolah. Uji keberterimaan ini dengan cara meminta pendapat serta penilaian dari kepala sekolah, wakil kepala sekolah dan ditambah dua orang guru. Hasil penilaian dari pihak sekolah dianalisis menggunakan Aiken. Berdasarkan hasil analisis didapatkan  $V$  terendah 0,833 dan yang tertinggi 1,000. Dengan rater 6 orang dan kategori rating 5 diperoleh  $V_t$  sebesar 0,79 dengan  $p = 0,029$ , sehingga dapat disimpulkan bahwa produk penelitian ini menurut pengguna sudah layak untuk dilaksanakan. Selain penilaian dari pimpinan sekolah dan guru juga diminta tanggapan mereka terhadap strategi pengembangan SMA INS Kayutanam. Pimpinan dan guru menyatakan bahwa strategi ini sudah layak untuk dilaksanakan berdasarkan skala prioritas.

Secara umum strategi pengembangan SMA INS Kayutanam dapat dirangkum sebagai berikut: (1) perumusan kurikulum terintegrasi yang memuat tujuan dan sasaran pendidikan Sjafei; (2) perumusan metode mengajar yang sesuai dengan tujuan dan sasaran pendidikan Sjafei; (3) pengelolaan aset produktif dengan pemanfaatan modal sosial; (4) pemanfaatan basis alumni. (5) membangun kerja sama dengan instansi terkait seperti sekolah lain, pemerintah daerah, dan pihak swasta.

#### Pembahasan

Pendidikan akhlak mulia bisa dilaksanakan bersamaan dengan program akademik dan program keterampilan dengan cara menintegrasikan materi akhlak mulia dengan akademik begitu juga dengan keterampilan. Kurikulum terintegrasi memungkinkan siswa untuk aktif menggali dan menemukan konsep dan prinsip-prinsip secara holistic bermakna dan otentik baik secara individual ataupun secara berkelompok. Temuan pe-

nelitian Feng, A.X; Tassel-Baska, V., dkk (2005, p.83) kurikulum terintegrasi (*Integrated Curriculum Model/ICM*) mampu meningkatkan prestasi siswa dari tahun 1996-2002.

Kurikulum merupakan suatu hal yang tidak dapat dipisahkan dari pendidikan dan pengajaran, termasuk dalam kegiatan pelatihan. Mengingat pentingnya kurikulum, termasuk untuk pelatihan keterampilan, di SMA INS Kayutanam perlu dibentuk tim perumus kurikulum keterampilan yang akan merumuskan silabus, RPP, materi dan termasuk strategi pelatihan yang akan digunakan. Sebagai acuan untuk merumuskan kurikulum program keterampilan bisa menggunakan Peraturan Menteri Pendidikan Nasional Nomor 28 Tahun 2009 tentang Standar Kompetensi Kejuruan Sekolah Menengah Kejuruan (SMK)/Madrrasah Aliyah Kejuruan (MAK). Meski SMA INS Kayutanam bukan sekolah kejuruan tetapi tidak salah jika sekolah ini memberikan program keterampilan kepada siswa, selain itu sekolah ini juga disebut SMA Plus, serta didukung oleh UU No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional pasal 36 ayat 3 butir d mengamanatkan bahwa kurikulum disusun sesuai dengan memperhatikan keragaman daerah dan lingkungan.

Berdasarkan hasil evaluasi dan diskusi dengan pimpinan beserta guru di SMA INS Kayutanam ditemukan bahwa kompetensi guru pada program akademik masih perlu ditingkatkan supaya guru-guru yang mengajar adalah guru yang benar-benar profesional dibidannya. Salah satu penyebab rendahnya kualitas pendidikan menurut King (1998, p.84-85) adalah tenaga pengajar yang tidak memenuhi standar, kemudian Edward (1998, p.471) mengemukakan salah satu temuan *Inter-American Development Bank*, faktor yang mempengaruhi lembaga pendidikan guru dan tenaga administrasi yang kurang akuntabel. Guru yang profesional menurut UU Guru dan Dosen No. 14 Tahun 2005 dan Permendiknas No. 16 Tahun 2007 adalah guru yang memiliki kompetensi pedagogik, kepribadian, sosial dan profesional. Peningkatan kompetensi ini bisa dilakukan

melalui (1) pelatihan/workshop, (2) pertemuan Musyawarah Guru Mata Pelajaran (MGMP) 1 kali dalam satu minggu, dan (3) melanjutkan pendidikan ke jenjang yang lebih tinggi.

Berdasarkan hasil evaluasi ternyata kegiatan MGMP oleh guru belum optimal. MGMP merupakan salah satu kegiatan untuk meningkatkan kompetensi guru, hal ini berdasarkan hasil temuan penelitian Alfian (2012, p.70) terdapat hubungan positif yang signifikan antara pengembangan profesionalisme guru dalam bentuk MGMP dengan peningkatan sekolah di MTs se Kota Binjai. Variabel pengembangan profesional guru dalam bentuk MGMP memberikan sumbangan efektif sebesar 16,68%. Temuan penelitian Tanjung (2009, p.ii) terdapat hubungan positif dan signifikan antara pemberdayaan MGMP dengan kinerja guru matematika dengan angka korelasi 0,386. Temuan penelitian Gunarno (2012, p.84) kegiatan *lesson study* berbasis MGMP meningkatkan persentase ketercapaian kompetensi pedagogik guru IPA MTs se-Kota Binjai rata-rata sebesar 6,84%. Temuan Penelitian Tasraida (2009, p.ii): (1) peran serta dalam MGMP berhubungan signifikan dengan kemampuan pelaksanaan pembelajaran dengan indeks 0,462; (2) komunikasi antar pribadi dan peran serta dalam MGMP berhubungan signifikan dengan kemampuan pelaksanaan pembelajaran dengan indeks 0,634, (3) motivasi kerja dan peran serta dalam MGMP berhubungan signifikan dengan kemampuan pelaksanaan pembelajaran dengan indeks 0,505. Berdasarkan beberapa temuan penelitian tersebut, salah satu untuk meningkatkan kompetensi guru adalah dengan mengikuti kegiatan MGMP 1 kali dalam satu minggu.

Berdasarkan hasil evaluasi fasilitas pendidikan di SMA INS Kayutanam secara umum termasuk kategori cukup. Fasilitas pendidikan baik yang secara langsung digunakan dalam proses pendidikan berupa peralatan pembelajaran dan yang secara tidak langsung seperti gedung perkantoran, gedung belajar, asrama, lapangan serta yang lainnya akan mempengaruhi proses pendi-

dikan. Prasarana yang lengkap, layak, bersih, dan nyaman akan membuat siswa merasa puas dan senang. Kepuasan siswa dan orang tua sebagai salah satu pengguna fasilitas merupakan suatu pendekatan manajemen untuk keberhasilan jangka panjang (Isumura et. al, 2000, p.176), Haselden (2003, p.188) juga mengemukakan dua dari empat prinsip menggunakan manajemen mutu terpadu adalah memperhatikan kepuasan pelanggan dan perbaikan secara terus menerus baik fisik maupun nonfisik. Berdasarkan hasil evaluasi fasilitas berupa gedung perkantoran, pembelajaran, asrama dan laboratorium perlu dilakukan perbaikan secara berkala dengan mempertimbangkan skala prioritas.

Kemudian dari segi peralatan laboratorium fisika dan biologi diperlukan inovasi agar bisa mengoptimalkan peralatan yang ada serta memanfaatkan sumber dari alam sesuai dengan falsafah "*alam takambang jadi guru*". Peralatan yang tidak bisa diambilkan dari lingkungan sekitar secara bertahap dilengkapi berdasarkan skala prioritas.

Rasio siswa dengan guru menurut Peraturan pemerintah No. 74 Tahun 2008 tentang guru, untuk SMA atau yang sederajat 20:1, dan SMK adalah 15:1. Sementara di SMA INS Kayutanam rasionya 2:1 sampai 5:1 untuk program keterampilan, sedangkan untuk program akademik rasio guru dengan siswa 10:1 sampai 17:1. Jumlah siswa sedikit untuk satu paket program keterampilan karena variasi program cukup banyak sementara jumlah siswa sedikit. Pada program keterampilan jumlah ini tidak efisien. Pada sisi lain peralatan untuk program keterampilan ada yang kurang lengkap dan ada juga yang sudah lengkap. Berdasarkan rasio guru dengan siswa dan kondisi peralatan program keterampilan perlu dilakukan pembatasan variasi program keterampilan. Program keterampilan yang mungkin untuk dilaksanakan adalah kriya meubel, teknik las, dan elektronika. Pemilihan program keterampilan ini didasarkan oleh peralatan yang mendukung dan dilihat dari aspek kebutuhan pasar.

Temuan penelitian Prabhakar (2011, pp.107-118) dengan 188 sampel dari 557

populasi ditemukan bahwa di sekolah negeri orang tua/masyarakat terlibat aktif dalam perencanaan kerangka kerja, sedangkan di sekolah swasta dalam perencanaan kerja orang tua dan masyarakat tidak terlibat. Temuan penelitian ini tidak jauh berbeda kondisinya dengan SMA INS Kayutanam yaitu peranan orang tua dan masyarakat masih kategori kurang. Lebih lanjut, Prabhakar menyampaikan bahwa dalam perencanaan kerangka kerja harus dikembangkan oleh pemangku kepentingan seperti kepala sekolah, guru, orang tua, dan masyarakat.

Pelibatan masyarakat dalam kerangka kebijakan pendidikan di sekolah merupakan salah satu dari inti manajemen berbasis sekolah (Mulyasa, 2009, p.24). Masyarakat menginginkan pendidikan yang bermutu, untuk itu sekolah harus dapat menjalin kerja sama dengan keluarga siswa dan masyarakat (Sagala, 2011, p.251). Salah satu wadah pelibatan masyarakat yang bisa melibatkan masyarakat dalam pendidikan sekolah adalah komite sekolah. Landasan hukum komite sekolah adalah Permendiknas No. 044/U/2002 tentang Dewan Pendidikan dan Komite Sekolah. Dalam rencana kerja sekolah Komite berfungsi sebagai penasehat, badan pendukung, pengontrol, dan penghubung. Agar peranan masyarakat lebih optimal terhadap SMA INS Kayutanam memfungsikan komite sekolah menurut semestinya adalah suatu kebutuhan.

Berdasarkan hasil evaluasi kerja sama sekolah dengan pihak lain masih termasuk kategori kurang. Kerja sama sekolah dengan pihak lain merupakan hubungan kerja sama antara sekolah dengan mitranya, di antara mitra tersebut bisa perseorangan, perusahaan, yayasan, organisasi nirlama, lembaga pendidikan, universitas, dan lain-lain (Sagala, 2011, p.98) bertujuan kepada perubahan yang lebih baik dalam penyelenggaraan pendidikan di sekolah. Dasar hukum yang bisa digunakan terkait kerja sama dengan pihak ketiga ini adalah UU No. 20 tahun 2003 tentang Sistem Pendidikan Nasional ayat 6 yang berbunyi "Memberdayakan semua komponen masyarakat berarti pendidikan diselenggarakan oleh pemerintah dan ma-

syarakat dalam suasana kemitraan dan kerja sama yang saling melengkapi dan memperkuat.

SMA INS Kayutanam membutuhkan kerja sama dengan pihak lain seperti sekolah negeri ataupun swasta terutama sekolah SMP/MTs, dan Pondok Pesantren. Kemudian juga bisa membangun kerja sama dengan sekolah yang setingkat dan termasuk juga kerja sama perguruan tinggi. Selain kerja sama dengan pihak sekolah dan perguruan tinggi SMA INS Kayutanam juga bisa membangun kerja sama dengan perusahaan seperti PT Semen Padang, PT.DOCK di Teluk Bayur dan lain-lain. Dalam rangka membangun kerja sama tersebut SMA INS Kayutanam terlebih dahulu membentuk tim yang bertugas untuk penelusuran kerja sama dengan instansi lain. Tim kerja sama ini bisa langsung dipimpin oleh wakil kepala sekolah bagian hubungan masyarakat. Penambahan tugas wakil kepada bidang hubungan masyarakat ini sedikit merubah struktur menjadi wakil kepala bidang humas dan kerja sama.

Berdasarkan hasil evaluasi ditemukan bahwa SMA INS Kayutanam kurang memiliki kemampuan dari segi pembiayaan, baik biaya operasional ataupun biaya investasi. Salah satu dari inti manajemen berbasis sekolah adalah pemberian kewenangan kepada sekolah untuk mengelola sumber daya dan sumber dana dengan memprioritaskan kebutuhan serta tanggap terhadap kebutuhan setempat (Mulyasa, 2009, p.24). Slamet PH (2014, p.6) mengemukakan bahwa salah satu dari tujuan rencana pengembangan sekolah adalah untuk menjamin tercapainya penggunaan sumber daya secara efisien, efektif, berkeadilan dan berkelanjutan.

Meskipun SMA INS Kayutanam memiliki kemampuan yang kurang untuk pembiayaan operasional dan investasi, namun sekolah ini memiliki sumber daya yang bisa dikelola. Pengelolaan sumber daya ini secara profesional merupakan salah satu sumber dana bagi sekolah. Di antara sumber daya yang dimiliki sekolah adalah lahan pertanian, perkebunan, peternakan, perikanan, waserda, restoran, diklat, bumi perkemahan,

pemasaran hasil karya siswa, dan penyediaan pembuatan perabot dan kebutuhan bangunan. Pada sisi lain sekolah ini memiliki karyawan selain guru sebanyak 24 orang. Pengelolaan sumber daya yang ada tersebut bisa dengan mengalihkan tugas pokok dan fungsi sebagian karyawan sekolah dan mengikutsertakan masyarakat dilingkungan sekolah. Terkait pengelolaan sumber daya di sekolah dipimpin oleh wakil kepala bidang sarana dan prasarana. UU No. 20 tahun 2003 tentang Sistem Pendidikan Nasional Pasal 51 ayat (1) bisa dijadikan sebagai dasar hukum yaitu pengelolaan satuan pendidikan anak usia dini, pendidikan dasar, dan pendidikan menengah dilaksanakan berdasarkan standar pelayanan minimal dengan prinsip manajemen berbasis sekolah/madrasah.

Berdasarkan hasil evaluasi ternyata kepala sekolah belum memiliki kewenangan penuh dalam mengelola sekolah, kadang-kadang masih ada tumpang tindih tanggung jawab. Salah satu temuan dari *Inter-American Development Bank* yang mempengaruhi lembaga pendidikan adalah kurang otomom dalam mengelola sekolah (Edward, 1998, p.471). Kemudian King (1998, p.85) mengemukakan faktor yang mempengaruhi lembaga pendidikan salah satunya adalah tumpang tindih tanggung jawab. Pemberian kewenangan penuh kepada sekolah dalam mengelola sumberdaya dan kebutuhan akan efektif memperbaiki mutu sekolah. Manajemen yang memberikan peluang otomomi kepala sekolah adalah manajemen berbasis sekolah. Penggunaan manajemen berbasis sekolah adalah suatu hal yang tepat terhadap peningkatan sekolah dan prestasi siswa Gammage (2008, pp.664-675). Kemudian dikemukakan oleh Botha (2006, p.341) bahwa beberapa penelitian menunjukkan MBS secara konsisten memperbaiki mutu sekolah seperti yang dikemukakan pada halaman 64 disertasi.

Pembagian kewenangan yang jelas antara yayasan dengan sekolah perlu diperjelas batasan-batasannya sehingga tidak terjadi tumpang tindih. Pemberian kewenangan penuh kepada kepala sekolah untuk mengelola

pendidikan dan mengelola sumber daya yang ada merupakan langkah yang tepat, karena sehari-hari yang lebih banyak di sekolah adalah kepala sekolah beserta staf. Pihak yayasan cukup memberikan target kepada kepala sekolah. Target yang diberikan itu bisa dievaluasi sesuai dengan waktu yang ditentukan. Apabila ada suatu instruksi yang akan diberikan ke sekolah cukup melalui kepala sekolah jangan langsung ke karyawan atau guru, karena hal ini akan menimbulkan dualisme kepemimpinan yang bisa menyebabkan karyawan atau guru kebingungan apabila instruksi itu berbeda dengan yang disampaikan oleh kepala sekolah. Edward (1998, p.471) mengemukakan salah satu masalah yang dihadapi oleh sekolah adalah kurang otonom dalam mengelola sekolah dan tidak adanya mekanisme evaluasi dan standar akademik.

Berdasarkan hasil evaluasi ternyata jumlah jam untuk mata pelajaran yang diterapkan di SMA INS Kayutanam. Jumlah jam program akademik sehari 5 jam per hari, sedangkan di sekolah yang sederajat 6-7 jam per hari. Kesepakatan waktu acara diskusi guru beserta pimpinan untuk hari Senin sampai Rabu pembelajaran dimulai 06.45-15.45, untuk hari Kamis jam pelajaran tetap seperti biasa yaitu 12.15 dan setelah salat zuhur pelaksanaan program keterampilan kemudian untuk hari Jumat dan Sabtu berjalan seperti biasa. Meskipun jam pelajaran dimulai pukul 06.45 kemungkinan siswa tidak akan terlambat masuk disebabkan siswa tinggal di asrama. Peak (2012, p.13) mengemukakan langkah yang perlu diperhatikan apabila sekolah mengalami penurunan salah satunya adalah meninjau ulang kalender kelas dan jadwal harian.

Program keterampilan dan program akhlak mulia di SMA INS Kayutanam belum memiliki instrumen yang terstandar untuk menilai proses dan hasil program keterampilan dan akhlak mulia. Jika sekolah ingin untuk meningkatkan kualitas program ini, sekolah akan kesulitan karena kegiatan yang sudah berjalan belum terukur keberhasilannya atau belum diketahui tingkat keberhasilannya. Untuk itu langkah utama yang

harus dilakukan di sekolah ini adalah pembentukan tim pengembangan instrumen yang bisa untuk menilai proses dan hasil program keterampilan dan program akhlak mulia. Kartowagiran (2013, p.19) mengemukakan bahwa evaluasi merupakan kegiatan yang sistematis untuk mengidentifikasi, mengklarifikasi dan mengaplikasikan suatu kriteria untuk menentukan keberhasilan suatu program. Jika suatu kegiatan dalam pembelajaran atau pelatihan belum memiliki instrumen untuk melakukan penilaian, sekolah tidak mungkin bisa mengetahui keberhasilan dari suatu kegiatan tersebut, sehingga untuk meningkatkan kualitas kerja atau produktivitas suatu lembaga dalam melaksanakan program tidak dapat dilaksanakan. Dikemukakan oleh Mardapi (2008, p.8) bahwa evaluasi merupakan suatu rangkaian kegiatan dalam meningkatkan kualitas kinerja, atau produktivitas suatu lembaga dalam melaksanakan program.

Berdasarkan ujian nasional ternyata daya serap siswa masih termasuk kategori kurang. Di antara strategi yang dapat meningkatkan kemampuan siswa yang dikemukakan oleh guru dan pimpinan adalah meningkatkan kedisiplinan guru dan siswa. Salah satu dari hasil penelitian Suparman (2003, pp.133-134) dan diterbitkan tahun 2012 adalah pola pembinaan disiplin yang dilakukan SMU Negeri 1 Cisarua berkerja sama dengan Yayasan Dharmaloka bagi siswa yang diasramakan cenderung lebih baik dan akan membawa hasil yang optimal jika dibandingkan dengan pembinaan disiplin melalui sekolah secara regular. Pembinaan disiplin melalui asrama ini sangat mungkin dilakukan di SMA INS Kayutanam karena siswa sudah tinggal di asrama. Namun, disiplin siswa saja tidak cukup berdasarkan salah satu dari temuan penelitian Yusuf (1987, p.126) yang diterbitkan tahun 2012, ternyata terdapat ketergantungan disiplin diri dalam belajar siswa terhadap penanaman disiplin yang dilakukan oleh guru.

Hasil belajar siswa ternyata tidak saja dipengaruhi oleh disiplin siswa tetapi juga dipengaruhi oleh disiplin guru. Hasil penelitian Siswatmaja (1987, pp.132-137) dan di-

terbitkan tahun 2012 adalah prestasi siswa dalam bidang studi Ilmu Pengetahuan Sosial kelas V Sekolah Dasar Negeri Kotamadya Pontianak lebih banyak tergantung pada disiplin guru dalam kegiatan proses belajar mengajar yang dilakukannya.

Selain peningkatan disiplin siswa perlu juga peningkatan motivasi belajar siswa dengan metode yang menarik, karena motivasi akan mempengaruhi hasil belajar. Strategi ini juga didukung oleh beberapa hasil penelitian diantaranya Hutapea (2012, p.129) (1) hasil belajar siswa yang diajar dengan pembelajaran tutor teman sebaya lebih tinggi daripada pembelajaran ekspositori, (2) hasil belajar siswa yang mempunyai motivasi berprestasi tinggi lebih baik dari siswa yang mempunyai motivasi berprestasi rendah, (3) terdapat interaksi antara strategi pembelajaran dengan motivasi berprestasi terhadap hasil belajar. Kemudian, hasil penelitian Sinaga (2007, p.115) ternyata prestasi belajar uji dapat ditingkatkan dengan meningkatkan pengelolaan Pendidikan Sistem Ganda (PSG) dan motivasi belajar. Temuan penelitian Mardiana (2014, p.ii) di antaranya adalah terdapat perbedaan hasil belajar fisika siswa yang memiliki motivasi tinggi dan siswa yang memiliki motivasi belajar yang rendah, terdapat interaksi antara model pembelajaran pembelajaran *group investigation* dan ekspositori dengan motivasi belajar dalam meningkatkan hasil belajar fisika.

Filosofi INS Kayutanam adalah “jangan minta buah mangga ke pohon rambutan, tapi jadikanlah setiap pohon berbuah manis”. Semenjak awal sekolah ini didirikan pendirinya sudah menyadari bahwa manusia/peserta didik memiliki potensi atau bakat yang berbeda-beda, hal ini terlihat dari rumusan filosofi sekolah. Bakat tersebut dapat diklasifikasikan menjadi delapan yaitu *intelligensi*, bakat *verbal*, numerik, skolastik, relasi ruang, mekanik, abstrak, dan bahasa. Dengan mengetahui kemampuan dasar siswa dan diarahkan ke jurusan yang tepat siswa akan bisa memahami materi pelajaran dalam tempo yang relatif singkat. Selain itu, perlu juga diketahui tipe pekerjaan yang te-

pat bagi siswa tersebut. John L Holland membagi tipe kepribadian manusia menjadi enam bagian yaitu *realistic, investigative, artistic, social, enterprising, dan conventional*.

Berdasarkan keterangan sebelumnya, maka salah satu strategi agar siswa SMA INS Kayutanam bisa diterima di Perguruan Tinggi adalah mengidentifikasi bakat yang dimiliki oleh siswa semenjak dari siswa masuk ke sekolah ini. Kegiatan identifikasi bakat siswa ini bisa dengan cara melakukan kerja sama dengan perguruan tinggi terkait. Setelah diketahui bakat siswa, kemudian dilakukan bimbingan belajar serta bimbingan karier sesuai dengan bakat siswa. Kerja sama dengan perguruan tinggi tidak hanya sekedar untuk mengidentifikasi bakat siswa tetapi juga sebagai sosialisasi tentang cara belajar diperguruan tinggi sehingga akan menimbulkan motivasi bagi siswa untuk melanjutkan pendidikannya, selain itu juga berfungsi untuk pengenalan jurusan yang cocok dengan bakat yang dimiliki oleh siswa.

Semenjak tahun 2014 SMA INS Kayutanam sudah terakreditasi A yang sebelumnya akreditasi sekolah ini B, namun tidak banyak masyarakat yang mengetahuinya. Masih ada anggapan masyarakat bahwa sekolah ini adalah sekolah tukang. Berdasarkan salah satu hasil penelitian Sufyarma (2002, p.499) persepsi masyarakat terhadap Ruang Pendidik INS Kayutanam adalah sebagai sekolah tukang, sekolah seni, tempat rehabilitasi mental, dan masyarakat yang bermental feodal tidak mau anaknya masuk sekolah kerja, karena sulit melanjutkan studi dan sulit menjadi pegawai negeri sipil.

Persepsi masyarakat sesuai dengan penelitian Sufyarma tersebut perlu diluruskan bahwa sekolah ini bukan sekolah tukang, tetapi sekolah ini telah setara dengan sekolah sederajat lainnya bahkan masih ada sekolah yang tarafnya di bawah sekolah ini. Berdasarkan Keputusan Badan Akreditasi Provinsi Sekolah/Madrasah Sumatera Barat Nomor: 1145/BAP-SM/LL/2013 tentang hasil penetapan hasil akreditasi sekolah/madrasah di Sumatera Barat, dari 105 sekolah/madrasah yang terakreditasi A sebanyak 22,86% dan SMA INS Kayutanam terma-

suk ke dalam 22,86%, dengan kata lain ada sebanyak 77,14% sekolah/madrasah yang dibawah SMA INS Kayutanam BAP-SM tanggal 21 Desember 2013.

Strategi peningkatan kepercayaan masyarakat tersebut berkaitan dengan peningkatan prestasi siswa baik berdasarkan hasil ujian nasional ataupun seberapa banyak lulusan sekolah ini yang diterima di perguruan tinggi terutama perguruan tinggi favorit, termasuk prestasi siswa di bidang keterampilan dan akhlak mulia yaitu sejalan dengan strategi peningkatan daya serap siswa pada ujian nasional dan strategi agar alumni minimal diterima diperguruan tinggi 60%.

Peningkatan kepercayaan masyarakat bisa juga dengan penataan ruang kampus menjadi sekolah yang nyaman dan indah, jaminan mutu pendidikan, keamanan sekolah. Semua kegiatan yang dilaksanakan tersebut tidak akan dikenal oleh masyarakat luas kalau tidak dipublikasikan dan disosialisasikan ke masyarakat. Untuk itu seluruh prestasi, program, kondisi sekolah dari segi keamanan, kenyamanan, keindahan, serta jaminan mutu pendidikan harus dipublikasikan kemasyarakat. Kegiatan publikasi dan sosialisasi ini merupakan tanggung jawab wakil kepala bidang humas. Mengoptimalkan peranan alumni merupakan strategi meningkatkan minat masyarakat untuk memasukkan anaknya ke sekolah ini.

## **Simpulan dan Saran**

Berdasarkan temuan data dan hasil analisis dapat diperoleh simpulan bahwa komponen-komponen yang sudah termasuk kategori baik: (1) kurikulum program bidang akademik, perencanaan kegiatan, struktur organisasi dan uraian tugas, pengawasan oleh pimpinan dan tim supervisor, sosialisasi penilaian diawal semester, dan penilaian hasil kerja siswa; (2) pembinaan peningkatan ibadah, kesetiakawanan kejujuran, dan kepemimpinan serta kedisiplinan; (3) peringkat hasil ujian nasional di kabupaten dan provinsi, hasil karya program keterampilan, dan akhlak mulia siswa di kampus; (4) rasio guru program akademik, akhlak mulia. Kualifikasi, latar belakang pendidikan serta sertifi-

kasi guru program akademik, keterampilan dan akhlak mulia; (5) ruang belajar, sebagai ruang keterampilan, ruang laboratorium, aula, ruang makan, masjid; (6) peranan pemerintah (bantuan guru, dana hibah, peralatan keterampilan bus operasional).

Komponen-komponen yang masih memerlukan perbaikan/perhatian: (1) kurikulum kegiatan keterampilan dan akhlak mulia, pelibatan warga sekolah dalam merumuskan visi misi, kesesuaian perencanaan dan pelaksanaan serta pengelolaan kegiatan kesiswaan, otonomi kepala sekolah untuk mengelola sumber daya yang ada, kompetensi guru. Instrumen penilaian proses dan hasil untuk program keterampilan; (2) jumlah jam pada program akademik, proses kegiatan akhlak mulia masih kegiatan tersendiri; (3) daya serap siswa pada terhadap materi ujian nasional; (4) rasio siswa dengan guru pada program keterampilan; (5) fasilitas gedung perkantoran seperti ruang kepala sekolah, wakil, asrama siswa, rumah guru dan pimpinan. Ruang sanggar teater, otomotif, musikalisasi puisi, peralatan keterampilan (kecukupan, relevansi, dan kemutakhiran). Alat peraga pada kegiatan akademik, perawatan lapangan olahraga, lokasi waserda. Pengolahan lahan pertanian, perkebunan, peternakan, perikanan; (6) peranan komite sekolah, peranan alumni, kerja sama dengan instansi terkait termasuk perusahaan.

Strategi pengembangan secara garis besar terdiri dari: (a) Perumusan kurikulum terintegrasi yang memuat tujuan dan sasaran pendidikan yang diprakarsai oleh Sjafei; (b) perumusan metode mengajar yang sesuai dengan tujuan dan sasaran pendidikan yang diprakarsai Sjafei; (c) pengelolaan aset produktif dengan pemanfaatan modal sosial; (d) Pemanfaatan basis alumni; (e) membangun kerja sama dengan instansi terkait seperti sekolah lain, pemerintah daerah, dan pihak swasta.

Berdasarkan hasil evaluasi dan strategi pengembangan yang dirumuskan maka dapat disarankan sebagai berikut: (1) strategi pengembangan SMA INS Kayutanam ini dapat dijadikan acuan untuk pengembangan

SMA INS Kayutanam di masa yang akan datang, karena strategi ini dikembangkan melalui evaluasi dan melibatkan pemangku kepentingan di SMA INS Kayutanam; (2) SMA INS Kayutanam dituntut untuk mengikuti perkembangan ilmu pengetahuan dan teknologi termasuk perubahan kurikulum, namun keunikan sekolah ini harus tetap dipertahankan yaitu dari segi filosofi dan sasaran pendidikannya.

## Daftar Pustaka

- Alfian, M. (2012). *Hubungan kompetensi guru dan pengembangan profesionalisme guru dalam bentuk Musyawarah Guru Mata Pelajaran (MGMP)*. (Tesis magister, tidak diterbitkan). Universitas Negeri Medan, Medan
- Botha, N., (2006). Leadership in school-base management: a case study in selected schools, *South African Journal of Education*, vol. 26, No. 3, pp.. 341-353
- Edward, Beatrice (1998). Neoliberalism and educational reform in Latin America. *Nature, Society, and Thought* 11.4, h. 471
- Feng, A.X; Tassel-Baska, V. (2005). A Longitudinal Assessment of Gifted Students' Learning Using the Integrated Curriculum Model (ICM): Impact and Perceptions of the William and Mary Language Arts and Science Curriculum. *Winter 2005: 27,2; ProQuest Professional Education* pg. 78
- Gammage, D.T. (2008). Three decades of implementation of school-based management in the Australian Capital Territory and Victoria in Australia, *The International Journal of Educational Management*, 664-675
- Gunarno. (2012). *Peningkatan kompetensi pedagogic guru Ilmu Pengetahuan Alam (IPA) melalui kegiatan lesson study berbasis MGMP di Madrasah Tsanawiyah (MTs) se-Kota Binjai*. (Tesis magister, tidak diterbitkan). Universitas Negeri Medan, Medan.

- Haselden, Polly G. (2003). *Use of affinity diagrams as instructional tools in inclusive classrooms*. Preventing School Failure 47.4: 187+. Gale Education, Religion and Humanities Lite Package. Web. 22 Dec. 2012, dari <http://go.galegroup.com/ps/i.do?id=GALE%7CA108548143&v=2.&u=ptn080&it=r&p=GPS&sw>
- Hutapea, F. (2012). Pengaruh strategi pembelajaran dan motivasi berprestasi terhadap hasil belajar membuat hiasan busana siswa SMK Negeri 8 Medan. *Tabularasa, Vol. 9. No. 2, Desember 2012*.
- INS. (t.th.), SMP-SMA 'PLUS'. Ruang Pendidik INS Kayutanam
- Isomura, Y., Yamamoto, M., & Yasuda, K. (2000). *Application of Total Quality Management (TQM) for Diabetes Patient Education Class in Primary Care: Improvement of Motivation for Faculties and Blood Glucose Control Levels for Patients*. Diabetes May 2000: A176. Gale Education, Religion and Humanities Lite Package. Web. 23 Dec. 2012. dari <http://go.galegroup.com/ps/i.do?id=GALE%7CA62891957&v=2.1&u=ptn080&it=r&p=GPS&sw=w>
- Kartowagiran, B. (2013). *Optimalisasi evaluasi pembelajaran teknik mesin melalui logic model untuk meningkatkan soft skills lulusan*. Pidato Pengukuhan Guru Besar
- King, D.Y. (1998). Reforming basic educational and struggle for decentralized educational administrasi in Indoensia. *Journal of Political and Military Sociology: Summer; 26, 1* ProQuest Sociologi, 83-95
- Mardiana. (2014). *Efek model pembelajaran kooperatif tipe group investigation dan motivasi terhadap hasil belajar fisika siswa SMA*. (Tesis magister, tidak diterbitkan). Universitas Negeri Medan, Medan.
- Mertens, D. M. (2004). *Research and Evaluation in Education and Psychology Integrating Diversity With Quantitative, Qualitative, and Mixed Methods*, Thousand, CA: Sage.
- Mulyasa, E. (2009). *Manajemen berbasis sekolah*. Bandung: Rosdakarya.
- Navis, A. A. (1996). *Filsafat dan strategi pendidikan M. Sjafei*. Jakarta: Gramedia Widiasarana Indonesia.
- Peak, M.H., (2012). TQM transforms the classroom, Management Review Sept. 1995: 13. Gale Education, Religion and Humanities Lite Package. Web 23 Oct. 2012
- Prabakar, & Rao, K.V. (2011). School based management: an analysis of the planning framework and community participation. *Journal of Arts, Science & Commerce, vol. II, Issue-3*, 107-118.
- Rossi, H.P., Freeman, H.E., & Lipsey, M.W. (2004). *Evaluation: a systematic aproach*. London: Sage Publications
- Raharja, S. (2008). Penyelenggaraan pendidikan Indonesia Nederlandche School (INS). *Jurnal Manajemen Pen-didikan. No. 01/ThIV/April/2008*. P. 9-19
- Sagala, S. (2011). *Kemampuan professional guru dan tenaga kependidikan*. Badung: Alfabeta.
- Siswatmadja. (1987). *Hubungan antara disiplin kerja guru dalam kegiatan proses belajar mengajar dengan prestasi siswa kelas V bidang studi ilmu pengetahuan sosial Kotamadya Pontianak* (Tesis magister). Universitas Pendidikan Indonesia, Bandung.
- Slamet PH. (7 September 2013). *Rencana pengembangan sekolah (RPS)*. Artikel. Diambil pada tanggal 7 Maret 2014, dari <http://lia.usersdocs.com/docs/124071/index-5789.html>
- Mardapi, D. (2008). *Teknik Penyusunan Instrumen Tes dan non Tes*. Yogyakarta: Mitra Cendikia Press.

- Marsidin, S. (2002). *Revitalisasi Sistem Pendidikan INS Kayutanam dengan memanfaatkan model MBS dalam pengembangan siswa yang mandiri: studi kasus pada Ruang Pendidik INS Kayutanam Sumatera Barat*. (Disertasi doctor, tidak diterbitkan). Universitas Pendidikan Indonesia, Bandung.
- Suparman (2003). *Pengaruh pola pendidikan berbasis disiplin terhadap perkembangan moral siswa di sekolah*. (Tesis magister, tidak diterbitkan). Universitas Pendidikan Indonesia, Bandung.
- Tanjung, A. (2009). *Hubungan antara iklim organisasi, perbedayaan MGMP dan motivasi berprestasi dengan kinerja guru*. (Tesis magister, tidak diterbitkan). Universitas Negeri Medan, Medan
- Tasraida. (2009). *Hubungan komunikasi antar pribadi, motivasi kerja, dan peran serta dalam Musyawarah Guru Mata Pelajaran (MGMP) dengan kemampuan pelaksanaan pembelajaran guru Biologi di SMA Negeri Kota Medan*. (Tesis magister, tidak diterbitkan). Universitas Negeri Medan, Medan
- Tilaar H.A.R., & Nurgroho, R. (2008). *Kebijakan pendidikan, pengantar untuk memahami kebijakan pendidikan sebagai kebijakan publik*. Yogyakarta: Pustaka Pelajar.
- Tirtarahardja, U., Sula, S.L.L. (2012). *Pengantar pendidikan*. Jakarta: Rikeka Cipta.
- Yusuf, S.Y. (1987). *Disiplin diri dalam belajar dibubungkan dengan penanaman disiplin yang dilakukan orang tua dan guru: Studi deskriptif-analitik tentang perilaku bertanggung jawab sebagai tujuan bimbingan terhadap para siswa SMA Negeri di Kotamadya Bandung* (Tesis magister). Universitas Pendidikan Indonesia, Bandung.

## PENGEMBANGAN TES BERPIKIR KRITIS DENGAN PENDEKATAN *ITEM RESPONSE THEORY*

<sup>1)</sup>Fajrianti, <sup>2)</sup>Wiwin Hendriani, <sup>3)</sup>Berlian Gressy Septarini

<sup>1, 2, 3)</sup>Fakultas Psikologi Universitas Airlangga

<sup>1)</sup>fajrianti@psikologi.unair.ac.id, <sup>2)</sup>wiwin.hendriani@psikologi.unair.ac.id,

<sup>3)</sup>gressy.septarini@psikologi.unair.ac.id

### Abstrak

Penelitian ini bertujuan untuk menghasilkan sebuah alat ukur (tes) berpikir kritis yang valid dan reliabel untuk digunakan, baik dalam lingkup pendidikan maupun kerja di Indonesia. Tahapan penelitian dilakukan berdasarkan tahap pengembangan tes menurut Hambleton dan Jones (1993). Kisi-kisi dan pembuatan butir didasarkan pada konsep dalam tes *Watson-Glaser Critical Thinking Appraisal* (WGCTA). Pada WGCTA, berpikir kritis terdiri dari lima dimensi yaitu *Inference*, *Recognition Assumption*, *Deduction*, *Interpretation* dan *Evaluation of arguments*. Uji coba tes dilakukan pada 1.453 peserta tes seleksi karyawan di Surabaya, Gresik, Tuban, Bojonegoro, Rembang. Data dikotomi dianalisis dengan menggunakan model IRT dengan dua parameter yaitu daya beda dan tingkat kesulitan butir. Analisis dilakukan dengan menggunakan program statistik Mplus versi 6.11. Sebelum melakukan analisis dengan IRT, dilakukan pengujian asumsi yaitu uji unidimensionalitas, independensi lokal dan *Item Characteristic Curve* (ICC). Hasil analisis terhadap 68 butir menghasilkan 15 butir dengan daya beda yang cukup baik dan tingkat kesulitan butir yang berkisar antara -4 sampai dengan 2.448. Sedikitnya jumlah butir yang berkualitas baik disebabkan oleh kelemahan dalam menentukan *subject matter experts* di bidang berpikir kritis dan pemilihan metode skoring.

**Kata kunci:** Pengembangan tes, berpikir kritis, *item response theory*

## DEVELOPING CRITICAL THINKING TEST UTILISING ITEM RESPONSE THEORY

<sup>1)</sup>Fajrianti, <sup>2)</sup>Wiwin Hendriani, <sup>3)</sup>Berlian Gressy Septarini

<sup>1, 2, 3)</sup>Fakultas Psikologi Universitas Airlangga

<sup>1)</sup>fajrianti@psikologi.unair.ac.id, <sup>2)</sup>wiwin.hendriani@psikologi.unair.ac.id,

<sup>3)</sup>gressy.septarini@psikologi.unair.ac.id

### Abstract

The present study was aimed to develop a valid and reliable instrument in assessing critical thinking which can be implemented both in educational and work settings in Indonesia. Following the Hambleton and Jones's (1993) procedures on test development, the study developed the instrument by employing the concept of critical thinking from *Watson-Glaser Critical Thinking Appraisal* (WGCTA). The study included five dimensions of critical thinking as adopted from the WGCTA: *Inference*, *Recognition Assumption*, *Deduction*, *Interpretation* dan *Evaluation of arguments*. 1453 respondents from Surabaya, Gresik, Tuban, Bojonegoro and Rembang were used for trailing the test. The dichotomous data were analyzed using the *Item Response Theory* with two parameter logistic model using statistical program Mplus ver. 6.11. Several assumptions were tested prior the IRT analysis; the test of unidimensionality, local independency and *Item Characteristic Curve* (ICC). Amongst 68 items only 15 items had good discrimination parameter. Difficulty item level ranged from - 4.95 to 2.448. The study was limited in producing high number of qualified items due to its failure in finding *subject matter experts* in critical thinking area and inadequate choice in scoring method.

**Keywords:** test development, critical thinking, *Item response theory*

## Pendahuluan

Pemecahan masalah merupakan kegiatan rutin yang berlangsung sepanjang kehidupan manusia. Efektifitas pemecahan masalah tergantung pada banyak faktor. Walaupun demikian, para ahli meyakini bahwa berpikir kritis merupakan salah satu faktor yang penting dalam pemecahan masalah baik di bidang pendidikan maupun bidang kehidupan lainnya. Bahkan, dalam beberapa tahun terakhir, penggunaan konstruk berpikir kritis sebagai prediktor keberhasilan di dunia pendidikan maupun di dunia kerja semakin banyak dilakukan (Wagner, 2002).

Berpikir kritis merupakan keterampilan penting yang dibutuhkan dalam dunia kerja di abad 21 (*Essential Skills for the 21st Century Workplace*, 2014). Keterampilan ini bahkan menduduki urutan pertama dalam daftar keterampilan yang dibutuhkan. Keterampilan komunikasi, kolaborasi, kesadaran global (*global awareness*), penguasaan teknologi, keterampilan dalam hidup dan karir, kemampuan belajar dan inovasi membutuhkan fondasi berpikir kritis yang baik.

Berpikir kritis dipandang sebagai hasil (*outcome*) dari pembelajaran di perguruan tinggi. Hal ini tidak hanya menjadi isu di pihak eksternal atau lingkungan masyarakat namun juga didalam perguruan tinggi itu sendiri. Isu utama yang muncul adalah tentang bagaimana akuntabilitas dari hasil pembelajaran berpikir kritis tersebut (Stassen, Herington, Henderson, 2011). Hal ini terkait dengan hasil survey yang dilakukan oleh Association of American Colleges and Universities (AAC&U) pada tahun 2008 yang menyatakan bahwa 73% dari karyawan menginginkan agar perguruan tinggi memberikan lebih banyak penekanan pada berpikir kritis dan penalaran analitis. Bahkan, survei AAC&U di tahun 2009 menunjukkan bahwa 74% responden menyatakan bahwa berpikir kritis merupakan inti dari tujuan pembelajaran dalam program pendidikan di kampus (Stassen dkk., 2011).

Pentingnya konstruk berpikir kritis dalam dunia pendidikan maupun pekerjaan memang tidak lagi diperdebatkan. Meskipun demikian, pendefinisian maupun bagaimana

konstruk berpikir kritis tersebut diukur masih menjadi perdebatan para ahli di bidang Psikologi, Filsafat maupun Pendidikan. Hal ini terkait dengan beragamnya definisi dan pengukuran konstruk berpikir kritis yang ada saat ini (Wagner, 2002; Stassen dkk., 2011).

Terdapat berbagai definisi tentang berpikir kritis, di antaranya Sternberg (1986) yang menyatakan bahwa berpikir kritis adalah proses mental, strategi dan representasi yang digunakan individu untuk memecahkan, membuat keputusan dan mempelajari konsep baru. Berpikir kritis merupakan investigasi yang bertujuan untuk mengeksplorasi situasi, fenomena, pertanyaan atau masalah untuk menjadi hipotesis atau kesimpulan melalui pengintegrasian seluruh informasi yang tersedia sehingga memiliki justifikasi yang meyakinkan (Kurfiss, 1988). Berpikir kritis mencakup kemampuan berpikir yang masuk akal (*reasonable*) dan reflektif yang berfokus pada keputusan tentang apa yang akan dipercaya atau dilakukan (Noris dan Ennis, 1989).

Angelo dan Cross (1995) menyebutkan berpikir kritis sebagai kegiatan berpikir tingkat tinggi, yang meliputi kegiatan menganalisis, mensintesis, mengenali permasalahan dan pemecahannya, menyimpulkan serta mengevaluasi. Kegiatan ini dilakukan terhadap berbagai informasi yang didapat dari hasil observasi, pengalaman, refleksi, di mana hasil proses ini digunakan sebagai dasar individu saat mengambil tindakan (Walker dan Finney, 1996). Sedikit berbeda dari penjelasan sebelumnya, Hossoubah (2007) mendefinisikan berpikir kritis sebagai kemampuan memberi alasan secara terorganisasi dan mengevaluasi kualitas suatu alasan secara sistematis.

Hasil penelitian pendahuluan menunjukkan bahwa perdebatan mengenai definisi berpikir kritis lebih terletak pada deskripsi tentang indikator daripada perbedaan yang lebih bersifat fundamental. Hal ini didukung oleh laporan hasil penelitian yang ditulis dalam *Critical Thinking: A Literature Review* (Lai, 2011). Dalam *Literature Review* ini dinyatakan bahwa di antara berbagai perbedaan

an mengenai definisi berpikir kritis terdapat kesepakatan bahwa berpikir kritis merupakan kemampuan berpikir yang terdiri dari komponen: keterampilan dalam menganalisis argumen, membuat kesimpulan baik secara induktif maupun deduktif, mengevaluasi dan membuat keputusan atau memecahkan masalah (Lai, 2011).

Pengukuran konstruk berpikir kritis di Indonesia cukup beragam, terutama terkait dengan konteks pengukurannya. Selama ini juga terdapat perdebatan tentang apakah berpikir kritis merupakan konstruk yang spesifik pada setiap *setting* atau bersifat general. Sejauh pengamatan peneliti selama ini, pengukuran berpikir kritis di Indonesia lebih sering menggunakan tes-tes yang dikembangkan di dunia barat. Misalnya, *California Critical Thinking Skills Test* dan *Cornell Critical Thinking Tests* yang digunakan tanpa adanya pengujian terlebih dahulu, apakah tes ini mengandung bias budaya atau tidak. Selain itu pengembangan pengukuran berpikir kritis di Indonesia lebih banyak dilakukan pada *setting* pendidikan matematika dan fisika. Sampai saat ini, di Indonesia belum ada tes berpikir kritis yang terstandar dikembangkan dengan menggunakan *setting* yang general/umum, terutama dengan menggunakan pendekatan *Item Response Theory* (IRT).

IRT adalah sebuah model probabilitas yang berusaha menjelaskan hubungan antara respon seseorang terhadap sebuah butir dengan variabel laten (kemampuan/*ability* atau sifat/*trait*) yang diukur oleh tes tersebut. Dalam hal ini, respon atau kinerja peserta tes merupakan hal yang dapat diamati (*observable*) sedangkan sifat atau kemampuan merupakan sesuatu yang tidak tampak (*unobservable*) yang mendasari kinerja pada tes tersebut (Embretson & Reise, 2000). Variabel laten (misalnya kemampuan) dalam IRT disebut sebagai theta ( $\theta$ ) dimana semakin tinggi  $\theta$  yang dimiliki seseorang maka semakin tinggi pula probabilitasnya untuk menjawab soal dengan benar (Baker, 2001). Probabilitas seorang peserta tes Berpikir Kritis untuk menjawab benar sebuah butir merupakan fungsi dari tingkat kemampuan berpikir kritis yang dimiliki dengan tingkat ke-

sulitan dari butir tersebut. Berikut merupakan model matematis dari IRT dengan 2 parameter.

$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}}$$

$i = 1, 2, \dots, n$

Keterangan :

- $\theta$  : tingkat kemampuan peserta tes
- $P_i$  : probabilitas peserta tes yang memiliki kemampuan  $\theta$  menjawab benar soal  $i$
- $a_i$  : indeks daya pembeda soal ke  $i$
- $b_i$  : indeks kesulitan soal ke- $i$
- $e$  : bilangan natural yang nilainya mendekati 2,718
- $n$  : banyaknya soal tes
- $D$  : faktor penskalaan yang harganya 1,7.

Lebih lanjut, kelemahan utama dari teori tes klasik adalah hasil pengukurannya yang bersifat *group dependent* dan *item dependent*. Parameter tingkat kesulitan soal dalam teori tes klasik tergantung pada peserta yang menjawab soal tersebut. Artinya, sebuah soal akan menjadi soal yang mudah jika dijawab oleh peserta yang pandai, namun soal tersebut sekaligus menjadi soal yang sulit jika dijawab oleh peserta yang kurang pandai. Demikian pula pada parameter kemampuan peserta, seorang peserta akan menjadi pandai saat mengerjakan soal yang mudah namun sekaligus menjadi kurang pandai saat mengerjakan soal yang sulit. Dengan demikian, tidak pernah diketahui tingkat kemampuan yang sebenarnya dari peserta tes. Seperti halnya juga tidak diketahui apakah sebuah soal memiliki tingkat kesulitan yang tinggi atau rendah, karena tergantung pada kemampuan peserta yang menjawabnya. Pada IRT hal yang demikian tidak terjadi, karena kita benar-benar dapat mengukur tingkat kemampuan seseorang dengan lebih akurat, begitu pula dalam mengestimasi parameter soal/butir.

Berdasarkan uraian latar belakang yang telah dipaparkan, dapat disimpulkan dua hal yang menjadi dasar penelitian ini. Pertama, pentingnya konstruk berpikir praktis dalam memprediksi keberhasilan seseorang dalam

pendidikan ataupun pekerjaan memunculkan kebutuhan akan tersedianya tes berpikir praktis yang valid dan reliabel sehingga dapat memberikan hasil ukur yang akurat. Tes-tes berpikir kritis yang telah ada di Indonesia lebih banyak merupakan penerjemahan dari tes-tes yang beredar di Barat (Amerika dan Eropa) tanpa didahului pengujian tentang bias budayanya. Hal ini akan berdampak pada keakuratan hasil ukurnya. Kedua, pengembangan tes sendiri di Indonesia masih lebih banyak menggunakan pendekatan Teori Tes Klasik (*Classical Test Theory*) yang memiliki sejumlah kelemahan. Oleh karena itu, penelitian ini bermaksud menindaklanjuti kebutuhan tersebut dengan mengembangkan tes berpikir kritis yang valid dan reliabel dengan menggunakan pendekatan IRT. Tujuan penelitian ini adalah mengembangkan tes Berpikir Kritis yang valid dan reliabel untuk digunakan dalam lingkungan pendidikan dan kerja di Indonesia.

**Metode Penelitian**

Tipe penelitian ini adalah penelitian survei, yaitu pengumpulan data dari kelompok masyarakat, yang datanya kemudian dianalisis untuk mengembangkan tes berpikir kritis yang dapat digunakan oleh kelompok masyarakat tersebut.

Tahapan penelitian yang dilakukan berdasarkan tahap pengembangan tes menurut Hambleton dan Jones (1993), yaitu:

*Menyiapkan Spesifikasi Tes (Preparation of Test Specifications)*

Tahap pertama ini diawali dengan proses identifikasi perilaku yang merepresentasikan konstruk atau domain perilaku. Pada tahap ini, pengembang tes perlu menentukan satu atau lebih jenis perilaku yang diyakini merupakan manifestasi dari konstruk yang akan diukur, kemudian merumuskannya ke dalam bentuk butir yang dapat mengungkap perilaku tersebut.

Terdapat beberapa cara dalam menentukan domain perilaku dari sebuah konstruk, pada penelitian ini peneliti memilih

cara berdasarkan tinjauan terhadap hasil penelitian (*Review of research*) tentang tes berpikir kritis yang telah ada saat ini. Berdasarkan tinjauan terhadap berbagai penelitian di bidang tes berpikir kritis, peneliti memutuskan untuk menggunakan konsep yang digunakan dalam tes *Watson-Glaser Critical Thinking Appraisal* (WGCTA). Pada WGCTA, berpikir kritis terdiri dari lima dimensi yaitu *Inference, Recognition Assumption, Deduction, Interpretation* dan *Evaluation of arguments*. Berdasarkan kelima dimensi tersebut peneliti membuat *blue print* sebagaimana tertuang dalam Tabel 1.

Tabel 1. *Blue Print* Tes

Dimensi	Jumlah Butir
Kesimpulan ( <i>Inference</i> ): Kemampuan dalam menilai tingkat probabilitas ketepatan/kebenaran sebuah kesimpulan berdasarkan informasi yang tersedia.	30
Pengenalan asumsi ( <i>Recognition Assumption</i> ) Kemampuan dalam mengidentifikasi asumsi yang tersirat pada sebuah pernyataan	36
Deduksi ( <i>Deduction</i> ) Kemampuan dalam menentukan apakah kesimpulan dibuat secara logis berdasarkan informasi yang tersedia	20
Interpretasi ( <i>Interpretation</i> ) Kemampuan dalam menilai sebuah bukti ( <i>evidence</i> ) dan membuat keputusan apakah generalisasi / kesimpulan yang dihasilkan dijamin berdasarkan data yang tersedia	20
Evaluasi argumen ( <i>Evaluation of arguments</i> ) Kemampuan dalam mengevaluasi kekuatan dan relevansi sebuah argument terkait dengan sebuah isu atau masalah tertentu	30

*Menyiapkan Butir-Butir Tes (Preparation of The Test Item Pool)*

Peneliti menggunakan berbagai sumber literatur untuk mengembangkan skena-

rio yang akan digunakan sebagai batang soal. Selanjutnya berdasarkan masing-masing batang soal dibuat 4-6 buah butir. Jumlah butir yang dibuat sesuai dengan yang tertera di *blue print*. Setelah itu dilakukan review terhadap soal yang telah dibuat. Review dilakukan oleh para expert judges yang memahami tentang konsep tes berpikir kritis.

#### Uji Coba Butir Tes di Lapangan (*Field Testing The Items*)

Uji coba dilakukan untuk mengetahui apakah instruksi tes dapat dipahami dengan baik serta pernyataan dalam butir tidak memiliki pengertian yang ambigu. Kegiatan ini melibatkan 10 orang mahasiswa S1 Psikologi yang sedang mengikuti magang di Lembaga Pengkajian dan Psikologi Terapan (LP3T) Unair. Pemilihan subjek tersebut didasarkan pertimbangan bahwa mereka telah terbiasa mengenal dan memahami tes-tes Psikologi yang digunakan dalam kegiatan seleksi dan promosi karyawan.

#### Revisi Butir Tes (Revision of the Test Items)

Berdasarkan masukan dari 10 mahasiswa pada uji coba, dilakukan revisi yang terkait dengan administrasi dan kalimat pada butir. Pada tahap ini tes dibuat menjadi dua bagian yang paralel, sehingga masing-masing bagian tes hanya berisi separuh dari jumlah butir yang terdapat dalam setiap dimensi. Hal ini dilakukan agar peserta tes tidak merasa jenuh dan kehilangan konsentrasi karena disebabkan oleh banyaknya soal yang harus dikerjakan. Setelah dilakukan revisi, dilakukan pengujian di lingkungan internal Fakultas Psikologi dengan menggunakan subyek 30 mahasiswa S1 yang mengikuti mata kuliah Psikologi Industri dan Organisasi dan 38 mahasiswa peserta mata kuliah Metode Penelitian Kualitatif. Hasil dari uji coba ini dilakukan analisis dengan menggunakan *software* ITEMAN untuk mengetahui tingkat kesulitan, daya beda dan fungsi masing-masing distraktor. Berdasarkan hasil analisis tersebut dirakit tes berpikir kritis format final yang hanya berisi butir dengan kualitas yang baik.

#### Pengembangan Tes (*test development*)

Proses berikutnya adalah melakukan pengambilan data di lapangan. Data diperoleh dengan menggunakan subjek yang mengikuti tes seleksi karyawan Semen Gresik dan Angkasa Pura. Jumlah seluruh subjek adalah 1.453 orang.

Lebih lanjut, analisis data dilakukan dengan pendekatan Item Response Theory, menggunakan program Mplus versi 6.00. Hasil dari penelitian ini berupa dokumen berupa hasil analisis butir yang akan memberikan estimasi parameter butir yaitu tingkat kesulitan dan daya beda butir. Hasil analisis juga akan memberikan data tentang tingkat kemampuan berpikir kritis dari masing-masing subjek.

#### Hasil Penelitian dan Pembahasan

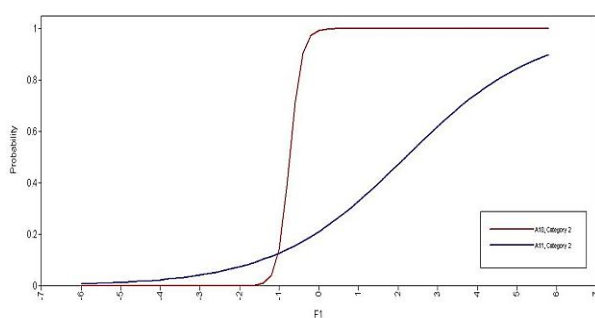
Sebelum melakukan analisis butir dengan menggunakan IRT, dilakukan uji asumsi yaitu unidimensionalitas, independensi lokal dan *Item Characteristic Curve* (ICC) (Hambleton dkk., 1991; Baker, 2001). Pengujian asumsi unidimensionalitas dilakukan dengan menggunakan CFA satu faktor dengan metode estimasi yaitu *robust weighted least squares* (WLS/WLSMV) yang sesuai dengan jenis data kategori/dikotomi. Pengujian dilakukan dengan *software* statistic Mplus versi 6.11.

Berdasarkan hasil pengujian CFA, diketahui bahwa subtes 1 sesuai dengan model satu faktor dengan nilai RMSEA = 0,026. Subtes 2 juga sesuai dengan model satu faktor dengan nilai RMSEA = 0,048. Subtes 3, datanya sesuai dengan model satu faktor dengan nilai RMSEA = 0,070. Subtes 4 juga sesuai dengan model satu faktor dengan nilai RMSEA = 0,054. Subtes 5 tidak sesuai dengan model satu faktor dengan nilai RMSEA = 0,111. Menurut MacCallum (dalam Hooper dkk., 2008) nilai RMSEA < 0,08 menunjukkan model yang *good fit*. Nilai RMSEA > 0.1 menunjukkan model yang *poor fit*. Artinya, dalam hal ini model subtes 5 tidak sesuai dengan data, sehingga subtes 5 bukan merupakan model satu faktor.

Hambleton dkk. (1991) mengemukakan bahwa jika asumsi unidimensionalitas dipenuhi maka asumsi independensi lokal juga terpenuhi. Berdasarkan hal tersebut maka keempat subtes pada tes berpikir kritis memenuhi asumsi independensi lokal.

Analisis butir tes kemampuan berpikir kritis dilakukan dengan menggunakan *software* statistik Mplus versi 6.11 Metode estimator yang digunakan dalam analisis ini adalah *Robust Maximum Likelihood* (MLR). Berikut adalah hasil dari analisis tersebut yang dilakukan pada setiap subtes.

Berdasarkan hasil analisis pada subtes 1 (kemampuan menilai ketepatan sebuah kesimpulan) diketahui bahwa hanya terdapat dua butir dengan daya beda yang lebih besar dari 0,3 yaitu butir nomor 10 dan 11. Kedua butir tersebut memiliki daya beda yang cukup baik, artinya kedua butir tersebut dapat membedakan antara individu yang memiliki kemampuan yang rendah dengan yang tinggi pada kondisi dimana kemampuan (*theta*) sama dengan tingkat kesulitan butir (b). Dengan demikian, kedua butir tersebut mampu membedakan antara individu yang memiliki kemampuan tinggi dalam menilai ketepatan sebuah kesimpulan dengan individu yang berkemampuan rendah. ICC dari butir nomor 10 dan 11 disajikan dalam grafik pada Gambar 1.



Gambar 1. ICC Subtes 1

Nilai daya beda yang rendah ditunjukkan dengan bentuk ICC yang datar, hal tersebut menunjukkan bahwa probabilitas menjawab benar antara individu dengan kemampuan yang tinggi dengan individu yang berkemampuan rendah adalah sama (Baker, 2001).

Nilai daya beda yang bertanda positif menunjukkan bahwa probabilitas menjawab benar butir tersebut meningkat seiring dengan meningkatnya taraf kemampuan (Baker, 2001). Sebaliknya, daya beda yang bernilai negatif menunjukkan bahwa probabilitas menjawab benar butir tersebut, menurun seiring dengan meningkatnya kemampuan dari tingkat yang rendah ke tingkat yang lebih tinggi (Baker, 2001). Butir yang memiliki daya beda negatif merupakan butir yang buruk. Hal tersebut dapat terjadi karena kesalahan pada saat penulisan butir, dimana terdapat kesalahan informasi yang membuat individu dengan kemampuan yang tinggi gagal memahami maksud dari butir tersebut sehingga salah dalam menjawab butir (Baker, 2001).

Tabel 2. Hasil Estimasi Parameter Butir Subtes 1

No. Butir	Daya Beda (a)	Tingkat Kesulitan (b)
1.	0,013	3,145
2.	-0,023	0,458
3.	-0,009	1,385
4.	0,103	1,345
5.	0,094	-0,200
6.	0,065	1,555
7.	-0,271	1,570
8.	0,042	1,067
9.	-0,054	2,430
10.	3,981	-4,959
11.	0,355	1,322
12.	0,176	-0,942
13.	-0,074	2,104
14.	-0,074	0,435
15.	0,063	1,065

Berdasarkan tingkat kesulitan butir dapat diketahui bahwa butir nomor 1 merupakan butir dengan tingkat kesulitan tertinggi sedangkan butir nomor 10 merupakan butir yang paling mudah. Tingkat kesulitan butir umumnya berada pada rentang -3 sampai dengan 3. Nilai kesulitan butir yang berada di luar rentang tersebut menunjukkan bahwa butir tersebut bermasalah yang mungkin terkait dengan isi maupun

penulisan butir (Sudol & Studer, 2010). Tingkat kesulitan item yang berada di luar rentang -3 sampai dengan 3 tidak sesuai dengan rentang kemampuan peserta yang umumnya berada pada rentang -3 sampai dengan 3 dengan nilai rata-rata 0 dan varians 1.

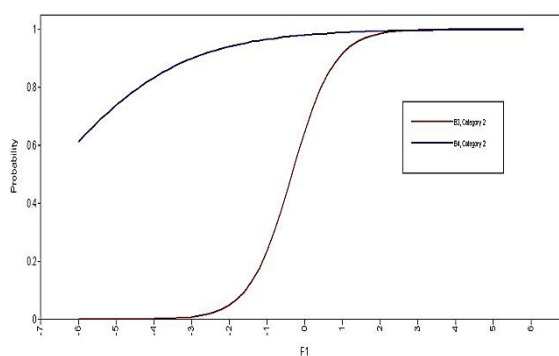
Tabel 3. Hasil Estimasi Parameter Butir Subtes 2

No. Butir	Daya Beda (a)	Tingkat Kesulitan (b)
1.	-1,281	0,082
2.	-0,577	2,448
3.	1,054	-0,609
4.	0,340	-3,920
5.	-0,150	-1,918
6.	-0,026	1,010
7.	0,065	0,032
8.	0,124	-0,379
9.	0,150	-0,456
10.	0,141	-0,666
11.	-0,077	-1,931
12.	-0,166	0,247
13.	0,091	0,673
14.	0,118	2,590
15.	0,001	-1,615
16.	-0,014	-0,673
17.	-0,048	-0,268
18.	0,119	0,214

Berdasarkan hasil analisis pada subtes 2 (kemampuan mengenali asumsi) diketahui bahwa hanya terdapat dua butir yang memiliki daya beda > 0,3, yaitu butir nomor 3 dan 4. Hal ini menunjukkan bahwa kedua butir tersebut yang dapat dengan cukup baik membedakan probabilitas menjawab benar antara individu yang berkemampuan dalam mengenali asumsi dengan individu yang berkemampuan tinggi. ICC dari butir nomor 3 dan 4 yang disajikan pada Gambar 2.

Dari 18 butir dalam subtes 2, terdapat 8 butir yang memiliki daya beda bertanda negatif. Hal tersebut mengindikasikan terjadi kesalahan dalam penulisan butir yang membuat individu yang berkemampuan tinggi dalam mengenali asumsi, gagal me-

mahami maksud dari butir tersebut. Butir tersulit terdapat pada butir no 2 dengan tingkat kesulitan 2,448 artinya hanya individu dengan kemampuan minimal sebesar 2,448 yang memiliki 50% peluang untuk menjawab butir dengan benar. Butir yang termudah adalah butir no 4 dengan nilai tingkat kesulitan -3,920, artinya peserta dengan kemampuan dalam mengenali asumsi sebesar -3,920 berpeluang 50% menjawab benar butir tersebut.



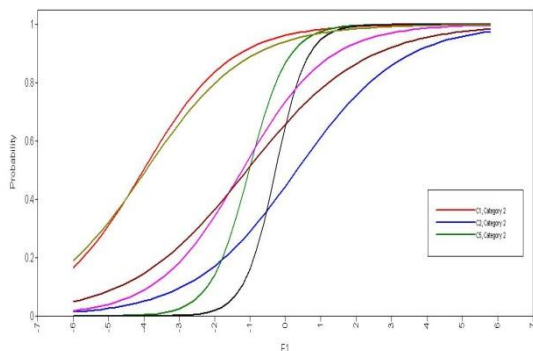
Gambar 2. ICC Subtes 2

Tabel 4. Hasil Estimasi Parameter Butir Subtes 3

No. Butir	Daya Beda (a)	Tingkat Kesulitan (b)
1.	0,477	-3,254
2.	0,398	0,220
3.	-0,276	0,479
4.	-0,376	2,164
5.	1,086	-1,888
6.	1,340	-0,630
7.	0,493	-1,023
8.	0,355	-0,656
9.	0,414	-2,777
10.	0,038	-2,063

Berdasarkan hasil analisis terhadap subtes 3 (deduksi: kemampuan menilai kelogisan sebuah kesimpulan) diketahui bahwa terdapat 7 butir yang memiliki nilai daya beda lebih besar dari 0,3. Hal tersebut menunjukkan bahwa sebagian besar butir pada subtes 3 dapat membedakan antara individu dengan kemampuan deduksi yang tinggi dengan yang rendah. Ketujuh butir tersebut memiliki bentuk ICC dengan kemiringan yang cenderung curam. Hal ini me-

nunjukkan bahwa butir tersebut memiliki daya beda yang cukup baik sehingga dapat menggambarkan individu yang berkemampuan tinggi memiliki peluang menjawab benar yang lebih besar daripada yang berkemampuan rendah. ICC dari butir nomor 1, 2, 5, 6, 7, 8, 9 disajikan pada Gambar 3.



Gambar 3. ICC Subtes 3

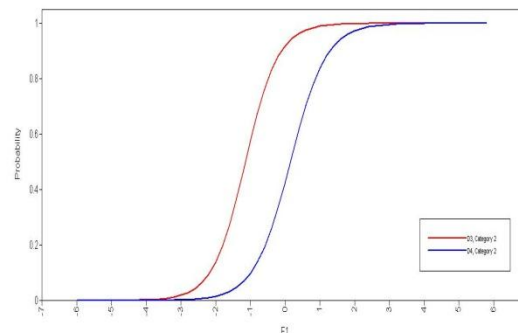
Pada subtes 3 terdapat dua butir yang memiliki daya beda bertanda negatif, yaitu butir nomor 3 dan 4. Hal ini mengindikasikan terjadinya kesalahan dalam penulisan butir, sehingga gagal dipahami oleh individu dengan kemampuan yang tinggi. Butir nomor 4 merupakan butir dengan tingkat kesulitan paling tinggi sedangkan butir yang termudah adalah butir nomor 1.

Tabel 5. Hasil Estimasi Parameter Butir Subtes 4

No. Butir	Daya Beda (a)	Tingkat Kesulitan (b)
1.	0,101	1,761
2.	0,258	0,583
3.	1,256	-2,435
4.	1,139	0,290
5.	-0,120	0,179
6.	-0,111	0,344
7.	-0,176	0,049
8.	0,013	-1,581
9.	-0,051	1,162
10.	-0,089	-1,177

Pada subtes 4 (interpretasi: kemampuan menilai sebuah *evidence*) hanya terdapat dua butir yang memiliki daya beda yang cukup baik ( $> 0,3$ ) yaitu butir nomor 3 dan 4. Artinya, kedua butir ini dapat membeda-

kan antara individu yang berkemampuan tinggi dengan yang berkemampuan rendah didalam kemampuan menilai apakah bukti yang digunakan dalam penarikan kesimpulan telah sesuai dengan data yang tersedia. ICC untuk butir nomor 3 dan 4 disajikan pada Gambar 4.



Gambar 4. ICC Subtes 4

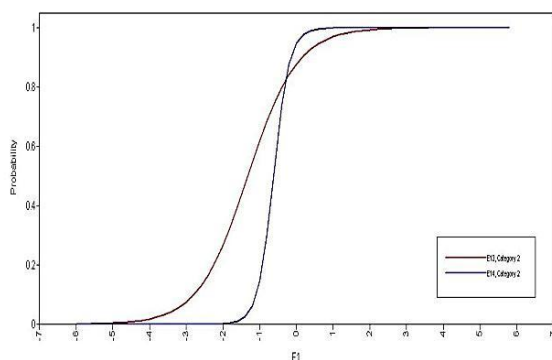
Terdapat 5 butir yang memiliki daya beda dengan nilai negatif. Kelima butir ini tergolong butir yang buruk sehingga individu dengan kemampuan yang tinggi gagal memahami maksud dari butir tersebut. Butir yang tersulit terdapat pada butir nomor 1 sedangkan yang termudah terdapat pada butir nomor 3.

Tabel 6. Hasil Estimasi Parameter Butir Subtes 5

No. Butir	Daya Beda (a)	Tingkat Kesulitan (b)
1.	-0,058	1,037
2.	-0,103	-0,453
3.	0,056	-0,752
4.	-0,069	0,221
5.	0,135	-1,188
6.	0,234	-0,666
7.	0,058	-0,251
8.	-0,011	-0,079
9.	0,220	0,018
10.	-0,043	-0,318
11.	0,143	-0,274
12.	0,120	-0,096
13.	0,880	-1,975
14.	2,764	-2,904
15.	-0,495	-0,544

Pada subtes 5 (kemampuan mengevaluasi argumen) hanya terdapat dua butir

yang memiliki daya beda yang tinggi, yaitu butir nomor 13 dan 14. Kedua butir ini mampu dengan baik membedakan individu yang berkemampuan tinggi dalam menilai sebuah argumen dengan individu yang berkemampuan rendah. Dengan demikian, ICC pada kedua butir ini akan memiliki kecuraman tinggi sehingga menunjukkan perbedaan bahwa individu yang berkemampuan tinggi akan memiliki peluang menjawab benar yang lebih besar daripada individu yang berkemampuan rendah. ICC dari butir nomor 13 dan 14 disajikan pada Gambar 5.



Gambar 5. ICC Subtes 5

Terdapat 6 butir yang memiliki daya beda negatif, artinya terjadi kesalahan penulisan yang membuat kesalahan informasi sehingga individu yang berkemampuan tinggi gagal memahami butir. Kesulitan tertinggi terdapat pada butir nomor 1 dan butir yang termudah adalah butir nomor 14. Revisi terhadap butir-butir dengan nilai daya beda yang rendah maupun negatif akan dilakukan pada tahap pengembangan tes berikutnya.

Berdasarkan hasil analisis dari kelima subtes dalam tes kemampuan berpikir kritis, diperoleh data bahwa dari total 68 butir hanya terdapat 15 butir yang memiliki daya beda lebih besar dari 0.3. Hal ini menunjukkan bahwa ke 15 butir tersebut dapat membedakan probabilitas menjawab benar antara individu yang berkemampuan tinggi dengan yang berkemampuan rendah.

Dari keseluruhan butir terdapat 27 butir dengan daya beda yang negatif. Hal ini menunjukkan bahwa banyak terjadi kesalahan dalam penulisan butir sehingga terjadi kesalahan dalam memahami informasi pada

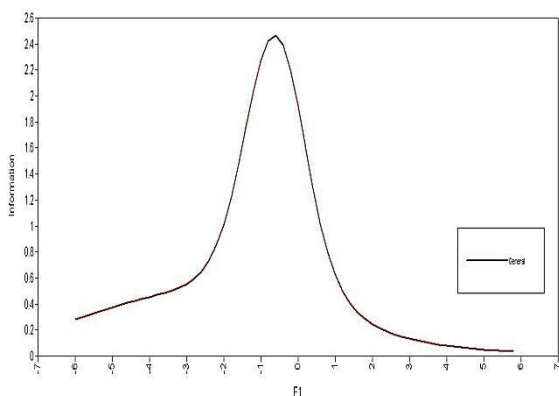
individu yang memiliki kemampuan berpikir kritis tinggi. Banyaknya butir yang berdaya beda rendah serta bertanda negatif menunjukkan bahwa terdapat banyak kesalahan dalam proses penulisan butir. Hal ini dapat terjadi karena pemilihan topik yang digunakan sebagai skenario pada batang soal kemungkinan besar bukanlah topik yang dapat digunakan untuk mengukur kemampuan berpikir kritis.

Pemilihan topik pada penelitian ini telah dilakukan oleh para ahli yaitu para dosen sekaligus Psikolog di bidang Industri dan Organisasi serta Psikolog di bidang Pendidikan dan Perkembangan. Pemilihan pakar tersebut mungkin saja kurang tepat karena belum tentu menguasai bidang berpikir kritis dengan baik. Artinya, terjadi kesalahan dalam menentukan *Subject Matter Experts* di bidang berpikir kritis (Possin, 2014).

Kesalahan lainnya dimungkinkan terjadi dalam menentukan jawaban benar. Soal-soal yang terdapat pada tes kemampuan berpikir kritis berbeda dengan soal-soal yang terdapat pada tes objektif misalnya di bidang matematika, ilmu alam maupun pengetahuan lain yang jawaban benar dari soal tersebut telah diketahui dengan pasti. Soal pada tes yang sejenis kemampuan berpikir kritis memiliki cara skoring yang berbeda. Pada jenis tes yang tidak memiliki cara skoring baku terdapat tiga pilihan cara skoring, yaitu (a) *target-scoring*, (b) *expert-scoring*, (c) *consensus-scoring* (MacCann, 2006). Dalam penelitian ini digunakan cara skoring dengan metode *expert-scoring*. Kondisi ini memungkinkan terjadinya kesalahan karena prosedur kesalahan dalam menentukan *Subject Matter Experts* di bidang berpikir kritis. Kesalahan-kesalahan ini akan diperbaiki pada proses penelitian selanjutnya yaitu pada tahap penelitian tahun kedua.

Seperti yang tampak pada gambar 6, dapat diketahui bahwa nilai fungsi informasi Tes Berpikir Kritis akan maksimum pada tingkat kemampuan  $-0,5$  yaitu sebesar 2,4. Berdasarkan grafik di atas juga dapat diketahui bahwa tes Berpikir kritis yang dikembangkan dapat mengestimasi dengan

tepat tingkat kemampuan yang berada pada rentang -3 sampai dengan 1.



Gambar 6. Test Information Function

### Simpulan

Meskipun belum optimal dan masih terdapat sejumlah catatan yang perlu diperbaiki pada penelitian berikutnya, namun penelitian ini telah menghasilkan tes berpikir kritis yang dapat digunakan dalam proses diagnosis maupun seleksi dalam lingkungan akademik dan kerja. Terkait proses, setiap tahapan dalam pengembangan tes juga telah dilakukan sesuai dengan yang dikemukakan oleh Hambleton dan Jones (1993).

Penelitian lanjutan di tahun kedua sangat diperlukan untuk memperbaiki hasil yang telah diperoleh dengan memperhatikan berbagai catatan yang telah diberikan pada bagian pembahasan

### Daftar Pustaka

- Angelo, T.A. & Cross, P. (1995). *Classroom assessment techniques: a Handbook for college teachers* (2<sup>nd</sup> edition). San Francisco: Jossey-Bass Publishers.
- Baker, F. B. (2001). *The basics of butir response Theory* (2<sup>nd</sup> edition). USA: ERIC Clearinghouse on Assessment and Evaluation.
- Embretson, S. E., & Reise, S. P. (2000). *Butir response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of butir*

*response theory*. Newbury Park, CA: Sage Publications.

- Hambleton, R. K., & Jones, R. W. (1993). An NCME instructional module on: Comparison of classical test theory and butir response theory and their applications to test development. *Educational Measurement: Issues and Practice*, 12(3). 38-47.
- Hossoubah, Z. (2007). *Developing creative and critical thinking skills (terjemahan)*. Bandung: Yayasan Nuansa Cendia.
- Kaupp, J., Frank, B., & Chen, A. (2014). *Evaluating critical thinking and problem solving in large classes: Model eliciting activities for critical thinking development*. Toronto: Higher Education Quality Council of Ontario.
- Kurfiss, J. G. (1988). *Critical thinking: Theory, research, practice, and possibilities*. (ASHE-ERIC Higher Education Report No. 2). Association for the Study of Higher Education. <http://www.eric.ed.gov/PDFS/ED304041.pdf> (diunduh 14 April 2014).
- Lai, E. R. (2011). *Critical thinking: a Literature review*. Author, Pearson Assessments. <http://www.pearsonassessments.com/hai/images/tmrs/CriticalThinking> (diunduh 14 April 2014).
- MacCann, C.E. (2006). New approaches to measuring emotional intelligence: Exploring methodological issues with two new assessment tools. *Doctoral dissertation*. Sydney: University of Sydney.
- Norris, S. P. & Ennis, R. H. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Midwest Publications.
- Possin, K. (2014). Critique of the Watson-Glaser critical thinking appraisal test: The more you know, the lower your score. *Informal Logic*, Vol 34, No 4.
- Stassen, M.L.A., Herrington, A., & Henderson, L. (2011). *To improve the academy*. San Francisco, CA: Jossey-Bass Publishers.

- Sternberg, R. J. (1986). *Critical thinking: Its nature, measurement, and improvement*. National Institute of Education. <http://eric.ed.gov/PDFS/ED272882.pdf> (diunduh 14 April 2014).
- Essential Skills for the 21st Century Workplace: Keys to Succeeding in the Global Economy. [www.21stcenturyskills.org](http://www.21stcenturyskills.org) (diunduh 14 April 2014)
- Wagner, T. A., (2002). Critical thinking: The development of a new measure. *Thesis*. Blacksburg, Virginia: Virginia Polytechnic Institute and State University.
- Walker, P. & Finney, N. (1999). *Skill development and critical thinking in higher education*. London: Higher Education Research & Development Unit, University College.

## PENGEMBANGAN MODEL EVALUASI PENGELOLAAN PONDOK PESANTREN

<sup>1)Khuriyah, 2)Zamroni, 3)Sumarno</sup>

<sup>1)IAIN Surakarta, 2)Universitas Negeri Yogyakarta, 3)Universitas Negeri Yogyakarta</sup>

<sup>1)khuriyah\_stainslo@yahoo.co.id, 2)zamronihardjowiriono@yahoo.com, 3)sumarno\_unj@yahoo.co.uk</sup>

### Abstrak

Penelitian ini bertujuan untuk menghasilkan pedoman evaluasi IBSQ dan mendeskripsikan: (1) kriteria model evaluasi pengelolaan pondok pesantren yang baik; (2) efektivitas penggunaan model evaluasi pengelolaan pondok pesantren; dan (3) pengelolaan pondok pesantren. Metode penelitian yang digunakan yaitu *research and development* (R & D). Subjek penelitian yaitu pimpinan, pengelola, ustaz, santri dan alumni Pondok Pesantren Al Falah Sidoharjo Sragen, Pondok Pesantren Assalaam, dan Pondok Pesantren Al Muayyad Surakarta. Hasil penelitian menunjukkan bahwa telah dihasilkan sebuah model evaluasi yang dapat mengevaluasi pengelolaan pondok pesantren yang terdiri dari lima buah buku dengan kriteria: (1) karakteristik instrumen dalam model IBSQ: (a) memiliki format yang baik; (b) memenuhi substansi model evaluasi; (c) memiliki tingkat validitas konstruk yang dapat diandalkan; (d) memiliki tingkat reliabilitas yang tinggi; (2) model IBSQ memiliki tingkat efektivitas yang baik; (3) pengelolaan setiap pondok pesantren secara nyata terjadi perbedaan: (a) pengelolaan di Pondok Pesantren Al Falah masih sangat tradisional dengan bukti kegiatan administrasi masih manual; (b) pengelolaan di Pondok Pesantren Assalaam sudah memanfaatkan IT sebagai sarana dalam menjalankan pengelolaan; (c) pengelolaan di Pondok Pesantren Al Muayyad masih campuran antara manual dan berbasis komputer.

**Kata kunci:** *model evaluasi, pengelolaan, pondok pesantren*

## DEVELOPING THE EVALUATION MODEL OF ISLAMIC BOARDING SCHOOL MANAGEMENT

<sup>1)Khuriyah, 2)Zamroni, 3)Sumarno</sup>

<sup>1)IAIN Surakarta, 2)Universitas Negeri Yogyakarta, 3)Universitas Negeri Yogyakarta</sup>

<sup>1)khuriyah\_stainslo@yahoo.co.id, 2)zamronihardjowiriono@yahoo.com, 3)sumarno\_unj@yahoo.co.uk</sup>

### Abstract

The study was to generate the IBSQ evaluation guidelines and to describe: (1) the criteria of good evaluation model for the Islamic Boarding School management; (2) the effectiveness of using the evaluation model for the Islamic Boarding School management; and (3) the management of Islamic Boarding School. The method that was used was the research and development (R&D). The subjects in the study were the leaders, the managers, the *ustaz*, the students and the alumni of Al Falah Islamic Boarding School Al Falah Sidoharjo Sragen, of Assalaam Islamic Boarding School and of Al Muayyad Surakarta Islamic Boarding School. The result of the study showed that there had been an evaluation model that might evaluate the Islamic boarding school management and the evaluation model consisted of five books with the following criteria: (1) the IBSQ evaluation model had the following characteristic instruments: (a) having a good format; (b) meeting the substances of an evaluation model; (c) having reliable level of construct validity; and (d) having high level of reliability; (2) the IBSQ evaluation model had a good level of effectiveness; (3) there had been actual differences in each Islamic boarding school: (a) the management in the Al Falah Islamic Boarding School was still very traditional. It was indicated by manual administration; (b) the management in Assalaam Islamic Boarding School had benefitted the information technology as the tools in performing managerial tasks; and (c) the management in the Al Muayyad Islamic Boarding School used a combination between the manual and the computer-based way.

**Keyword:** *evaluation model, management, Islamic boarding school*

## Pendahuluan

Pondok pesantren merupakan lembaga pendidikan tertua dalam perjalanan kehidupan Indonesia sejak enam abad yang lalu hingga sekarang. Pondok pesantren di Indonesia dikenal sebagai tempat belajar mengajar yang intensif dan paling sesuai dengan kultur masyarakat Islam Indonesia. Pendidikan dan pengajaran di pesantren berurat akar ke bawah, mendapatkan dukungan dari masyarakat, dan hidup di tengah masyarakat serta mengabdikan pada kepentingan rakyat.

Pondok pesantren memiliki ciri khas dibandingkan dengan lembaga pendidikan lain di Indonesia. Pondok pesantren merupakan lembaga pendidikan yang mengkhususkan pada pendalaman ilmu keagamaan. Dalam perkembangannya, kini terjadi banyak perubahan mulai dari segi keilmuan (dengan mulai mengadopsi mata pelajaran umum), juga pada kapasitas-kapasitas lain kelembagaan pesantren yang kini menjadi sebuah institusi yang memiliki kelengkapan fasilitas untuk membangun potensi santri tidak hanya pada segi akhlak, nilai, intelektual, dan spiritualnya, tetapi juga pada peralatan yang ada di dalamnya. Hal ini sejalan dengan tujuan pokok didirikannya pondok pesantren adalah mencetak ulama atau mencetak orang yang mendalami ilmu agamanya atau *tafaqqub fi ad-diin* (Nafi', et al., 2007, p.5).

Pondok pesantren merupakan lembaga pendidikan yang paling variatif mengingat adanya kebebasan dari pendirinya untuk mewarnai pesantrennya dengan penekanan pada kajian tertentu. Penekanan ini menurut Qomar (2007, p.58) didasarkan atas keahlian dari kiai pengasuhnya. Banyaknya variasi pesantren tersebut perlu diadakan pembedaan secara kategorial. Kategori ini dapat ditinjau dari berbagai perspektif, segi rangkaian kurikulum, tingkat kemajuan dan kemodernan, keterbukaan terhadap perubahan dan dari sudut sistem pendidikannya.

Menurut Arifin (1991, p.251) dari segi kurikulum, pesantren dapat digolongkan menjadi pesantren modern, pesantren ta-

hassus dan pesantren campuran. Menurut Dhofier (1985, p.41) dari perspektif keterbukaan terhadap perubahan-perubahan yang terjadi, pesantren dibagi menjadi pesantren salafi dan khalafi. Pesantren salafi bersifat konservatif, sedangkan pesantren *khalafiy* bersifat adaptif (Qomar, 2007).

Lebih lanjut dijelaskan oleh Qomar (2007) perbedaan pesantren tradisional dan modern dapat diidentifikasi dari perspektif manajerialnya. Pesantren modern telah dikelola dengan rapi dan sistematis dengan kaidah manajerial yang umum. Sementara itu, pesantren tradisional berjalan secara alami tanpa berupaya melakukan pengelolaan secara efektif.

Pesantren salafiyah merupakan bentuk lembaga pendidikan Islam tertua di Indonesia. Keberadaannya mengiringi kehadiran Islam sebagai salah satu alat berdakwah yang efektif dalam *menggembeng* santri agar memiliki pengetahuan agama yang mapan sehingga nantinya dapat mengajarkan pada orang lain. Hanya saja, usia pesantren yang sangat tua tidak dibarengi dengan kemajuan manajemennya. Masih banyak kondisi manajemen pesantren tradisional yang sangat memprihatinkan. Atau dengan kata lain pola manajemen pesantren tradisional cenderung dilakukan secara insidental dan kurang memperhatikan tujuan-tujuan yang telah disistematisasikan secara *hierarkis* (Rahim, 2001, p.154).

Ditinjau dari perspektif manajerial, landasan tradisi dalam mengelola suatu lembaga, termasuk pesantren menyebabkan produk pengelolaan itu asal jadi, tidak memiliki fokus strategi yang terarah, dominasi personal terlalu besar dan cenderung eksklusif dalam pengembangannya. Pada segi pendidikan, pesantren salafiyah pun belum ada standar kurikulum dan pengawasan mutu.

Selain itu pendidikan pesantren salafiyah pun masih memerlukan kajian dari sistem belajarnya dan cakupan materi ajar, karena tidak ada standar kurikulum dan pengawasan mutu. Di bidang agama pesantren belum tentu dapat menjamin lulusan pesantren menjadi ulama, sebaliknya di bidang umum

pun pengetahuan yang dimiliki santri sangat minim. Asumsi ini muncul karena ada anggapan bahwa ilmu-ilmu keduniaan tidak terlalu penting karena tidak dibawa ke akhirat. Kesan seperti ini terdengar eksklusif, namun ini adalah realita secara umum pada pesantren *salafiah*.

Berdasarkan temuan-temuan tersebut, untuk mewujudkan pondok pesantren yang memiliki pola manajemen yang baik perlu diawali dengan pelaksanaan evaluasi. Evaluasi pengelolaan pondok pesantren sangat penting dilakukan untuk memperoleh satu kriteria tentang mekanisme pengelolaan pondok pesantren yang terstandar.

Terkait dengan hal tersebut, penelitian ini bertujuan untuk menghasilkan sebuah model evaluasi pengelolaan pondok pesantren yang tepat untuk mengevaluasi semua tipe pondok pesantren dan mendeskripsikan: (1) kriteria model evaluasi pengelolaan pondok pesantren yang baik; (2) efektivitas penggunaan model evaluasi pengelolaan pondok pesantren; (3) pengelolaan di pondok pesantren.

## Metode Penelitian

Penelitian ini dikategorikan sebagai penelitian pengembangan dengan mengkombinasikan model evaluasi CIPP dan model *goal free evaluation*. Pemilihan model CIPP dan *goal free evaluation* didasarkan beberapa alasan: (1) model CIPP merupakan sebuah model evaluasi yang bertujuan untuk memperbaiki atau menyempurnakan program yang berjalan di pondok pesantren; (2) model *goal free evaluation* merupakan sebuah model evaluasi yang bertujuan untuk membuktikan keberlangsungan program. Oleh karena itu, peneliti tidak perlu terlebih dahulu melihat tujuan dari program tersebut. Hal ini dimaksudkan agar data temuan tidak bias; (3) melalui model *goal free evaluation*, data yang diperoleh evaluator dapat dijadikan sebagai pelengkap dari data yang diambil menggunakan model lain.

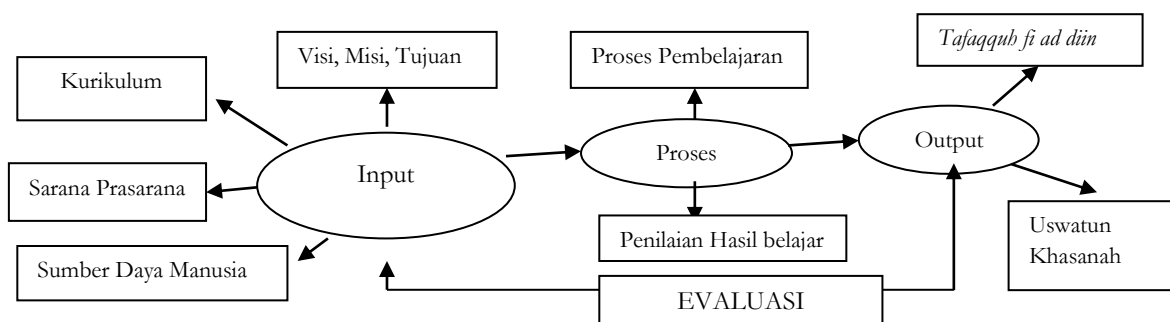
Borg & Gall (1989, p.781) menyatakan *research and development* terdiri dari suatu siklus untuk mengembangkan suatu produk, diuji lapangan, dan direvisi berdasarkan data

uji lapangan. Produk yang dihasilkan dalam penelitian ini adalah suatu model evaluasi yang tepat untuk mengevaluasi pengelolaan pondok pesantren yang diberi nama *Islamic Boarding School for Quality* atau IBSQ. Siklus ini terdiri dari langkah-langkah sebagai berikut: (1) pengumpulan informasi; (2) perencanaan; (3) mengembangkan produk awal; (4) pengujian lapangan awal; (5) revisi pada produk utama; (6) pengujian lapangan utama; (7) revisi produk operasional; (8) pengujian lapangan operasional; (9) revisi produk akhir; (10) diseminasi dan distribusi.

Selanjutnya, sepuluh langkah tersebut diringkas menjadi empat tahap yaitu sebagai berikut. Pertama, kegiatan melakukan studi pendahuluan, yaitu kegiatan pengumpulan informasi baik secara literatur maupun kajian lapangan mengenai pengelolaan yang berlangsung di pondok pesantren. Kedua adalah penyusunan draf instrumen evaluasi pengelolaan pondok pesantren. Sebagaimana telah dikemukakan, bahwa sampai saat ini belum terdapat instrumen evaluasi yang tepat digunakan untuk mengevaluasi pengelolaan pondok pesantren. Oleh karena itu, dalam penyusunan draf instrumen ini didasarkan atas hasil observasi yang peneliti lakukan. kemudian dilakukan uji kelayakan dilihat dari keterbacaannya.

Ketiga, mengadakan uji coba, meliputi uji coba terbatas yang dilakukan di Pondok Pesantren Wali Songo Sragen dan Pondok Pesantren Al Fatah Kartasura. Uji coba lebih luas dilakukan di Pondok Pesantren Al Fatah Sragen, Pondok Pesantren Assalaam Surakarta, dan Pondok Pesantren Al Muayyad Surakarta. Keempat adalah revisi produk akhir.

Penelitian ini memiliki tiga komponen yang dievaluasi, yakni komponen *input* pondok pesantren, kualitas proses dan *output* alumni. Komponen *input* berhubungan dengan evaluasi konsep pondok pesantren (visi, misi, tujuan), kompetensi kiai, kompetensi ustaz, dan sarana prasarana. Komponen kualitas proses meliputi evaluasi pembelajaran dan evaluasi penilaian. Komponen *output* membahas tentang kompetensi alumni pondok pesantren.



Gambar 1. *Hypothetic Construct model IBSQ*

Gambar 1 menunjukkan bahwa pengelolaan di pondok pesantren meliputi aspek *input*, proses dan *output*. Aspek input merupakan segala sesuatu yang menjadi dasar pelaksanaan pengelolaan pondok pesantren meliputi visi, misi, tujuan, kurikulum, sarana prasarana maupun sumber daya manusia yang terdiri dari ustaz, santri bahkan seorang kiai. Aspek lain yang terkait dalam pengelolaan sebuah lembaga pendidikan adalah proses. Aspek proses di sini ditekankan hanya pada kegiatan proses pembelajaran dan penilaian hasil belajar. Dengan melihat karakteristik ustaz, santri dan sarana prasarana yang tersedia akan berdampak pada proses pembelajaran yang berlangsung. Begitu juga dengan penilaian. Semua pondok pesantren menerapkan sistem penilaian untuk mengetahui kemampuan santri dalam memahami materi. Hanya saja mekanismenya yang sedikit berbeda. Aspek berikutnya adalah *output*. *Output* santri setelah belajar dari pondok pesantren adalah menjadi seorang santri yang memiliki kemampuan memahami ilmu agama dengan baik (*tafaqquh fi addin*) dan berakhlak *karimah*.

Penelitian ini dilakukan di wilayah eks Karesidenan Surakarta. Mengingat keterbatasan peneliti, maka tidak seluruh pondok pesantren diteliti. Pondok pesantren yang diteliti adalah Pondok Pesantren Al Falah Sidoharjo Sragen, Pondok Pesantren Assalaam Surakarta, dan Pondok Pesantren Al Muayyad Surakarta. Penentuan pondok pesantren didasarkan atas karakteristik masing-masing pondok pesantren yang secara umum sama antara wilayah satu dan yang lainnya. Unit analisis dalam penelitian ini adalah penyelenggara Sekolah Dasar Islam

terpadu dengan subjek penelitian sebagai berikut: (1) kiai, kepala sekolah, dan ustaz sebagai penyelenggara pondok pesantren. Dalam penelitian ini berperan sebagai narasumber dalam pengembangan model IBSQ; (2) santri sebagai partisipan dalam kegiatan pembelajaran juga berperan sebagai narasumber dalam pengembangan model IBSQ; (3) alumni pondok pesantren berperan sebagai narasumber dalam pengembangan model IBSQ; (4) pakar dan praktisi evaluasi, pengelola dan peneliti pondok pesantren berperan sebagai penilai instrumen model IBSQ dan narasumber dalam pengembangan model IBSQ.

Jumlah keseluruhan responden dari pondok pesantren sebanyak 181 orang. Metode pengumpulan data dalam instrumen model evaluasi yang dikembangkan dalam penelitian ini antara lain angket, dokumentasi, dan interview. Pencarian data menggunakan angket dalam penelitian ini dilaksanakan untuk menggali pandangan, persepsi, maupun sikap responden mengenai indikator yang ada dalam model evaluasi pengelolaan pondok pesantren. Dokumentasi digunakan untuk menggali data yang berhubungan dengan data pondok pesantren di Jawa Tengah dengan kategori tradisional, kombinasi dan modern, serta data lain yang meliputi jumlah santri, ustaz, dan data lain yang berkaitan dengan penelitian. Sedangkan interview digunakan untuk memperoleh data langsung dari subjek mengenai pengelolaan pondok pesantren.

Penelitian pengembangan ini dirancang untuk dilakukan dalam dua tahapan besar. Tahap pertama yaitu tahap pendahuluan yang meliputi kegiatan melakukan

pengumpulan informasi berdasarkan teori, dan hasil survei sementara. Kegiatan ini bertujuan untuk mengetahui kebutuhan akan perlunya sebuah model evaluasi pengelolaan pondok pesantren. Tahap kedua adalah tahap pengembangan yang terdiri dari langkah kedua sampai keempat.

Pada tahap pertama atau kegiatan pendahuluan, peneliti melakukan penelitian awal dilanjutkan dengan pengembangan. Kegiatan pendahuluan ini dilakukan dalam rangka menganalisis kebutuhan terhadap model yang dikembangkan dengan melibatkan pakar, dan praktisi pondok pesantren. Berdasarkan hasil penelitian awal ini akan dikembangkan instrumen model evaluasi pengelolaan pondok pesantren.

Pada tahap kedua, dilakukan pengembangan mulai dari penyusunan draf instrumen hingga pelaksanaan uji coba dan analisis. Untuk menghasilkan instrumen yang berkualitas, peneliti menggunakan teknik Delphi. Uji coba instrumen dilakukan untuk mengetahui validitas, mengestimasi reliabilitas serta penilaian responden mengenai model IBSQ. Uji coba dilakukan dalam dua tahap, yaitu uji coba terbatas dan uji coba lebih luas.

Dengan demikian model IBSQ dinilai dengan dua cara yaitu pertama, secara objektif yang akan menentukan validitas dan reliabilitas instrumen; kedua, secara subjektif berdasarkan penilaian pengguna (*reviewer*) model IBSQ berupa gambaran deskriptif model yang meliputi penilaian mengenai format, substansi, dan prosedur evaluasi.

Validitas instrumen yang digunakan dalam penelitian ini adalah validitas isi dan validitas konstruk. Teknik untuk memperoleh validitas isi dilakukan dengan menggunakan teknik Delphi.

Validitas konstruk dilakukan untuk menguji *construct validity* yaitu untuk mengetahui sejauhmana instrumen tersebut mengukur sifat konstruk tertentu. Untuk menganalisis validitas empirik instrumen dilakukan melalui metode *Exploratory Factor Analysis* (EFA) dengan bantuan program SPSS 17.

Estimasi reliabilitas instrumen dilakukan menggunakan *Cronbach Alpha* dengan bantuan program SPSS 17. Menurut Nunally (1981, p.230) apabila nilai *Alpha Cronbach* pada setiap tabel *output* yang diperoleh  $> 0,70$  maka instrumen tersebut termasuk reliabel.

Teknik analisis yang digunakan yaitu analisis deskriptif yang berfungsi untuk menganalisis data hasil penelitian dalam bentuk yang sederhana sehingga mudah mendapatkan gambaran hasil penelitian (Syamsuddin, 2002, p.19). Sementara itu, data yang diperoleh melalui observasi dan *interview* akan dianalisis secara kualitatif. Melalui analisis ini menurut Mishler (Denzin & Lincoln, 1994) bertujuan agar peneliti dapat mendeskripsikan dan menjelaskan pola hubungan yang hanya dapat dilakukan dengan seperangkat konsep yang spesifik.

Pengelolaan pondok pesantren dapat dikatakan baik apabila telah memenuhi enam komponen yang meliputi dokumen pesantren, kompetensi kiai, kompetensi ustaz, proses pembelajaran, penilaian hasil belajar, dan komponen alumni. Kriteria pengelolaan tersebut dapat dilihat pada Tabel 1.

Tabel 1. Kriteria Pengelolaan Pondok Pesantren

No	Skor	Kategori
1	1,000 - 1,999	Sangat Buruk
2	2,000 - 2,999	Buruk
3	3,000 - 3,999	Baik
4	4,000-5,000	Sangat Baik

## Hasil Penelitian dan Pembahasan

Mengacu pada empat tahapan penelitian dan pengembangan sebagaimana telah diuraikan sebelumnya, laporan dari masing-masing tahap adalah sebagai berikut.

Tahap pertama adalah studi pendahuluan. Studi ini dilakukan untuk mengetahui pendapat para pemerhati, peneliti, maupun pengelola pondok pesantren mengenai perlu tidaknya pengelolaan pondok pesantren dievaluasi.

Menurut peneliti sekaligus pemerhati pondok pesantren dari Litbang Jawa Tengah, Mudistaruna mengemukakan bahwa evaluasi pengelolaan pondok pesantren sangat perlu dilakukan. Menurutnya manajemen pengelolaan perlu mengikuti perkembangan tuntutan zaman sehingga mampu beradaptasi/bersaing dengan lembaga pendidikan formal yang cenderung dinamis (hasil wawancara pada tanggal 29 Januari 2015).

Perlunya evaluasi pengelolaan pondok pesantren juga dikemukakan oleh Fajar Shodiq yang merupakan dosen IAIN Surakarta sekaligus pemerhati pondok pesantren. Fajar Shodiq mengatakan bahwa semua lembaga termasuk pondok pesantren harus dievaluasi. Hal ini bertujuan untuk kemajuan pesantren sekaligus meminimalisasi kekurangan (hasil wawancara pada tanggal 30 Januari 2015).

Begitu juga pendapat dari pengelola Pondok Pesantren Al Falah (KH. Mahbub) dan Pondok Pesantren Al Fattah (Gus Muh Najib) yang mengatakan bahwa evaluasi sangat diperlukan. Tujuannya adalah sebagai sarana untuk meningkatkan mutu dan menghilangkan citra negatif tentang pesantren (hasil wawancara pada tanggal 30 Januari 2015).

Menurut Ibu Hj. Tasnim M (salah satu pengajar di pondok pesantren di Surakarta) evaluasi pengelolaan pondok pesantren dapat disebut juga dengan *muhasabah*. Kegiatan ini perlu dilakukan untuk mengetahui ketercapaian target yang telah ditentukan, sekaligus perbaikan pada langkah berikutnya (hasil wawancara pada tanggal 30 Januari 2015).

Menurut Subandji (Dosen dan pengelola pondok pesantren) bahwa evaluasi pengelolaan pondok pesantren sangat perlu dilakukan, terutama bagi pesantren tradisional. Karena manajemen di pesantren salafiah selama ini masih *top down*, statis dan kurang terbuka dengan pengaruh perkembangan lembaga dan santri. Lebih-lebih di era globalisasi yang penuh persaingan dan tantangan, pengelolaan perlu dikembangkan menjadi lebih demokratis, progressif, dan terbuka dengan dunia luar (hasil wawancara

pada tanggal 3 Februari 2015). Berdasarkan hasil *need assessment* tersebut dapat dinyatakan bahwa evaluasi pengelolaan pondok pesantren memang perlu dilakukan.

Tahap kedua adalah penyusunan draf instrumen. Penyusunan draf awal dilakukan dengan mempertimbangkan temuan-temuan hasil studi pendahuluan. Draft rancangan awal dimaksudkan untuk memperoleh kriteria pengelolaan pondok pesantren yang penting untuk dievaluasi.

Para ahli dan praktisi pondok pesantren yang ikut memberikan kontribusi adalah H.A. Fauzi, Jakfar Assagaf, Sukirman (Alumni Pondok Pesantren), Subandji, K. H. Moh Mahbub (Pengelola Pondok Pesantren) dan Retno Wahyuningsih (Mahasiswa S3 PEP).

Pada draft kisi-kisi instrumen, terdapat sekitar 78 kriteria dari 34 komponen. Secara keseluruhan, draft kisi-kisi instrumen diterima oleh para praktisi. Hanya beberapa butir kriteria saja yang menurut mereka penting untuk dimasukkan sebagai bahan evaluasi.

Selain penilaian tersebut, para praktisi juga memberikan masukan. Ahmad Fauzi memberikan masukan, untuk komponen kriteria kiai perlu ditambah dengan indikator "ahli agama", sementara untuk komponen kriteria ustaz ditambah dengan indikator "keluarga kiai atau alumni".

Menurut Subandji, indikator dari kriteria kiai perlu ditambah dengan indikator "keilmuan, dan kepemimpinan", sedangkan untuk komponen kriteria ustaz ditambah dengan "loyalitas atau kesetiaan", "kesalehan", dan "keilmuan".

Menurut Jakfar Assagaf, untuk komponen kriteria kiai perlu ditambah dengan indikator "keilmuan", komponen metode pemilihan kiai ditambah dengan indikator "penunjukan", dan untuk komponen kriteria ustaz perlu ditambah dengan indikator "alumni".

Setelah draft instrumen dinyatakan baik oleh para narasumber, selanjutnya instrumen secara utuh dikonsultasikan kepada para ahli yang lain. Pada kegiatan Delphi putaran kedua para ahli yang berpartisipasi terdiri dari

ahli evaluasi dan ahli dalam bidang pondok pesantren.

Praktisi pendidikan yang ahli dalam bidang pondok pesantren diambil dari Universitas Islam Negeri Sunan Kalijaga Yogyakarta yakni: (1) Abdurrahman Assegaf, (2) Sutrisno, dan dari IAIN Surakarta yaitu Toto Suharto. Sementara pakar evaluasi diambil dari Universitas Negeri Yogyakarta yaitu Badrun Kartowagiran.

Pemberian masukan draf rancangan awal model evaluasi ini dilaksanakan selama bulan April dan Mei 2014. Menurut Bapak Abdurrahman Assagaf bahwa ciri-ciri umum pondok pesantren hendaknya dimasukkan dalam instrumen penelitian seperti: kiai, santri, Masjid, kitab kuning, dan asrama; Madrasah sebagai bagian dari pondok pesantren, juga hendaknya dimasukkan dalam instrumen.

Menurut Sutrisno, aspek keteladanan perlu masuk sebagai salah satu kriteria kiai. Untuk menilai perilaku santri diperlukan pengamatan. Untuk kinerja alumni ditambah dengan kesesuaian pekerjaan dengan ilmu yang dipelajari di Pesantren.

Menurut Toto Suharto bahwa perlu diungkap ideologi pendidikan pondok pesantren. Banyak pondok pesantren yang berafiliasi dengan jaringan Timur Tengah yang ditengarai memiliki ideologi tertentu; perlu diungkap *hidden* kurikulum masing-masing pesantren; pembiayaan pesantren perlu transparan, karena banyak pesantren yang memperoleh sumber biaya dari Timur Tengah/Luar Negeri; perlu diperbaiki tata tulis yang masih salah. Misalnya: instrumen, kriteria, dan sebagainya.

Menurut Bapak Badrun Kartowagiran dalam instrumen tersebut evaluator kiai tidak hanya kiai saja, tetapi juga orang lain seperti: ustaz, dan santri; evaluator ustaz juga tidak hanya ustaz yang bersangkutan tetapi perlu ditambah orang lain seperti kiai/pimpinan, sesama ustaz, dan santri. Untuk penskoran, perlu diberi pilihan ya dan tidak.

Selain masukan secara tertulis, para pakar juga memberikan penilaian secara umum terhadap buku panduan evaluasi dan

penilaian terhadap komponen instrumen. Instrumen dinilai menggunakan rentang 1-5, sehingga skor maksimal adalah 5.

Penilaian secara umum merupakan penilaian para pakar terhadap pedoman melakukan evaluasi pengelolaan pondok pesantren yang meliputi unsur tampilan sampai dengan kemungkinan ketercapaian evaluasi dengan menggunakan model IBSQ.

Hasil penilaian terhadap pedoman evaluasi pengelolaan pondok pesantren dari para pakar menunjukkan bahwa secara keseluruhan pedoman evaluasi pengelolaan pondok pesantren dinilai sangat baik. Uraian dari hasil tersebut sebagaimana tertulis di bawah ini.

Pedoman evaluasi mempunyai 14 dimensi. Dari 14 dimensi tersebut, 12 di antaranya berkategori sangat baik yaitu tampilan depan, tampilan isi, cakupan ruang lingkup, keluasan penjabaran komponen, tingkat keterbacaan, kemudahan untuk dipahami, sistematika penulisan, penggunaan bahasa, pemilihan huruf, ketebalan halaman, kepraktisan, dan ketercapaian evaluasi. Komponen yang dinilai baik yaitu "*lay out*" dan "efektivitas waktu mengerjakan" yang berkategori baik. *Lay out* instrumen memang dibuat sangat sederhana sehingga terkesan monoton. Sementara untuk efektivitas waktu mengerjakan, beberapa dari responden mengomentari jika penilaian terhadap buku pedoman evaluasi memakan waktu yang cukup banyak karena dalam mengisi harus berpikir, tidak bisa langsung mengisi tanpa memahami isinya.

Berdasarkan hasil perhitungan diketahui bahwa instrumen pedoman evaluasi secara keseluruhan memperoleh rerata skor 4,71. Perolehan rerata tersebut menunjukkan bahwa instrumen pedoman evaluasi berkategori sangat baik.

Selain penilaian secara umum, para pakar juga diminta untuk menilai instrumen yang telah disusun. Sebagaimana telah dijelaskan sebelumnya bahwa instrumen evaluasi pengelolaan pondok pesantren terdiri dari aspek *input*, proses, dan *output*. Aspek *input* meliputi komponen kiai, ustaz, santri, perencanaan, kurikulum, sarana prasarana,

dan pembiayaan. Hasil penilaian para pakar terhadap poin-poin komponen dan kriteria evaluasi pengelolaan pondok pesantren dijelaskan melalui uraian berikut ini:

Menurut penilaian para pakar, semua komponen dalam aspek *input* dinilai sangat baik untuk disertakan dalam penyusunan instrumen evaluasi pengelolaan pondok pesantren. Hanya beberapa indikator dari komponen tersebut yang dinilai baik yaitu metode pemilihan kiai dan jumlah ustaz.

Aspek proses meliputi komponen proses pembelajaran dan sistem penilaian. Semua indikator dalam komponen tersebut dinilai baik. Aspek terakhir yaitu *output* diwakili oleh komponen kompetensi alumni. Para pakar menilai baik untuk semua indikator pada komponen tersebut.

Berdasarkan hasil perhitungan instrumen penelitian, secara keseluruhan diperoleh rerata skor 4,17. Perolehan rerata tersebut menunjukkan bahwa instrumen berkategori sangat baik dan dapat digunakan.

Tahap ketiga adalah uji coba meliputi uji coba terbatas dan uji coba lebih luas.

#### Uji Coba Terbatas

Uji coba terbatas dilakukan selama bulan Mei dan Juni 2014 di Pondok Pesantren Al Fattah Kartasura dan Pondok Pesantren Walisongo Sragen. Pada uji coba terbatas ini terdapat 52 responden, akan tetapi hanya 13 orang yang menilai keterbacaan instrumen model IBSQ. Hasil uji coba terbatas terhadap kedua Pondok Pesantren disajikan menyatu sebagaimana disajikan pada Tabel 2.

Tabel 2. Hasil Penilaian terhadap Instrumen Pedoman Evaluasi Model IBSQ

FORMAT SECARA UMUM			
No	Komponen	Rerata	Kriteria
1	Pengemasan dan tampilan model evaluasi	2,762	Menarik
2	Lay out penulisan	3,076	Baik
3	Pemilihan huruf, font, dan spasi	3,307	Sangat baik
4	Sistematika penulisan	3,0	Baik
5	Penggunaan bahasa	3,076	Baik
6	Ketebalan	2,384	Cukup tebal
7	Tingkat keterbacaan	3,615	Sangat mudah dibaca
8	Kemudahan untuk dipahami	2,846	Mudah dipahami
SUBSTANSI MODEL EVALUASI			
1	Kemudahan memahami pedoman	2,923	Mudah dipahami
2	Cakupan ruang lingkup evaluasi	3,384	Telah mencakup
3	Keluasan penjabaran komponen	3,230	Cukup menjabarkan komponen
4	Petunjuk mengerjakan instrument	3,538	Mudah dimengerti
5	Kemudahan untuk dikerjakan	2,769	Cukup mudah dikerjakan
6	Waktu untuk mengerjakan	2,615	Agak menyita waktu
7	Kemaanfaatan	3,307	Sangat bermanfaat
8	Urgensi evaluasi	3,307	Sangat dibutuhkan untuk mengevaluasi
9	Ketercapaian evaluasi	3,153	Cukup mampu mengevaluasi
10	Dibandingkan dengan evaluasi diri pesantren	2,769	Memiliki kesulitan yang sama
11	Dibandingkan dengan model evaluasi lain	3,000	Memiliki kesulitan yang sama
PROSEDUR EVALUASI			
1	Persiapan dan perencanaan	2,846	Agak merepotkan
2	Pelaksanaan evaluasi	2,538	Mudah dilakukan
3	Analisis data evaluasi	2,461	Cukup rumit
4	Penetapan kriteria hasil evaluasi	2,538	Mudah dilakukan
5	Penyusunan laporan hasil evaluasi	2,846	Mudah dilakukan

Sumber: Hasil Analisis Data Primer

Aspek ketebalan, diakui oleh beberapa pihak pedoman model evaluasi IBSQ cukup tebal sehingga bagi responden yang menerimanya merasa malas untuk membacanya. Perlu dikaji cara yang efektif yang dapat menghilangkan kesan tebal pada buku model IBSQ tersebut. Sementara untuk analisis data dinilai cukup rumit.

Hal ini dapat dimaklumi karena penilai berasal dari pondok pesantren yang tidak mempunyai latar belakang pendidikan berbasis ilmu-ilmu eksakta. Sebagai masukan, perlu dipertimbangkan cara yang lebih efektif untuk membantu responden memahami analisis data dengan lebih mudah.

Hasil evaluasi pengelolaan pondok pesantren menggunakan model IBSQ pada tahap uji coba terbatas juga memperoleh hasil yang baik. Komponen yang dievaluasi dalam penelitian ini terdiri dari enam komponen yaitu dokumen, kiai, ustaz, proses pembelajaran, penilaian hasil belajar, dan kompetensi alumni.

Komponen dokumen memperoleh rerata skor 3,5 berkategori sangat baik. Komponen kiai memperoleh rerata skor 2,89 berkategori baik. Komponen ustaz memperoleh rerata skor 3,38 berkategori sangat baik.

Selanjutnya, komponen proses pembelajaran memperoleh rerata skor 3,08 berkategori baik. Komponen penilaian hasil belajar memperoleh rerata skor 3,81 berkategori sangat baik, dan kompetensi alumni memperoleh rerata skor 3,27.

Berdasarkan hasil uji coba terbatas instrumen pedoman evaluasi IBSQ dinyatakan baik dan hasil evaluasi pengelolaan pondok pesantren juga baik, maka tahap berikutnya adalah melakukan uji coba lebih luas.

#### Uji Coba Lebih Luas

Uji coba lebih luas dilakukan di Pondok Pesantren Al Falah Sidoharjo Sragen, Pondok Pesantren Assalaam Surakarta, dan Pondok Pesantren Al Muayyad Surakarta. Penelitian dilakukan selama bulan Agustus dan September 2014.

Dengan menggunakan Tabel 3 akan diketahui hasil penilaian para responden dari Pondok Pesantren Al Falah Sidoharjo Sragen, Pondok Pesantren Assalaam, dan Pondok Pesantren Al Muayyad Surakarta.

Tabel 3. Konversi Nilai Uji Coba Lebih Luas

No	Skor	Kategori
1	1,000 - 1,7499	Sangat Buruk
2	1,75 - 2,499	Buruk
3	2,500 - 3,249	Baik
4	3,25 - 4,000	Sangat Baik

Berdasarkan hasil penilaian para pengguna bahwa pedoman evaluasi pengelolaan pondok pesantren secara keseluruhan dinilai baik dengan perolehan rerata skor 3,15. Penjelasan secara rinci diuraikan berikut ini:

Pada aspek format evaluasi secara umum diperoleh hasil sebagai berikut: (1) pengemasan dan tampilan model evaluasi memperoleh rerata skor 2,89 atau dinilai cukup menarik; (2) *lay out* penulisan memperoleh rerata skor 3,17 atau dinilai baik; (3) pemilihan huruf, font, dan spasi memperoleh rerata skor 3,32 atau dinilai sangat baik; (4) sistematika penulisan memperoleh rerata skor 3,23 atau dinilai baik; (5) penggunaan bahasa memperoleh rerata skor 3,67 atau dinilai sangat baik; (6) ketebalan memperoleh rerata skor 2,20 atau dinilai cukup tebal; (7) tingkat keterbacaan memperoleh rerata skor 3,79 atau dinilai mudah dibaca; dan (8) kemudahan untuk dipahami memperoleh rerata skor 3,25 atau dinilai agak sulit untuk dipahami.

Pada aspek substansi model evaluasi diperoleh hasil sebagai berikut: (1) kemudahan memahami pedoman memperoleh rerata skor 3,06 atau dinilai mudah untuk dipahami; (2) cakupan ruang lingkup evaluasi memperoleh rerata skor 3,5 atau dinilai telah mencakup ruang lingkup evaluasi; (3) keluasan penjabaran komponen memperoleh rerata skor 3,21 atau dinilai cukup menjabarkan komponen; (4) petunjuk mengerjakan instrumen memperoleh rerata skor 3,64 atau dinilai mudah dimengerti; (5) kemudahan untuk dikerjakan memperoleh

rerata skor 3,97 atau dinilai cukup mudah dikerjakan; (6) waktu untuk mengerjakan memperoleh rerata skor 3,22 atau dinilai tidak menyita waktu; (7) kemanfaatan memperoleh rerata skor 3,51 atau dinilai sangat bermanfaat; (8) urgensi evaluasi memperoleh rerata skor 3,26 atau dinilai cukup penting; (9) ketercapaian evaluasi memperoleh rerata skor 3,14 atau dinilai cukup mampu mengevaluasi; (10) dibandingkan dengan evaluasi diri pesantren memperoleh rerata skor 2,75 atau dinilai agak sulit untuk digunakan; dan (11) dibandingkan dengan model evaluasi lain memperoleh rerata skor 2,96 atau dinilai memperoleh kesulitan yang sama.

Pada aspek prosedur evaluasi diperoleh hasil sebagai berikut: (1) persiapan dan perencanaan memperoleh rerata skor 3,5 atau dinilai tidak merepotkan; (2) pelaksanaan evaluasi memperoleh rerata skor 3,12 atau dinilai mudah dilakukan; (3) analisis data evaluasi memperoleh rerata skor 2,69 atau dinilai mudah dilakukan; (4) penetapan kriteria hasil evaluasi memperoleh rerata skor 2,85 atau dinilai mudah dilakukan; (5) penyusunan laporan hasil evaluasi memperoleh rerata skor 2,78 atau dinilai mudah.

Berdasarkan hasil penilaian responden dari ketiga pondok pesantren dapat dinyatakan bahwa pedoman evaluasi pengelolaan pondok pesantren secara keseluruhan dinilai baik dan dapat digunakan oleh pondok pesantren lainnya.

Setelah mengevaluasi pedoman evaluasi IBSQ, berikutnya adalah mengujicobakan instrumen tersebut pada wilayah yang lebih luas. Objek penelitian kali ini adalah Pondok Pesantren Al Falah Sidoharjo Sragen, Pondok Pesantren Assalaam Sukoharjo, dan Pondok Pesantren Al Muayyad Surakarta.

Pertama, hasil analisis secara keseluruhan dari evaluasi pengelolaan Pondok Pesantren Al Falah memperoleh skor rerata total sebanyak 2,87 dan berkategori baik. Meskipun demikian, terdapat beberapa komponen evaluasi yang memperoleh nilai rerata kurang baik yakni komponen dokumentasi, dan kompetensi kiai.

Untuk komponen dokumen, pengelola Pondok Pesantren Al Falah memang tidak mengisi instrumen evaluasi, karena menurut salah satu pengelola belum didokumentasikan. Misalnya; untuk kurikulum, pengelola hanya menunjukkan jadwal mengaji santri yang ditulis tangan. Hal tersebut dapat dimaknai bahwa secara praktis, pelaksanaan pembelajaran yang terdapat di Pondok Pesantren Al Falah dilandasi oleh tujuan awal pendirian pesantren. Dari tujuan tersebut, muncul sejumlah materi pelajaran (kurikulum) selama satu tahun. Untuk mengetahui tingkat pencapaian tujuan, pondok pesantren juga melakukan evaluasi. Hanya saja, kesemuanya itu belum didokumentasikan dengan baik.

Sementara untuk komponen kiai, memang banyak terdapat fakta yang tidak sesuai dengan pernyataan yang terdapat dalam instrumen tersebut. Misalnya; mekanisme pemilihan kiai, pembatasan masa kerja kiai, mekanisme perekrutan ustaz, perekrutan santri, dan pengadaan barang.

Kedua, hasil analisis secara keseluruhan dari evaluasi pengelolaan Pondok Pesantren Assalaam memperoleh skor rerata total sebanyak 3,48 dan berkategori sangat baik. Kategori ini sejalan dengan hasil observasi di Pondok Pesantren Assalaam. Pondok Pesantren Assalaam telah melakukan pendokumenan dengan baik, hampir semua kegiatan dapat didokumentasikan melalui *web*-nya sendiri. Kriteria kiai, ustaz dan santri juga dapat dideskripsikan dengan baik. Hal ini juga dapat diketahui melalui proses pembelajaran yang berjalan sangat baik ditunjang dengan fasilitas yang sangat memadai.

Begitu pun dengan alumni yang saat ini telah tersebar di berbagai daerah. Mereka telah menunjukkan kemampuannya di bidang agama kepada masyarakat dengan cara berdakwah, menjadi pengelola organisasi masyarakat dan sebagainya.

Ketiga, hasil analisis secara keseluruhan dari evaluasi pengelolaan Pondok Pesantren Al Muayyad memperoleh skor rerata total sebanyak 3,28 dan berkategori sangat baik. Meskipun sama-sama berkat-

gori sangat baik, akan tetapi terdapat perbedaan perolehan rerata skor dibandingkan hasil dari Pondok Pesantren Assalaam.

Dokumentasi di Pondok Pesantren Al Muayyad telah diarsip dengan cukup baik. Segala sesuatunya tengah diusahakan berbasis IT, meski masih terkendala dengan sumber daya. Proses pembelajaran juga telah menggunakan strategi yang bertujuan mengaktifkan santri, sarana prasarana diusahakan terpenuhi dengan baik.

Struktur kepengurusan pondok pesantren telah melibatkan pihak dari luar pesantren, meskipun demikian unsur kekeluargaan masih terasa dalam pengelolaan pondok pesantren.

Jika mencermati hasil analisis deskriptif secara keseluruhan, semua komponen evaluasi memiliki hasil yang baik. Hal ini ditunjang dari hasil hitung menggunakan program SPSS 17 dalam menguji validitas yang dilihat dari perolehan nilai KMO dan besarnya persentase yang dapat menjelaskan jumlah variasi serta perolehan nilai reliabilitasnya.

Komponen dokumen pesantren memperoleh nilai KMO 0,656 dengan muatan *loading factor* terendah 0,711 dan tertinggi 0,735, sedangkan reliabilitasnya sebesar 0,99. Untuk komponen kiai memperoleh nilai KMO 0,736 dengan muatan *loading factor* terendah 0,668 dan tertinggi 0,695, sedangkan reliabilitasnya sebesar 0,913.

Komponen ustaz memperoleh nilai KMO 0,862 dengan muatan *loading factor* terendah 0,523 dan tertinggi 0,851, sedangkan reliabilitasnya sebesar 0,934. Komponen proses pembelajaran memperoleh nilai KMO 0,742 dengan muatan *loading factor* terendah 0,527 dan tertinggi 0,800, sedangkan reliabilitasnya sebesar 0,845.

Komponen penilaian hasil belajar memperoleh nilai KMO 0,939 dengan muatan *loading factor* terendah 0,605 dan tertinggi 0,754, sedangkan reliabilitasnya sebesar 0,993. Hasil perhitungan tersebut menunjukkan bahwa semua komponen memperoleh nilai KMO terendah 0,517 dan tertinggi 0,939. Menurut Ghazali (2009, p.53) nilai KMO yang dikehendaki harus mem-

peroleh  $> 0,5$  untuk dapat dilakukan analisis faktor. Begitu juga dengan jumlah persentase yang melebihi 50% pada tiap komponennya. Selain itu dilihat dari muatan *loading factor*-nya menunjukkan bahwa *loading factor* terendah 0,523 dan tertinggi sebesar 0,996. Hal ini menunjukkan bahwa semua instrumen dinyatakan valid.

Parameter yang digunakan untuk mengetahui reliabilitas adalah dengan melihat nilai *Alpha Cronbach* pada setiap tabel *output* yang diperoleh, dengan ketentuan jika indeks Alphanya  $> 0,70$  maka instrumen tersebut termasuk reliabel (Nunally, 1981, p.230). Dari perhitungan tersebut diketahui semua komponen memperoleh  $\alpha > 0,70$ . Perolehan nilai *Alpha Cronbach* terendah sebesar 0,845 dan tertinggi sebesar 0,999. Hal ini menunjukkan bahwa instrumen tersebut memiliki tingkat keandalan yang tinggi.

#### Revisi Produk

Tahap keempat adalah revisi produk. Beberapa catatan yang harus diperhatikan dan diperbaiki dari uji coba terbatas adalah: (a) jumlah responden harus ditambah lebih banyak agar hasil analisis lebih maksimal. Pada ujicoba terbatas jumlah responden hanya sebanyak 54 orang yang terdiri dari kiai, ustaz, santri dan alumni. Dari jumlah tersebut, beberapa komponen evaluasi tidak dapat menunjukkan hasil yang maksimal; (b) penampilan *layout* instrumen kurang menarik, karena hanya berupa hasil *foto copy* dengan ukuran kertas folio yang memanjang; (c) pedoman evaluasi terlalu tebal, sehingga membuat responden merasa malas untuk mengisi karena cukup menyita waktu; (d) bahasa yang digunakan terdengar asing bagi responden, ada baiknya menggunakan bahasa yang lebih mudah dipahami.

Berdasarkan beberapa catatan tersebut maka perbaikan yang dilakukan pada tahap ini adalah: (1) menambah jumlah responden; (2) instrumen dijilid menjadi sebuah buku yang cukup menarik; (3) halaman dalam pedoman evaluasi diringkas lagi sehingga mengurangi ketebalan; dan (4) mengurangi bahasa yang terdengar asing bagi santri.

Setelah revisi uji coba terbatas dilakukan, dilanjutkan dengan revisi uji coba lebih luas. Beberapa catatan yang harus diperhatikan dan diperbaiki adalah: (a) bahasa yang digunakan masih belum sepenuhnya bisa dipahami oleh responden terutama santri; (b) penampilan instrumen masih terlalu sederhana, sehingga kurang menarik

Selain itu, berdasarkan hasil analisis data terdapat perubahan nama faktor pada beberapa komponen evaluasi: (1) pada komponen dokumentasi pondok pesantren, faktor pertama disebut sebagai "profil pesantren" dan faktor kedua disebut sebagai "kegiatan pondok pesantren"; (2) pada komponen proses pembelajaran, faktor pertama disebut sebagai faktor "persiapan"; faktor kedua disebut sebagai faktor "kegiatan inti dalam pembelajaran" dan faktor ketiga adalah "kegiatan mengoperasikan media pembelajaran"; (3) pada komponen penilaian hasil belajar, faktor pertama disebut sebagai "faktor kerangka konsep penilaian" dan faktor kedua disebut sebagai "faktor bentuk dan mekanisme penilaian"; dan (4) pada komponen alumni, faktor pertama disebut sebagai "faktor hubungan alumni dengan masyarakat", dan faktor kedua disebut sebagai faktor "hubungan alumni dengan diri sendiri dan keluarga".

Catatan lain yang diperoleh, bahwa hampir semua responden terutama para ustaz menyatakan penting untuk dilakukan evaluasi. Meskipun demikian, mereka juga mengatakan bahwa sebenarnya pondok pesantren tidak dapat dibandingkan, karena semua memiliki karakteristik sendiri-sendiri.

#### Produk Akhir

Hasil akhir dari penelitian ini berupa sebuah model evaluasi pengelolaan Pondok Pesantren IBSQ terdiri dari enam komponen evaluasi, yaitu: dokumen, kompetensi kiai, kompetensi ustaz, proses, pembelajaran, penilaian dan kompetensi alumni. Keenam komponen tersebut disusun menjadi lima buku. Buku-buku tersebut adalah sebagai berikut.

- Buku 1: Berisi instrumen tentang evaluasi dokumen, yaitu dokumen visi, misi dan tujuan pondok pesantren, struktur organisasi pesantren, kurikulum, pembiayaan, pengelolaan program, pengawasan/control
- Buku 2: Berisi instrumen tentang pelaksanaan proses belajar mengajar
- Buku 3: Berisi instrumen tentang kompetensi kiai/pimpinan pondok pesantren
- Buku 4: Berisi instrumen tentang kompetensi ustaz dan penilaian hasil belajar
- Buku 5: Berisi instrumen tentang kompetensi alumni

Kelima buku tersebut diisi oleh responden dari pondok pesantren yang terdiri dari kiai, kepala sekolah, ustaz, pengelola, santri, dan alumni. Buku 1 berisi tentang dokumen pesantren terdiri dari 14 butir yang diisi oleh pengelola pondok pesantren.

Buku 2 berisi tentang proses pembelajaran terdiri dari 17 butir yang diisi oleh ustaz dan santri. Buku 3 berisi tentang kompetensi kiai yang terdiri dari 42 butir, diisi oleh kiai dan ustaz. Buku 4 berisi tentang kompetensi ustaz (25 butir) dan penilaian hasil belajar (36 butir) yang diisi oleh kiai/kepala sekolah, ustaz, dan santri. Buku 5 berisi tentang kompetensi alumni terdiri dari 8 butir yang diisi oleh alumni setiap pondok pesantren.

Sebagai sebuah model evaluasi, evaluasi pengelolaan pondok pesantren menggunakan model IBSQ dan perangkatnya telah diuji implementasinya. Hasil tersebut menunjukkan bahwa implementasi model ini cukup praktis dan objektif. Hasil uji secara kuantitatif dalam skala menengah menggunakan program SPSS menunjukkan bahwa instrumen evaluasi model IBSQ semuanya telah memenuhi koefisien reliabilitas dan butir-butir instrumennya telah valid. Berdasarkan analisis data secara kualitatif model ini cukup sederhana dalam implementasinya, praktis, objektif, dan disertai panduan evaluasi pembelajaran yang cukup singkat, jelas dan lengkap sehingga mem-

berikan kemudahan dalam pelaksanaan evaluasi pengelolaan pondok pesantren.

Adapun karakteristik selengkapnya dari model evaluasi ini sebagai berikut: (1) model ini digunakan untuk mengevaluasi pengelolaan pondok pesantren yang bertipe *salafiyah*, modern, dan kombinasi; (2) model evaluasi pengelolaan ini bersifat komprehensif, terdiri dari tiga komponen utama mencakup: *input*, *process*, dan *output*; (3) penggunaan evaluasi model ini tidak tergantung pada pendekatan pengelolaan tertentu yang dilaksanakan oleh pengelola pondok pesantren; (4) model evaluasi pengelolaan ini dapat digunakan sebagai *diagnostic evaluation*, yaitu menemukan kelemahan/kekurangan pengelolaan pondok pesantren, dengan demikian kekurangan/kelemahan tersebut dapat segera diperbaiki. Perbaikan tersebut terutama menyangkut komponen proses, sehingga diharapkan pada komponen *output* menjadi lebih baik.

Dengan demikian, model evaluasi pengelolaan pondok pesantren ini akan sangat baik digunakan oleh setiap pondok pesantren. Jika pedoman ini akan mengevaluasi semua komponen dalam pengelolaan pondok pesantren, penelitian Muiyasaroh (2014) lebih dikhususkan untuk mengevaluasi program tahfiz di pondok pesantren.

Meskipun berbeda, akan tetapi penemuan dari dua hasil penelitian tersebut sejatinya merupakan sebuah sinergi yang mampu meningkatkan kualitas pondok pesantren.

## **Simpulan**

Berdasarkan hasil penelitian dan analisis data yang telah diuraikan sebelumnya, maka dapat disimpulkan bahwa melalui pengembangan model evaluasi pengelolaan pondok pesantren telah dihasilkan lima buah buku yang mempunyai kriteria sebagai berikut: (1) karakteristik instrument dalam model IBSQ: (a) memiliki format yang baik (rata-rata 3,19); (b) telah memenuhi substansi model evaluasi (rata-rata 3,19); (c) memiliki tingkat validitas konstruk yang dapat diandalkan, terbukti pada uji coba lebih luas semua indikator mengelompok pada

satu faktor (unikomponen); (d) memiliki tingkat reliabilitas yang tinggi di atas 0,61; (2) model IBSQ memiliki tingkat efektivitas yang baik, hal ini terbukti dengan 76% pengguna dapat menggunakan dengan mudah; (3) gambaran pengelolaan pada setiap pondok pesantren secara nyata terjadi perbedaan: (a) pengelolaan di Pondok Pesantren Al Falah masih sangat tradisional dengan bukti salah satunya kegiatan administrasi masih manual; (b) pengelolaan di Pondok Pesantren Assalaam sudah memanfaatkan IT sebagai sarana dalam menjalankan pengelolaan selama ini; (c) pengelolaan di Pondok Pesantren Al Muayyad masih campuran antara manual dan berbasis komputer.

Sebagaimana telah disebutkan sebelumnya bahwa salah satu karakteristik model IBSQ adalah bersifat komprehensif, dan dapat digunakan bagi pondok pesantren salafiah, modern maupun kombinasi. Maka sudah saatnya pondok pesantren berbenah dan mulai terbuka untuk melakukan evaluasi diri. Hal ini dimaksudkan agar pelaksanaan evaluasi lebih mudah diterima dan dilaksanakan di pondok pesantren. Melalui model IBSQ, setidaknya pengelola pondok pesantren dapat mengkaji ulang proses yang selama ini telah dijalankan, untuk selanjutnya dapat memperbaiki serta meningkatkan kualitas dari kegiatan-kegiatan yang telah dilakukan.

Berdasarkan hasil penelitian juga dapat diketahui bahwa bahwa Pondok Pesantren Salafiah Al Falah tidak menggunakan prinsip-prinsip manajemen rasional sebagaimana pondok pesantren modern, akan tetapi dalam pelaksanaannya mampu menjalankan roda organisasi (pondok pesantren) dengan baik. Melihat realitas ini, sangat penting kiranya bagi peneliti berikutnya untuk melakukan penelitian secara mendalam tentang manajemen pesantren salafiah.

## **Daftar Pustaka**

Arifin, (1991). *Kapita selekta pendidikan (Islam dan umum)*. Jakarta: Bumi Aksara

- Borg, P. L & Gall, M. D (1983) *Educational research: an introduction (4rd ed)*. New York: Longman, Inc
- Denzin, N. K., Lincoln, Y. S. (2000) *Handbook of qualitative research, 2nd edition*. London: Sage Publication, Inc, International Educational and Professional Publisher
- Dhofier, Z. (1985) *Tradisi pesantren: studi tentang pandangan hidup Kyai*. Jakarta: LP3ES.
- Ghozali, I. (2009). *Aplikasi analisis multivariate dengan program SPSS*. Semarang: Badan Penerbit Universitas Diponegoro
- Muyasaroh, M., & Sutrisno, S. (2014). Pengembangan instrument evaluasi CIPP pada program pembelajaran tahfiz Al-Qur'an di pondok pesantren. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 18(2), 215-233. Retrieved from <http://journal.uny.ac.id/index.php/jpep/article/view/2862/2389>
- Nunally, J. C. (1981). *Psychometric theory (3<sup>rd</sup> ed)*. New York: McGraw-Hill Book Company
- Nafi', M. D., et al. (2007) *Praksis pembelajaran pesantren*. Yogyakarta: LkiS Pelangi Aksara
- Qomar, M. (2007) *Manajemen pendidikan Islam strategi baru pengelolaan lembaga pendidikan Islam*. Jakarta: Penerbit Erlangga
- Rahim, H. (2001). *Arab baru pendidikan Islam di Indonesia*. Jakarta: PT. Logos Wacana Ilmu

**PEMETAAN MUTU PAUD *FULLDAY* UNTUK  
MENINGKATKAN KUALITAS LAYANAN DAN KELEMBAGAAN PROGRAM PAUD**

<sup>1)</sup> Puji Yanti Fauziah, <sup>2)</sup> Sugito

<sup>1)</sup>pujiyanti@uny.ac.id, <sup>2)</sup>sugito@uny.ac.id

<sup>1,2)</sup>Fakultas Ilmu Pendidikan Universitas Negeri Yogyakarta

**Abstrak**

Penelitian ini bertujuan untuk: (1) memetakan Program PAUD *fullday*; (2) menghasilkan rekomendasi penyelenggaraan PAUD *fullday*. Metode yang digunakan dalam penelitian ini adalah survei. Penentuan lembaga yang dijadikan sampel dilakukan secara acak sejumlah 23 lembaga se-DIY. Adapun hasil penelitiannya: (1) Standar pencapaian perkembangan anak sudah baik tetapi hasil capaian perkembangan baru hasil observasi dan penilaian kegiatan pagi sampai siang dan belum dilakukan terintegrasi dan holistik sepanjang hari; (2) rekomendasi dalam penyelenggaraan Program PAUD *fullday* adalah: (a) diperlukan kurikulum model pembelajaran yang holistik dalam arti menstimulasi seluruh aspek perkembangan anak dari pagi sampai sore; (b) diperlukan model penilaian yang integratif dari pagi sampai sore dari keseluruhan aktivitas anak, dan (c) pentingnya pemberdayaan orang tua dalam pembelajaran *fullday* dan lembaga telah melakukan pemberdayaan dengan baik.

**Kata kunci:** PAUD *fullday*, kualitas layanan dan kelembagaan

**THE QUALITY MAPPING OF FULL DAY ECD FOR IMPROVING THE QUALITY  
OF SERVICE AND INSTITUTIONAL PROGRAMS ECD**

<sup>1)</sup> Puji Yanti Fauziah, <sup>2)</sup> Sugito

<sup>1)</sup>pujiyanti@uny.ac.id, <sup>2)</sup>sugito@uny.ac.id

<sup>1,2)</sup>Fakultas Ilmu Pendidikan Universitas Negeri Yogyakarta

**Abstract**

This research aims to: (1) mapping the full day early childhood program; (2) Recommendation implementation of full day early childhood education. The method used in this research is the survey. Determination institution randomized sample number 23 taken from the early childhood institutions throughout the District. The description of the implementation of full day early childhood are: 1 Achievement standard child development has been good but the results of the achievements of the new developments on the observation and assessment of activities in the morning until noon, has not carried out an integrated and holistic throughout the day. recommendations for the implementation of full day early childhood program; (a) required a holistic learning model in the sense of stimulating all aspects of child development from morning till evening; (b) assessment of the achievement of the child's development has not been documented in a neat necessitating a thorough assessment model of the overall activity of the child; (c) the importance of empowering parents in a full day of learning , and institutions have to empower parents with the good

**Keywords:** *full day early childhood , quality of service and institutional*



## Pendahuluan

Pembentukan kualitas sumber daya manusia ditentukan oleh tiga lingkungan, yaitu sekolah, keluarga dan masyarakat. Dewantara (1962, p.386) menyebutnya dengan tiga alam pendidikan. Berbagai penelitian ilmiah tentang pentingnya pendidikan keluarga menyebutkan bahwa keluarga sebagai pendidikan utama dan pertama bagi anak.

Perubahan sosial yang terjadi pada bentuk keluarga dari keluarga besar (*extended family*) menjadi keluarga inti (*nuclear family*) menyebabkan adanya perubahan pola asuh yang sangat bergantung pada pasangan suami istri. Hal ini senada dengan yang dikemukakan oleh Faturochman (2001, p. 2) yang diambil dari Ardiner & Gardiner, 1988 selama beberapa dekade terakhir, keluarga mengalami perubahan bentuk dari ukuran keluarga besar menjadi lebih kecil. Perubahan ini mempengaruhi aspek-aspek yang lainnya. Perubahan pertama adalah jumlah keluarga mengecil dari keluarga luas menjadi keluarga inti. Kedua selama masa transisi ini peran keluarga mengalami perubahan dari peran sosial emosional keluarga ke peran ekonomis. Fenomena peningkatan suami istri bekerja banyak mengurangi waktu keluarga sehingga anak lebih banyak berinteraksi dan dipengaruhi oleh lingkungan sekolah.

Peningkatan peran istri bekerja dari tahun ke tahun mengalami peningkatan, hal ini didukung oleh adanya revolusi industri yang banyak memberikan kesempatan perempuan untuk bekerja. Hasil penelitian di Amerika menunjukkan selama beberapa dekade perempuan bekerja mengalami peningkatan. Penelitian dilakukan pada keluarga yang memiliki anak usia di bawah 18 tahun. Pada tahun 1940 perempuan bekerja hanya berjumlah 8,6 %. Tahun 1946 naik menjadi 18,2%, tahun 1956 naik lagi menjadi 27,5%, 1966 mengalami kenaikan 35.5%, 1976 naik menjadi 48,8%, 1986 naik lebih besar menjadi 62,5%, dan pada tahun 1996 mencapai angka 70%. Perubahan dan peningkatan data ibu bekerja juga sangat berpengaruh pada status keluarga, status perempuan, sikap

orang tua dan perspektif gender anak (persepsi anak tergantung pada jenis kelamin dan efek besar terjadi pada keluarga menengah ke bawah (Hoffman, p.2).

Hoffman menemukan tentang perbedaan dampak ibu bekerja dan ibu tidak bekerja pada anak yaitu: (1) Anak perempuan yang memiliki Ibu bekerja ditemukan bahwa anak perempuan tersebut memiliki prestasi akademik yang lebih tinggi, karir yang lebih sukses, memiliki banyak pilihan karir nontradisional dan memiliki komitmen pekerjaan; (2) pada anak yang berasal dari keluarga miskin baik dari orang tua lengkap maupun *single parent*, ditemukan bahwa anak yang memiliki ibu bekerja memiliki skor kognitif yang lebih tinggi serta indikator sosial emosional yang lebih baik; (3) penemuan penelitian pada akhir-akhir tahun ditemukan bahwa pada keluarga menengah yang memiliki anak laki-laki dan ibu bekerja, anak laki-laki memiliki *performance* yang lebih rendah dibandingkan dengan anak yang ibunya tidak bekerja.

Hasil penelitian Hoffman (2008, p.3) menyebutkan bahwa ibu bekerja yang menikmati pekerjaannya memiliki interaksi positif dengan anaknya, lebih simpatik dan lebih sedikit memperlihatkan kemarahannya dalam situasi mendisiplinkan anak. Hal ini terjadi karena ibu bekerja merasa bersalah ketika meninggalkan anak dan berusaha menjadi ibu yang baik secara berlebihan. Anak merasakan pola asuh yang terlalu melindungi. Di sisi lain ibu yang tidak menikmati pekerjaannya cenderung menggunakan metode yang keras dalam mendidik anak, memperlihatkan sedikit kasih sayang dibanding dengan ibu-ibu yang dapat menikmati pekerjaannya.

Penelitian lain terkait dengan pekerjaan dan pendidikan orang tua adalah hasil penelitian dari Charlene yang menemukan bahwa terdapat pengaruh signifikan selain sekolah yaitu anak yang memiliki latar belakang pendidikan tinggi berpengaruh terhadap kemampuan membaca anak. Hal ini dipengaruhi oleh kebiasaan orang tua dalam hal membaca dan dilihat oleh anaknya. (Hildebrand).

Tahun 1999-1998 di Amerika mulai berkembang *fullday school kindergarten*, bahkan 61% lembaga *Kindergarten* sudah menyediakan layanan *fullday*. (NCE, 2014, p.15). Hal ini berbeda dengan di Indonesia karena konsep PAUD *Fullday* baru berkembang baru tahun 2010-an, bahkan sampai saat ini penyelenggaraan PAUD *Fullday* belum dapat diidentifikasi dan didata oleh pemerintah untuk mengetahui jumlah pasti lembaga yang menyelenggarakan PAUD *Fullday*.

Fenomena ibu bekerja juga terjadi di Indonesia, menurut data statistik Angka partisipasi pendidikan di Indonesia berdasarkan jenis kelamin pada tingkat SMU pada tahun 2011 seimbang antara laki-laki dan perempuan. Perempuan 48,31 % dan laki-laki 47,64%. Hal ini tentu saja akan berdampak pada kesempatan pekerjaan yang lebih besar bagi perempuan.

Adanya peningkatan perempuan bekerja dan perubahan bentuk keluarga yang lebih kecil atau menjadi keluarga inti menyebabkan keluarga membutuhkan mitra dalam pengasuhan anak. Saat ini terutama di kota-kota besar terjadi peningkatan jumlah lembaga *baby day care* dan sekolah-sekolah *fullday*. Menurut data statistik jumlah lembaga PAUD saat ini mencapai ribuan lembaga PAUD. Di Daerah Istimewa Yogyakarta (DIY) lembaga yang memfasilitasi program *full day* dan TPA semakin memiliki beragam jenis dan bentuk layanan. Mulai dari TK *Fullday*, TPA, SD kelas rendah *fullday* yang semuanya memfasilitasi bagi keluarga kecil yang pasangannya bekerja. Hal ini tidak hanya terjadi di Indonesia bahkan studi di Rhode Island pada tahun 2004-2005 menunjukkan bahwa 100% anak-anak mengikuti *fullday Kindergarten Program*.

Program *Fullday Kindergarten* adalah program pendidikan yang memberikan layanan minimum 6 jam dalam sehari yang dimulai pada pukul 08.00 sampai pukul 15.00. Hasil riset menemukan bahwa anak-anak yang mengikuti *fullday kindergarten* memiliki kesiapan dalam mengikuti sekolah dasar, dan memiliki kemampuan membaca jika dibandingkan dengan anak-anak yang mengikuti program *half day*. Anak-anak di *fullday*

memiliki kelebihan dan kesiapan belajar karena menurut hasil penelitian Guru memiliki lebih banyak waktu untuk dapat mengidentifikasi gaya belajar anak, kebutuhan anak serta permasalahan yang dihadapi anak. Dari perspektif anak-anak, mereka lebih banyak memiliki pilihan untuk melakukan berbagai aktivitas, pilihan belajar, dan memperdalam materi karena waktu yang lebih banyak.

Indonesia sebagai negara berkembang masih memiliki keterbatasan tentang penelitian yang berkenaan dengan *fullday* program baik efektivitas, pengaruh terhadap anak serta model-model pembelajaran yang dilakukan. Penelitian ini akan memetakan tentang program-program *full day* yang telah berjalan kemudian mengembangkan berbagai macam model layanan pembelajaran PAUD *Fullday*.

## Metode

Penelitian ini menggunakan pendekatan survei. Lembaga penyelenggara PAUD *Fullday* dari lima Kabupaten se DIY dipetakan dengan menggunakan instrumen kuesioner, wawancara melalui FGD dan studi dokumentasi.

Subjek dalam penelitian ini adalah lembaga-lembaga PAUD *Fullday* yang sudah memberikan layanan *fullday* pada masyarakat. Penentuan *sampling* dengan metode *purposive sampling* yaitu memilih lembaga-lembaga yang memiliki *fullday*. Dalam mengukur ketercapaian standar mutu, kuesioner menggunakan skala satu sampai dengan empat. Komponen yang dikembangkan dalam kuesioner adalah standar pencapaian perkembangan anak TK, isi proses dan penilaian dan standar pengelolaan dan pembiayaan.

## Hasil Penelitian dan Pembahasan

Keluarga merupakan pendidikan utama dan pertama bagi anak. Akan tetapi dengan adanya perubahan sosial yang terjadi pada bentuk keluarga dari keluarga besar (*extended family*) menjadi keluarga inti (*nuclear family*) menyebabkan adanya perubahan pola asuh yang sangat bergantung pada pasangan

suami istri. Hal ini senada dengan yang dikemukakan oleh Faturochman (2001, p.2) yang diambil dari ey ardiner & Gardiner, 1988 selama beberapa dekade terakhir, keluarga mengalami perubahan bentuk dari ukuran keluarga besar menjadi lebih kecil.

Peningkatan wanita bekerja di Indonesia terjadi karena dua faktor yaitu adanya penawaran dan permintaan. Penawaran terjadi karena adanya perbaikan pendidikan di kalangan perempuan, serta adanya perubahan dan pergeseran persepsi masyarakat yang memaklumi wanita bekerja. Penawaran terjadi karena semakin terbukanya lapangan pekerjaan bagi pekerja wanita dan penyebab utama adalah semakin tingginya biaya hidup yang menyebabkan pasangan suami istri harus bekerja, fenomena ini dapat dilihat terutama di kota-kota besar (Tjiptoheriyanto dalam Tjaja, 2000).

Peningkatan wanita bekerja pada kalangan TKI menimbulkan konsekuensi dimana pengasuhan anak diserahkan pada kerabat atau suami (Winda Yunitasari). Sedangkan pada wanita terdidik yang bekerja di sektor industri lebih memilih layanan pengasuhan *baby sitter* atau PAUD *fullday*. PAUD *fullday* membutuhkan kemampuan finansial yang relatif lebih besar. Hal ini sesuai dengan hasil penelitian di Amerika bahwa PAUD *fullday* lebih mahal jika dibandingkan dengan PAUD *half day* atau kelas reguler. *Fullday* lebih mahal karena pembiayaan untuk anak termasuk dalam biaya makan, pengasuhan yang lebih panjang menyebabkan jam kerja guru dan pengasuh lebih panjang, kebutuhan ruangan yang lebih luas (Plucker et all, 2004).

Menurut data statistik nasional jumlah Guru Taman kanak-kanak berjumlah 7633 orang, dengan rasio siswa dan guru di DIY adalah satu berbanding 10 siswa. Semakin kecil usia anak, maka rasio anak dan guru semakin kecil. Jumlah pendidik dan pegasuh di PAUD *fullday* berbanding dengan jumlah anak, dan DIY termasuk dalam kategori ideal. Sehingga biaya PAUD *fullday* lebih mahal jika dibandingkan dengan layanan PAUD reguler.

Salah satu kendala yang dihadapi dalam penelitian adalah sampai saat ini tidak ada data yang menyajikan tentang jumlah PAUD *fullday*, yang tersedia data jumlah lembaga Taman kanak-kanak, kelompok bermain, Taman penitipan anak dan SPS PAUD. (Statistik PAUD 2014).

Hal ini berbeda dengan hasil temuan Plucker tentang *fullday* di Indiana, dimana lembaga PAUD *fullday* sudah memiliki data sejak tahun 1978, kemudian pada tahun 2004 mulai menganalisis tentang dampak *fullday* pada kemampuan akademik anak, motivasi anak, kemampuan membaca, kemampuan matematik dimana anak-anak *fullday* memiliki kemampuan lebih baik jika dibandingkan dengan anak-anak *halfday*. Tetapi disisi finansial *fullday* memiliki biaya yang lebih tinggi (Plucker 2004, p.29).

Data UNESCO tahun 2005 menyebutkan bahwa lembaga PAUD lebih terkonsentrasi didaerah perkotaan, terutama pada layanan kelompok bermain. Sedangkan hasil survey UNESCO pada 200 layanan PAUD terungkap bahwa layanan pendidikan anak usia dini 99% dikelola oleh swasta dan sisanya adalah negeri atau pemerintah. Hal ini berbeda dengan kondisi Malaysia yang jumlah lembaga PAUD anantara pemerintah dan swasta perbandingannya 41% swasta dan lebih banyak pemerintah. (Laporan tahunan UNESCO 2005). Halini semakin menjelaskan bahwa investasi atau perhatian pemerintah terhadap dunia PAUD masih sangat minim. Sehingga jika masyarakat ingin mendapatkan layanan PAUD berkualitas masyarakat harus membayar lebih mahal.

Program *fullday Kindergarten* adalah program pendidikan yang memberikan layanan minimum 6 jam dalam sehari yang dimulai pada pukul 08.00 sampai pukul 15.00. Hasil riset menemukan bahwa anak-anak yang mengikuti full day kindergarten memiliki kesiapan dalam mengikuti sekolah dasar, dan memiliki kemampuan membaca jika dibandingkan dengan anak-anak yang mengikuti program *half day*. Anak-anak di *full day* memiliki kelebihan dan kesiapan belajar karena menurut hasil penelitian,

Guru memiliki lebih banyak waktu untuk dapat mengidentifikasi gaya belajar anak, kebutuhan anak serta permasalahan yang dihadapi anak. Sedangkan dari perspektif anak-anak, mereka lebih banyak memiliki pilihan untuk melakukan berbagai aktivitas, pilihan belajar, dan memperdalam materi karena waktu yang lebih banyak.

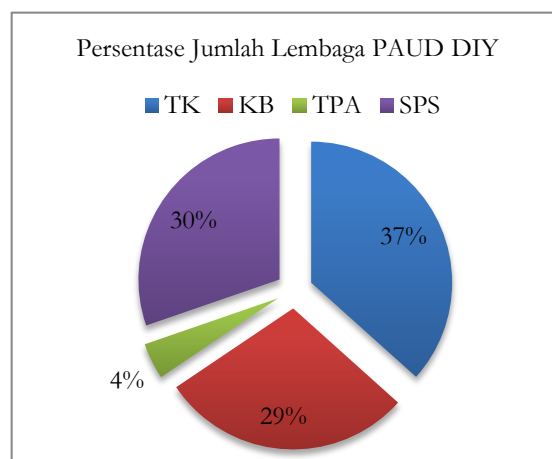
Indonesia sebagai Negara berkembang masih memiliki keterbatasan tentang penelitian yang berkenaan dengan *fullday program* baik efektivitas, pengaruh terhadap anak serta model-model pembelajaran yang dilakukan. Temuan Penelitian ini memetakan tentang program-program *fullday* yang telah dilaksanakan, kemudian mengembangkan berbagai macam pembelajaran yang dilakukan. Lembaga-lembaga PAUD Kelompok Bermain (KB), Taman Penitipan Anak (TPA) dan satuan PAUD Sejenis (SPS) berada di bawah koordinasi Himpaun Pendidik PAUD (HIMPAUDI DIY). HIMPAUDI DIY memiliki visi Terwujudnya pendidik dan tenaga kependidikan anak usia dini Indonesia yang tangguh, profesional dan berakhlak mulia pada tahun 2015, Adapun misi HIMPAUDI: (1) menghimpun pendidik dan tenaga kependidikan anak usia dini Indonesia untuk mencerdaskan bangsa; (2) meningkatkan profesionalisme pendidik dan tenaga kependidikan anak usia dini Indonesia; (3) meningkatkan kesejahteraan pendidik dan tenaga kependidikan anak usia dini Indonesia; (4) menjalin kemitraan dengan berbagai pihak dalam meningkatkan daya juang HIMPAUDI; Untuk jumlah lembaga PAUD di Indonesia tercatat 186.627 lembaga. ()

Menurut data PAUDNI jumlah peserta didik berdasarkan jenis satuan di DIY terbagi dalam empat jenis layanan yaitu TK, KB, TPA dan SPS. Adapun jumlah peserta didik adalah 132.951 yang terbagi atas 67.515 anak laki-laki dan 65.436 anak perempuan. Sedangkan jumlah anak PAUD di DIY 214.785 dan anak yang terlayani di lembaga PAUD sejumlah 210.352 artinya angka partisipasi kasar DIY sebanyak 97%. (Kemendikbud Final APK APM gabungan 2013). APK DIY termasuk APK tertinggi di

Indonesia. Walaupun terdapat perbedaan jumlah peserta didik antara data dari PAUDNI dan data dari Kemendikbud namun memiliki persamaan bahwa jumlah anak di DIY hampir seluruh anak sudah terlayani dari pelbagai satuan pendidikan PAUD yang ada baik TK, KB, TPA maupun SPS.

Adapun secara lebih rinci jumlah seluruh lembaga PAUD di DIY terdapat 5.575 lembaga PAUD yang terdiri dari 80 lembaga negeri dan 5.364 lembaga PAUD swasta. Sedangkan data lembaga PAUD berdasarkan jenis satuan terbagi atas 2042 TK, 1612 KB, 231 TPA dan 1688 SPS. Sehingga jumlah total keseluruhan lembaga adalah 5.575.

Dalam penelitian ini lembaga yang diteliti fokus pada lembaga yang memberikan layanan PAUD *fullday*. Kesulitan tim peneliti adalah data PAUD *fullday* tidak muncul dalam profil propinsi, sehingga pengambilan sampel dilakukan dengan *purposive sampling*. Jadi lembaga-lembaga PAUD yang diidentifikasi memberikan layanan PAUD *fullday* baik dari satuan TK, KB dan TPA diambil untuk dijadikan sampel. Lembaga yang diambil adalah 25 lembaga penyelenggara PAUD *fullday* berikut data pemetaan hasil penelitian.

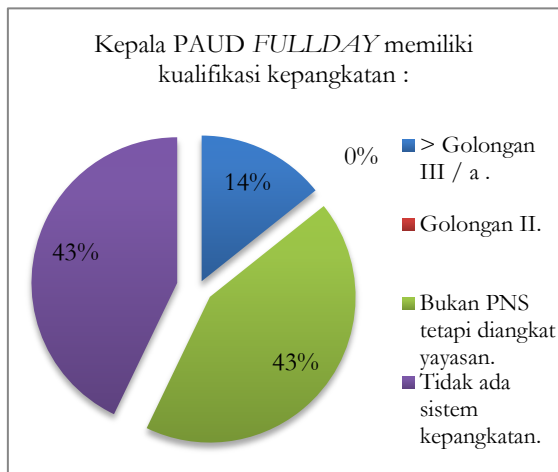


Gambar 1. Persentase Jumlah Lembaga PAUD DIY

#### Pendidik dan Tenaga Kependidikan

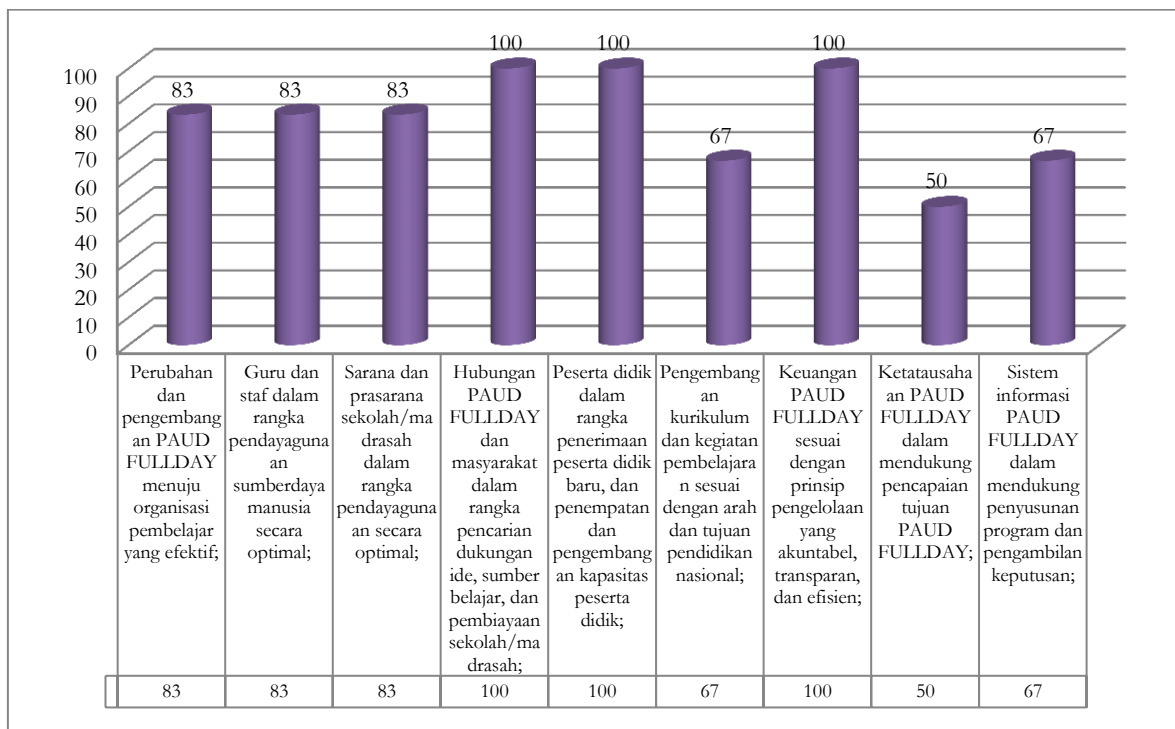
Berdasarkan hasil penelitian ditemukan kualifikasi dan kompetensi pendidik

terutama kepala lembaga PAUD. dilihat dari aspek pengalaman yang menjadi kepala lembaga PAUD 100% memiliki pengalaman lebih dari lima tahun, sedangkan dari sisi kualifikasi pendidikan 57% berpendidikan S1 dan 43 % berpendidikan D3. Dari aspek kepangkatan 14 % responden adalah PNS golongan III dan 43% nonPNS yang diangkat yayasan dan 43% lainnya kepala lembaga yang yayasannya tidak memiliki sistem kepangkatan seperti disajikan pada Gambar 2.

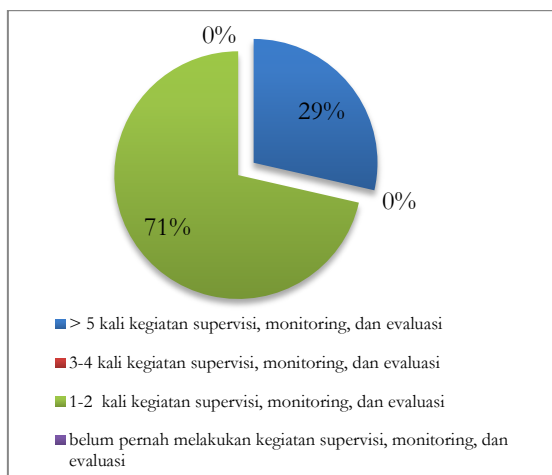


Gambar 2. Kualifikasi Kepangkatan Kepala Lembaga PAUD

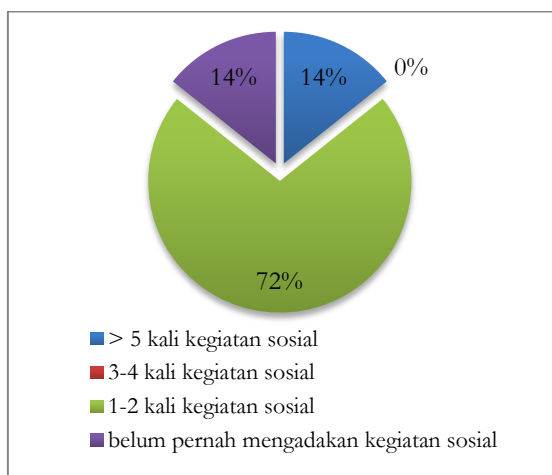
Persepsi guru dan tenaga kependidikan dilihat dari beberapa kompetensi kepala lembaga yang indikatornya terdiri dari: (1) menyusun perencanaan PAUD *fullday* untuk berbagai tingkatan perencanaan kepala lembaga memiliki persepsi yang sangat baik dengan pemberian skor nilai di atas 85; (2) mengembangkan organisasi PAUD *fullday* sesuai dengan kebutuhan mendapatkan persepsi yang cukup dengan skor 65, artinya kemampuan kepala lembaga masih perlu dikembangkan jika dikaitkan dengan pengembangan program untuk memenuhi kebutuhan dari masyarakat; (3) indikator memimpin PAUD *fullday* dalam rangka pendayagunaan sumber daya sekolah/madrasah secara optimal mendapatkan persepsi sangat baik; indikator menciptakan budaya dan iklim sekolah/madrasah yang kondusif dan inovatif bagi pembelajaran peserta didik; memanfaatkan kemajuan teknologi informasi bagi peningkatan pembelajaran dan manajemen PAUD *fullday*; dan melakukan monitoring, evaluasi, dan pelaporan pelaksanaan program kegiatan PAUD *fullday* dengan prosedur yang tepat, serta merencanakan tindak lanjutnya. Untuk lebih detail dapat terlihat pada Gambar 3.



Gambar 3. Kompetensi Manajerial Kepala PAUD *Fullday*



Gambar 4. Intensitas Kepala PAUD *Fullday* melakukan Supervisi, *monitoring*, dan Evaluasi dalam setahun



Gambar 5. Kegiatan Sosial yang Diselenggarakan PAUD *Fullday*

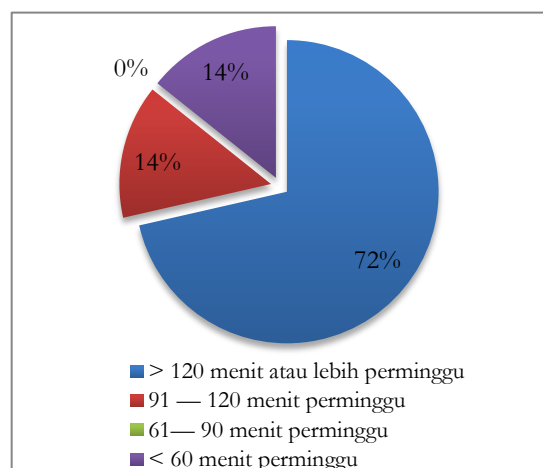
Intensitas Kepala PAUD *fullday* melakukan supervisi, monitoring, dan evaluasi dalam waktu 1 tahun: 29% responden menjawab bahwa kepala lembaga melakukan >5 kali kegiatan supervisi, monitoring, dan evaluasi; 3-4 kali kegiatan supervisi, *monitoring*, dan evaluasi; dan 72% 1-2 kali kegiatan supervisi, monitoring, dan evaluasi; 0% lembaga belum pernah melakukan kegiatan supervisi, monitoring, dan evaluasi. Artinya proses *monitoring* dan supervisi kepala lembaga pada umumnya sudah dilakukan tetapi dalam intensitas yang berbeda tiap tahunnya. Intensitas Kepala PAUD melakukan supervisi, *monitoring*, dan

evaluasi dalam waktu satu tahun disajikan pada Gambar 4.

Kepala PAUD *fullday* mempunyai kompetensi sosial, meliputi: bekerja sama dengan pihak lain untuk kepentingan PAUD *fullday*; berpartisipasi dalam kegiatan sosial masyarakat; dan memiliki kepekaan sosial terhadap orang atau kelompok lain yang ditunjukkan dengan 72% lembaga melakukan kegiatan sosial 1-2 kali dalam setahun, 14% lebih dari lima kegiatan sosial dan 14% melakukan kegiatan sosial 3-4 kegiatan sosial dalam setahun yang disajikan pada Gambar 5.

#### Peta Standar Isi, Proses dan Penilaian

Adapun struktur Program PAUD terdiri dari aspek. Pertama, aspek pengembangan anak. 100% lembaga telah mengembangkan lima aspek yang terdiri dari aspek nilai dan moral, sosial emosional, fisik motorik dan bahasa. Kedua, waktu pelayanan. 72% lembaga memberikan layanan lebih dari 120 menit per minggu, karena lembaga yang di undang adalah lembaga penyelenggaraan PAUD *fullday*. Aspek-aspek tersebut disajikan Gambar 6.

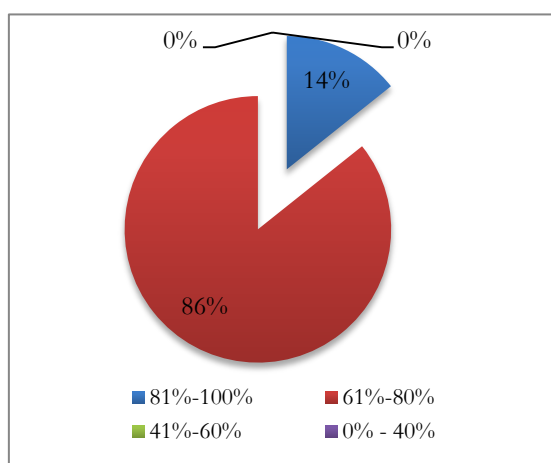


Gambar 6. Waktu Pelayanan PAUD *Fullday*

Struktur program kegiatan di PAUD *fullday* mencakup bidang pengembangan pembentukan perilaku dan bidang pengembangan kemampuan dasar melalui pembiasaan yang dimasukkan dalam program tahunan, semester dan bulanan. Sedangkan

untuk ruang lingkup pengembangan anak 100% lembaga telah mengembangkan aspek nilai, fisik, bahasa, sosial dan emosional.

Pengelompokan usia anak ditentukan oleh jenis layanan pendidikan yang diberikan. Pengelompokan usia adalah 0-3 bulan, 4-9 bulan, 10-16 bulan, 16-24 bulan, >24 bulan. Untuk TPA memberikan layanan 0-16 bulan, sedangkan untuk KB dan TPA memberikan layanan lebih dari 24 bulan. Untuk responden lembaga yang diteliti pada umumnya sudah memiliki tiga layanan PAUD yang terintegrasi 82% yaitu memiliki TPA, KB dan TK sedangkan sisanya fokus pada salah satu layanan yaitu TPA dan KB.



Gambar 7. Jumlah Lembaga yang Sudah Mengembangkan Pembelajaran Tematik

Perbandingan pengasuh atau pendidik di PAUD *fullday* usia 0- kurang dari satu tahun adalah 57% 1:4; 14% 1:6 dan 1:7 14%. Untuk usia 0 - < 1 tahun memiliki rasio pengasuh/pendamping dengan peserta didik bersifat fleksibel sesuai jenis layanan program, dengan jumlah minimal pendidik/pendamping 1 orang. Lebih jelas dapat dilihat dalam bagan 7. Untuk usia 1-2 tahun rasio pengasuh dan anak 72% responden memiliki perbandingan 1: 6 atau lebih kecil, dan 28% memiliki rasio lebih besar yaitu satu pengasuh dengan lebih dari enam anak. Untuk kelompok usia 2-3 tahun 87% responden memiliki perbandingan pengasuh 1 berbanding 8 anak dan 13% lainnya memiliki perbandingan yang lebih tinggi yaitu satu pengasuh dengan lebih dari 8 anak.

Setiap lembaga telah memiliki kalender pendidikan dengan komponen pengaturan kegiatan waktu pembelajaran selama satu tahun, hari libur dan hari penting pendidikan.

Proses pembelajaran dimulai dengan pembuatan rencana kegiatan harian yang dibuat, untuk RKH usia 0-2 tahun sebanyak 71% lembaga memperhatikan karakteristik anak ketika membuat RKH dan memperhatikan kebiasaan setiap anak; RKH hanya dibuat sampai jam pembelajaran yang sangat variatif mulai dari jam delapan pagi sampai dengan jam 10 atau jam 13.00. Setelah kegiatan pembelajaran di kelas selesai maka aktivitas *after school* lebih pada aktivitas bebas untuk anak, dari 25 lembaga hanya 25% lembaga yang telah memadukan aktivitas pembelajaran inti formal dengan kegiatan *after school* sampai dengan penjemputan. Sisanya 14% lembaga Tidak membuat RKH dan menyamakan seluruh kelompok anak dan lebih banyak pada pengasuhan. Sedangkan untuk kelompok satuan kelompok bermain dan TK 100% para pendidik telah membuat RKH dan RKM setiap hari.

Lembaga PAUD *fullday* menerapkan prinsip pembelajaran, meliputi: 71% lembaga memperhatikan tingkat perkembangan, kebutuhan, minat dan karakteristik anak artinya ketika lembaga menyusun rencana pembelajaran pembelajaran disesuaikan dengan tingkat perkembangan, kebutuhan minat dan karakteristik anak, tetapi dari hasil wawancara didapatkan informasi dan data tambahan bahwa pembelajaran yang diobservasi oleh pendidik atau guru terbatas pada jam pembelajaran inti di kelas, setelah selesai kegiatan inti kegiatan tambahan lebih banyak pada proses pengasuhan dan diserahkan pada pengasuh dengan kegiatan bebas. Terkait dengan integrasi kesehatan gizi, pendidikan dan pengasuhan jumlah lembaga yang memadukan dengan aspek gizi dan kesehatan anak baru sebanyak 43% lembaga, artinya kesehatan fisik anak sudah dilakukan dan difasilitasi tetapi tidak dimasukkan dalam lingkup pembelajaran secara interaktif. Cek kesehatan dilakukan secara berkala oleh semua lembaga tetapi hasil

tes kesehatan atau gizi jarang digunakan oleh guru sebagai bahan pembelajaran di kelas. Untuk metode pembelajaran 86% lembaga memakai Pembelajaran melalui metode bermain dan dilakukan secara bertahap, berkesinambungan, dan bersifat pembiasaan. Dari hasil wawancara diketahui bahwa lembaga-lembaga yang telah memberikan layanan PAUD *fullday* menggunakan pelbagai sumber pembelajaran dan kegiatan *outing class* atau *field trip*. Dari hasil penelitian ini diharapkan sesuai dengan teori bermain Hurlock bahwa *outing class* dan bermain menjadi metode pembelajaran yang sangat menarik. Karena bermain menurut Hurlock adalah proses yang membuat anak tahu tentang dunia-meniru, mengeksplorasi, menguji dan membangun (Hurlock 1999).

Dalam pembuatan RKH komponen yang menjadi pertimbangan dalam membuat rencana pembelajaran adalah metode permainan, alat bermain, sumber pembelajaran dan teknik penilaian serta alat penilaian. Alat permainan edukatif yang digunakan diantaranya seperti *playdough*, *puzzle*, boneka, APE dalam, APE luar, menara donat dan balok. Lingkungan sekitar juga sering dimanfaatkan untuk pembelajaran terutama ketika pembelajaran inti.

Terkait dengan keterlibatan orang tua, dari hasil penelusuran dokumentasi 100% orang tua anak yang mengikuti PAUD *fullday* ibu dan bapaknya bekerja dan orang tua yang sedang tugas belajar. Untuk komunikasi dengan orang tua dari hasil wawancara diketahui bahwa komunikasi dilakukan melalui media buku penghubung, sehingga komunikasi dilakukan dengan intensitas harian, bulanan dan tengah semester dan tahunan. Adapun materi yang dikomunikasikan adalah perkembangan anak terkait dengan aspek moral, nilai, sosial, emosional, bahasa dan motorik anak. Adapun intensitas komunikasi 50% lembaga melakukan komunikasi lebih dari empat kali komunikasi dengan orang tua selama satu semester, 33% lembaga berkomunikasi sebanyak dua kali, dan 17% sisanya berkomunikasi secara formal 3 kali dalam satu semester.

Metode penilaian dilakukan melalui catatan anekdot sebanyak 71% lembaga, 57% lembaga membuat catatan pengamatan, dan 43% melakukan ceklis, dan 60% melakukan penilaian melalui portofolio. Penilaian portofolio dilakukan dengan menganalisis hasil kerja anak, pada umumnya pendidik hanya mengumpulkan hasil karya anak tanpa melakukan analisa atau pengamatan terhadap hasil karya anak, tentang pencapaian yang telah dimiliki oleh anak.

Sedangkan materi yang dikomunikasikan orang tua tentang hasil penilaian yang telah dianalisa oleh pendidik sebanyak 72% lembaga hanya dilakukan kurang dari dua kali selama satu semester, artinya komunikasi awal yang ditanyakan peneliti terkait dengan intensitas komunikasi lebih banyak dilakukan terbatas pada aktivitas di kelas namun untuk hasil analisa perkembangan anak masih jarang dilakukan. Disamping itu masih ada 14% lembaga yang melakukan komunikasi dengan orang tua tentang analisa hasil penilaian terkait tumbuh kembang anak dilakukan lebih dari empat kali dalam satu semester, setelah diteliti lebih dalam lembaga tersebut memiliki wahana komunikasi melalui program parenting yang secara berkala telah dilakukan, tidak hanya pada pendidikan anak tetapi dimanfaatkan untuk laporan perkembangan anak bulanan.

Selain dilaporkan kepada orang tua lembaga PAUD memanfaatkan hasil penilaian untuk perbaikan program, perbaikan pembelajaran, peningkatan kompetensi pedagogik pendidik, penyediaan sarana pembelajaran dan kebersihan serta sarana prasarana bagi anak yang berkebutuhan khusus.

## Simpulan dan Saran

### Simpulan

Berdasarkan hasil temuan pemetaan yang dilihat dari empat standar sesuai dengan permen No 58 tahun 2009. Adapun hasil penelitiannya adalah sebagai berikut. Pertama, capaian standar pencapaian perkembangan sudah baik. Kualifikasi kepala sekolah sudah baik rata-rata memiliki kualifikasi pendidikan S1. Standar sarana prasarana sudah baik.

Kedua adalah aspek pembelajaran, materi, penilaian dan pengamatan belum dilakukan secara utuh dan terintegrasi dari pagi sampai sore. Ketiga, capaian perkembangan anak sulit ditemukan dilapangan karena seluruh lembaga PAUD tidak memiliki data rekapan pencapaian perkembangan anak dan deskripsi penilaian. Keempat, materi pembelajaran terbatas pada kegiatan inti dan kegiatan setelah pembelajaran lebih banyak pada pengasuhan, sehingga dari pengembangan model PAUD *Fullday* diperlukan kegiatan pembelajaran utuh dan berkesinambungan dari pagi sampai sore.

Saran.

Berdasarkan simpulan dari hasil penelitian dapat diberikan beberapa saran sebagai berikut: (1) perlu dibuatkan model pembelajaran utuh terintegrasi dan holistic dari pagi sampai sore; (2) perlu dibuatkan format silabi, RKH dan format penilaian sebagai kelengkapan dari model pembelajaran *fullday* yang integratif dan holistic; (3) perlu dibuat design kegiatan pada sore hari sehingga tidak hanya fokus pada pengasuhan tetapi ada proses stimulasi yang terencana dan teramati dengan baik; (4) perlu dibuat format untuk merekap data-data perkembangan siswa agar sekolah memiliki data yang dapat digunakan untuk pengembangan pembelajaran di sekolah

#### Daftar Pustaka

Dewantara, K.H. (1962): *Pendidikan Bagian I*. Yogyakarta: Taman Siswa.

- Faturochman. (2001). Revitalisasi peran keluarga. *Buletin Psikologi, tahun IX No 2, Desember 2001*, 39-47.
- Hurlock, E. B., (1999). *Perkembangan Anak*. Jakarta. Erlangga.
- Hoffman, L. W. 'The effects of the mother's employment on the family and the child.  
<http://parenthood.library.wisc.edu/Hoffman/Hoffman.html>. Di unduh pada tanggal 15 April 2013.
- Hildebrand, C. Effect all-day, half day kindergarten programming on reading, Writing, Math and classroom sosial behaviours. University of Nebraska. Diakses dari <http://nationalforum.com/Electronic%20Journal%20Volumes/Hildebrand,%20Charlene%20Effects%20of%20All-day%20and%20Half-day%20Kindergarten%20Programming.pdf>
- National Center for Education Statistik. (NCES). (2014). Half day and fullday in The United States.
- Plucker, J., et. all. (2004) : *The effect of fullday versus halfday Kindergarten*. Indiana.
- Sudjana, D. (2006). *Evaluasi Program pendidikan Luar Sekolah Untuk pendidikan Non Formal dan Pengembangan Sumber Daya Manusia*. Bandung: PT Remaja Rosda Karya.
- Tjaja, R.P. Naskah No 20 tahun (2004): Perempuan bekerja dan implikasi sosial. Bappenas Pusat Jakarta.
- UNESCO (2005) : Laporan tahunan PAUD di Indonesia.

## FAKTOR KOGNITIF DAN NON-KOGNITIF PADA SELEKSI MAHASISWA BARU SEBAGAI PREDIKTOR TERHADAP PRESTASI AKADEMIK

<sup>1)</sup>Tissa Octavira Permatasari, <sup>2)</sup>Yayi Suryo Prabandari, <sup>3)</sup>Tri Nur Kristina

<sup>1)</sup>Fakultas Kedokteran Universitas Swadaya Gunung Jati, Cirebon

<sup>2)</sup>Fakultas Kedokteran Universitas Gadjah Mada, Yogyakarta

<sup>3)</sup>Fakultas Kedokteran Universitas Diponegoro, Semarang  
tissaoctavira@yahoo.com

### Abstrak

Proses penerimaan mahasiswa baru di Fakultas Kedokteran Universitas Swadaya Gunung Jati (FK Unswagati) lebih menekankan pada hasil tes seleksi akademik sebagai faktor kognitif, sedangkan faktor non-kognitif kurang diperhitungkan. Penelitian ini bertujuan menganalisis faktor kognitif dan non-kognitif pada seleksi mahasiswa baru sebagai prediktor prestasi akademik. Metode yang digunakan adalah potong lintang dengan sampel total berjumlah 97 orang terdiri dari mahasiswa angkatan 2009 dan 2010. Prestasi akademik diukur melalui IPK S.ked, rerata nilai blok, dan *Objective Structured Clinical Examination* (OSCE) komprehensif. Tes seleksi akademik memiliki hubungan yang bermakna ( $r=0,40$ ) dan mempengaruhi IPK sebesar 16% dan nilai blok sebesar 11,9%, sedangkan tes seleksi akademik tidak berhubungan secara bermakna dengan nilai OSCE komprehensif. Karakteristik psikologi tidak memiliki hubungan yang bermakna dengan ketiga variabel prestasi akademik, tetapi motivasi dan strategi belajar menunjukkan pengaruh terhadap IPK sebesar 19,3% dan 30,6%. Maka, tes seleksi akademik memiliki *predictive validity* terhadap prestasi akademik, dan karakteristik psikologi dapat meningkatkan kekuatan prediksi terhadap prestasi akademik.

**Kata kunci:** prestasi akademik, non-kognitif, kognitif, predictive validity

## COGNITIVE AND NON-COGNITIVE FACTOR IN STUDENT ADMISSION AS PREDICTOR OF ACADEMIC ACHIEVEMENT

<sup>1)</sup>Tissa Octavira Permatasari, <sup>2)</sup>Yayi Suryo Prabandari, <sup>3)</sup>Tri Nur Kristina

<sup>1)</sup>Fakultas Kedokteran Universitas Swadaya Gunung Jati, Cirebon

<sup>2)</sup>Fakultas Kedokteran Universitas Gadjah Mada, Yogyakarta

<sup>3)</sup>Fakultas Kedokteran Universitas Diponegoro, Semarang

### Abstract

The admission process at FM SGJU is more emphasis on academic test results, whereas non-cognitive factors are still underestimated. The aim of this study is to analyze the cognitive and non-cognitive factors in the admission process of new student as predictor of academic achievement. The research method was cross sectional. The total sample consist of 97 students of batch 2009 and 2010. Academic achievement was measured through the GPA, average block grade, and comprehensive OSCE score. Academic test has significant correlation ( $r=0,40$ ) and affects 16% of the GPA and 11,9% of average block grade, while the test was not significantly associated with comprehensive OSCE score. Psychological characteristics have no significant association with the academic achievement, in the other hand motivation and learning strategies show the effect on the GPA amounted to 19,3% and 30,6%. Academic test on the admission process has predictive validity for academic achievement and psychological characteristics can improve the predictive power of academic achievement.

**Keywords:** academic achievement, non-cognitive, cognitive, predictive validity



## Pendahuluan

Prestasi akademik merupakan salah satu indikator keberhasilan proses pendidikan. Hasil belajar tersebut dapat berupa pengetahuan/kognitif, sikap/afektif dan keterampilan psikomotor (Kumara, 1990). Prestasi akademik dipengaruhi oleh 3 faktor utama, yaitu: konteks, *input* mahasiswa serta proses perkuliahan (Cassidy & Eachus, 2000; Bruinsma, 2003; Diseth & Martinsen, 2003). Berbagai institusi pendidikan kedokteran di seluruh dunia memiliki metode dan kriteria yang berbeda dalam memilih kandidat yang ideal (Ranasinghe *et al.*, 2012). Seleksi mahasiswa pada umumnya dilakukan untuk 2 tujuan, yaitu untuk menyesuaikan jumlah calon mahasiswa dengan tempat yang tersedia dan untuk mendapatkan mahasiswa yang dapat mengikuti program pendidikan yang sulit dengan sukses dan kemudian dapat menjadi anggota profesi yang efektif (Wilkinson *et al.*, 2008).

Banyak penelitian yang telah membuktikan bahwa kemampuan kognitif dan capaian akademik sebelumnya merupakan faktor prediktor kesuksesan mahasiswa dalam menjalani proses pendidikan (Ferguson *et al.*, 2002). Beberapa penelitian telah mencari hubungan antara gaya belajar, wawancara dan faktor non-kognitif lainnya dengan kesuksesan di bidang akademik dengan berbagai variasi hasil yang membutuhkan penelitian lebih lanjut (Ferguson *et al.*, 2002). Sebuah penelitian prospektif membuktikan bahwa tingkat kedewasaan dan motivasi intrinsik memiliki pengaruh yang signifikan (Frischenschlager *et al.*, 2005). Sebuah sistem yang mengkombinasikan antara capaian akademik sebelumnya (*prior knowledge*) dan faktor non-kognitif dapat membantu meningkatkan proses seleksi dan mendeteksi kegagalan lebih dini (Ranasinghe *et al.*, 2012).

Penelitian yang telah dilakukan sebagian besar menitikberatkan pada kekuatan prediksi faktor kognitif terhadap kesuksesan mahasiswa pada tahap sarjana kedokteran dan belum banyak yang memperhatikan pengaruh kriteria seleksi mahasiswa terhadap kompetensi mahasiswa setelah lulus sarjana.

Penelitian yang menganalisis pengaruh faktor non-kognitif terhadap kesuksesan mahasiswa masih harus dikembangkan (Ferguson *et al.*, 2002). Aspek kompetensi yang diukur dalam penelitian hanya berbasis pengetahuan, sehingga keterampilan klinik mahasiswa kedokteran sebagai salah satu kompetensi penting tidak dapat dinilai hubungannya dengan tes seleksi mahasiswa (Farrokhi-Khajeh-Pasha *et al.*, 2012).

Fakultas Kedokteran Unswagati Cirebon merupakan institusi pendidikan kedokteran swasta yang menyelenggarakan seleksi penerimaan mahasiswa baru menggunakan beberapa kriteria, yaitu: tes seleksi akademik, tes psikologi, pemeriksaan fisik, dan wawancara. Tes seleksi akademik terdiri dari 100 soal yang mewakili 5 disiplin ilmu (masing-masing 20 soal) yaitu biologi, kimia, fisika, matematika, dan bahasa Inggris. Tes psikologi dilakukan untuk mengetahui karakteristik individu mahasiswa yang merupakan faktor non-kognitif. Karakteristik individu mahasiswa diukur menggunakan tes intelegensi yang mengukur potensi kecerdasan, tes grafis mengukur penyesuaian diri, kemampuan bersosialisasi, dan tes kemampuan kerja mengukur motivasi, daya tahan terhadap stres, dan ketelitian.

Meskipun demikian, dalam proses penerimaan mahasiswa baru tersebut lebih menekankan pada hasil tes seleksi akademik sebagai faktor kognitif, sedangkan hasil tes lainnya sebagai faktor non-kognitif kurang diperhitungkan. Faktor kognitif lainnya seperti pencapaian akademik pada tingkat SMA tidak mendapatkan perhatian dalam proses seleksi penerimaan mahasiswa baru di fakultas tersebut. Hasil evaluasi di FK Unswagati selama 4 tahun terakhir menunjukkan prestasi akademik mahasiswa setiap angkatan rata-rata mengalami penurunan setelah memasuki tahapan yang membutuhkan tingkatan proses berpikir yang lebih tinggi.

Berdasarkan uraian tersebut, diperlukan suatu penelitian yang menganalisis prediksi dari kriteria seleksi mahasiswa, baik dari aspek kognitif yang mencakup pencapaian akademik di SMA maupun tes seleksi

bidang akademik serta aspek non-kognitif terhadap prestasi akademik mahasiswa kedokteran, baik dari segi kognitif maupun keterampilan klinik.

### Metode Penelitian

Rancangan penelitian ini adalah belah lintang (*cross-sectional*). Penelitian ini dilakukan terhadap mahasiswa Fakultas Kedokteran Unswagati Cirebon angkatan 2009 dan 2010 yang telah menyelesaikan pendidikan tahap sarjana dan telah mengikuti OSCE komprehensif. Metode *sampling* yang digunakan adalah *total sampling*, yaitu seluruh populasi Mahasiswa Fakultas Kedokteran Unswagati Cirebon angkatan 2009 sebanyak 27 orang dan angkatan 2010 sebanyak 70 orang.

Data sekunder dalam penelitian ini berupa nilai tes seleksi akademik didapatkan dari soal tes sejumlah 100 soal pilihan ganda dengan 5 pilihan jawaban dan 1 jawaban benar. Setiap jawaban benar mendapatkan skor 1 dan jawaban salah tidak mendapatkan pengurangan. Nilai rapor SMA didapatkan dari rerata nilai mahasiswa dalam disiplin ilmu matematika, biologi, fisika, kimia, dan bahasa Inggris selama 3 tahun di SMA. Nilai blok didapatkan dari ujian tulis blok yang berupa pilihan ganda dengan 5 pilihan jawaban dan 1 jawaban benar. Setiap jawaban benar mendapatkan skor 1 dan jawaban salah tidak mendapatkan pengurangan. Ujian tulis blok dilakukan dalam bentuk *Multiple Choice Questions (MCQs)*. Validasi isi ujian tulis berdasarkan *Blueprint Assessment* untuk level kognitif berdasarkan taksonomi Bloom yang telah disiapkan sebelumnya. Validasi konstruk dan reliabilitas ujian tulis blok dapat dilihat dari hasil *item analysis* tiap blok. Nilai IPK didapatkan dari angka mutu mahasiswa selama 7 semester, didapatkan dari mengalikan bobot sks tiap blok dengan angka mutu tiap blok. Nilai OSCE Komprehensif didapatkan dari rerata nilai 12 station yang dinilai menggunakan rubrik. Rubrik dan soal OSCE komprehensif dikembangkan berdasarkan *Blueprint Assessment* yang telah dibuat se-

belumnya dan direviu oleh tim materi OSCE komprehensif.

Pengumpulan data tentang motivasi mahasiswa dilakukan secara primer menggunakan *Archer's Health Professions Motivation Survey* terdiri dari 41 pertanyaan yang memuat tentang orientasi belajar tuntas, yang telah diterjemahkan dan divalidasi. Pada penelitian yang dilakukan oleh Marasabessy (2011) tentang peranan motivasi berprestasi terhadap proses kognitif, dilakukan uji coba kepada 30 orang mahasiswa tingkat 2 jurusan kesehatan lingkungan dan didapatkan hasil koefisien alpha 0,7115. Namun, dikarenakan ada perbedaan konteks dalam penelitian ini maka akan dilakukan uji coba kepada 30 orang mahasiswa FK Unswagati tingkat 2 untuk mengukur validitas dan reliabilitas instrumen *Archer's Health Professions Motivation Survey* yang telah dimodifikasi.

Pengumpulan data tentang gaya belajar dilakukan secara primer menggunakan kuesioner *Learning Style* dari Honey & Mumford yang terdiri dari 80 pertanyaan sebagai instrumen. Instrumen ini terdapat dalam bahasa Inggris sehingga akan dilakukan adaptasi dan pengukuran validitas serta reliabilitas, melalui beberapa tahap sebagai berikut: (a) kuesioner diterjemahkan ke dalam bahasa Indonesia dengan berkonsultasi dengan penterjemah dari Fakultas Keguruan Bahasa Inggris Unswagati dan ahli pendidikan kedokteran, (b) kuesioner yang telah diadaptasi akan dilakukan uji coba kepada 30 orang mahasiswa FK Unswagati tingkat 2. Kuesioner akan diperbaiki dan diuji coba ulang hingga valid dan reliabel.

Pengumpulan data tentang strategi belajar mahasiswa dilakukan secara primer menggunakan ASSIST yang terdiri dari 52 pertanyaan yang diterjemahkan dan divalidasi. Kuesioner ASSIST telah melalui uji realibilitas, pada penelitian yang dilakukan oleh Rochmawati (2011) yang mengukur strategi belajar pada mahasiswa keperawatan di Yogyakarta, hasil estimasi reliabilitas ASSIST pada penelitian itu memiliki koefisien alpha 0,912. Hasil tersebut didapatkan melalui proses penerjemahan dari bahasa Inggris ke dalam bahasa Indonesia dan uji

coba kepada 26 orang mahasiswa keperawatan di sebuah sekolah tinggi keperawatan. Berdasarkan hal tersebut, kuesioner ASSIST yang telah diterjemahkan dapat digunakan pada penelitian ini.

Karakteristik psikologi diambil dari hasil tes psikologi yang dilakukan oleh mahasiswa pada waktu seleksi masuk perguruan tinggi. Tes psikologi yang dilakukan berupa tes intelegensi yang mengukur potensi kecerdasan, tes grafis mengukur penyesuaian diri, kemampuan bersosialisasi, dan tes kemampuan kerja mengukur motivasi, daya tahan terhadap stres, dan ketelitian. Instrumen yang digunakan adalah Test Intelligensi Kolektif Indonesia – Pendidikan Tinggi (TIKI-T), Instruksi WZT, Baum, Pauli dan EPPS. Hasil pengukuran dirangkum dalam bentuk psikogram dan penilaian dikelompokkan menjadi disarankan, dipertimbangkan dan tidak disarankan.

### Hasil Penelitian dan Pembahasan

Penelitian ini dilakukan pada mahasiswa angkatan 2009 dan angkatan 2010 yang telah menyelesaikan pendidikan tahap akademik. Mahasiswa angkatan 2009 seluruhnya telah menyelesaikan pendidikan tahap akademik yaitu sebanyak 27 orang, sedangkan mahasiswa angkatan 2010 yang telah menyelesaikan pendidikan tahap akademik sebanyak 51 orang dari 70 orang. Oleh karena itu, data primer mengenai motivasi belajar, strategi belajar, dan gaya belajar didapatkan dengan membagikan kuesioner kepada 78 mahasiswa dengan *response rate* 100%.

Jumlah responden laki-laki dan perempuan tampak dalam tabel frekuensi berikut. Jumlah responden laki-laki sebanyak 33 orang dan responden perempuan sebanyak 45 orang, tersebar dengan berbagai rentang usia mulai dari 21 tahun sampai dengan 26 tahun.

Hasil analisis deskriptif nilai tes seleksi akademik 78 mahasiswa tersebut menunjukkan nilai paling rendah 15 dan paling 47 dari nilai tes maksimum 100, dengan rerata 27,99. Nilai rerata rapor SMA minimum 64 dan maksimum 90, dengan rerata 79,92.

Tabel 1. Distribusi Responden Berdasarkan Jenis Kelamin dan Usia

Usia	Jenis Kelamin		Total
	Laki – laki	Perempuan	
21	3	2	5
22	13	17	30
23	10	19	29
24	5	7	12
25	1	0	1
26	1	0	1
Total	33	45	78

Hasil analisis statistik terhadap soal tes akademik angkatan 2009 memiliki koefisien reliabilitas alpha 0,923, rata-rata tingkat kesukaran 0,224, dan rata-rata tingkat daya pembeda (*mean biserial*) 0,475. Hasil analisis statistik terhadap soal tes akademik angkatan 2010 gelombang 1 memiliki koefisien reliabilitas alpha 0,872, rata-rata tingkat kesukaran 0,213, dan rata-rata tingkat daya pembeda (*mean biserial*) 0,376. Hasil analisis statistik terhadap soal tes akademik angkatan 2010 gelombang 2 memiliki koefisien reliabilitas alpha 0,781, rata-rata tingkat kesukaran 0,282, dan rata-rata tingkat daya pembeda (*mean biserial*) 0,273. Hal ini menunjukkan soal tes seleksi yang diberikan tergolong sukar dan memiliki daya pembeda yang rendah. Hasil tes karakteristik psikologi mahasiswa menunjukkan 53 orang masuk dalam kriteria yang tidak disarankan, 21 orang masuk kriteria dipertimbangkan, dan hanya 4 orang yang masuk kriteria disarankan.

Motivasi belajar diteliti dengan menggunakan instrumen *Archer's Health Professions Motivation Survey* yang sebelumnya telah dilakukan uji coba kepada 50 mahasiswa fakultas kedokteran tingkat 2 dan 3. Hasil estimasi reliabilitas untuk item kuesioner yang mengukur orientasi pembelajaran tuntas (*mastery learning*) menunjukkan nilai *Cronbach's alpha* 0,843. Hasil estimasi reliabilitas untuk item kuesioner yang mengukur orientasi pembelajaran pada penampilan (*performance*) menunjukkan nilai *Cronbach's alpha* 0,826. Hasil estimasi reliabilitas untuk item kuesioner yang mengukur orientasi pembelajaran pada persyaratan akademik (*academic alienation*) menunjukkan nilai *Cronbach's alpha*

0,765. Hasil pengukuran skor motivasi belajar menggunakan instrumen tersebut dibagi menjadi 3 yaitu orientasi pembelajaran tuntas (*mastery learning*), orientasi pada penampilan (*performance*), dan orientasi pemenuhan syarat akademik (*academic alienation*). Hasil analisis motivasi berprestasi mahasiswa menunjukkan 67,9% memiliki orientasi belajar untuk penampilan, sedangkan 32,1% memiliki orientasi belajar tuntas.

Gaya belajar diteliti menggunakan instrumen *Learning Style* dari Honey & Mumford yang telah dilakukan uji coba. Uji coba dilakukan kepada 50 mahasiswa fakultas kedokteran tingkat 2 dan 3. Hasil reliabilitas kuesioner *Learning Style* dari Honey & Mumford memiliki koefisien *Cronbach's alpha* 0,752. Hasil analisis gaya belajar mahasiswa menunjukkan 42 orang memiliki gaya belajar reflektor, sedangkan sisanya terbagi secara merata pada 3 macam gaya belajar lainnya. Strategi belajar diteliti menggunakan instrumen ASSIST yang telah dilakukan pada penelitian terdahulu (koefisien alpha 0,912). Hasil analisis strategi belajar menunjukkan sebagian besar 83,3% menggunakan pendekatan belajar strategik, sedangkan yang menggunakan *deep learning* 8% dan *surface learning* 5%.

Hasil analisis korelasi antara variabel yang mewakili faktor kognitif dan non-kognitif dengan prestasi akademik dirangkum dalam Tabel 2.

Hubungan antara nilai tes seleksi akademik dengan nilai IPK dan rerata nilai blok dianalisis menggunakan korelasi *Pearson*. Hasil analisis menunjukkan nilai tes seleksi akademik memiliki hubungan yang bermakna dengan nilai IPK S.ked ( $r = 0,40$  dan  $p=0,0$ ) dan memiliki hubungan yang lemah dan bermakna dengan rerata nilai blok ( $r = 0,345$  dan  $p=0,002$ ). Sementara itu, hubungan antara nilai tes seleksi akademik dengan nilai OSCE komprehensif dianalisis menggunakan uji korelasi *Spearman*. Hasil analisis menunjukkan nilai tes akademik tidak memiliki hubungan yang bermakna dengan nilai OSCE komprehensif ( $r_s = 0,144$  dan  $p= 0,210$ ). Sementara itu, hasil analisis menunjukkan karakteristik psi-

kologi dan rerata nilai rapor SMA tidak memiliki hubungan yang bermakna dengan nilai IPK S. Ked, rerata nilai blok, dan OSCE komprehensif.

Tabel 2. Korelasi Antarafaktor Kognitif dan Non-kognitif dengan Prestasi Akademik

Variabel	Prestasi Akademik	r/r <sub>s</sub>	Sig. (p)
Nilai tes akademik	Nilai IPK	.400	.000*
	Rata-rata nilai IP semester	.345	.002*
	OSCE Komprehensif	.144	.210
Hasil psikotes	IPK S.Ked	.173	.130
	OSCE kompre	.161	.159
	Rerata nilai blok	.214	.061
Nilai rata-rata rapor	IPK S.Ked	.110	.336
	OSCE kompre	-.009	.938
	Rerata nilai blok	.125	.277
Jenis Kelamin	IPK S.Ked	.002	.984
	OSCE komprehensif	.069	.550
	Rerata nilai blok	-.018	.877
Usia	IPK S.Ked	-.243*	.032*
	OSCE komprehensif	.019	.867
	Rerata nilai blok	-.033	.774
Gaya Belajar	IPK S.Ked	-.195	.088
	OSCE komprehensif	-.033	.775
	Rerata nilai blok	-.209	.066
Motivasi belajar	IPK S.Ked	.450**	.000*
	OSCE kompre	.215	.059
	Rerata nilai blok	.387**	.000*
Strategi belajar	IPK S.Ked	.470**	.000*
	OSCE kompre	.221	.051
	Rerata nilai blok	.433**	.000*

Tabel 3 menunjukkan jenis kelamin dan gaya belajar tidak memiliki hubungan yang bermakna dengan prestasi akademik. Sementara itu, usia memiliki hubungan yang bermakna dengan IPK S.Ked. Motivasi belajar dan strategi belajar memiliki hubungan yang bermakna dengan IPK S.Ked dan rerata nilai blok. Namun demikian, tabel

tersebut menunjukkan tidak ada variabel yang memiliki hubungan yang bermakna dengan nilai OSCE komprehensif.

Hasil analisis data sebelumnya menunjukkan beberapa faktor kognitif dan non-kognitif memiliki hubungan secara bermakna dengan prestasi akademik, sehingga untuk mengetahui seberapa besar pengaruh faktor tersebut terhadap prestasi akademik dilakukan uji regresi linier.

Tabel 3. Hasil Uji Regresi Linier Faktor Kognitif dan Non-kognitif masing-masing terhadap IPK S.Ked

Variabel	R <sup>2</sup>	Sig. (p)
Nilai Tes Seleksi Akademik	0,160	0,000*
Usia	0,048	0,055
Motivasi belajar	0,193	0,000*
Strategi belajar	0,306	0,000*

Hasil analisis regresi linier menunjukkan bahwa yang memiliki pengaruh paling kuat terhadap nilai IPK adalah strategi belajar, dimana nilai  $R^2 = 0,306$  yaitu strategi belajar berpengaruh terhadap 30,6% dari nilai IPK S.Ked, diikuti oleh motivasi belajar (19,3%). Sementara itu, nilai tes seleksi akademik berpengaruh terhadap 16% dari nilai IPK S.Ked. Pengaruh yang lebih besar terhadap IPK S.Ked. didapatkan secara signifikan ketika nilai tes seleksi akademik dianalisis bersama usia, motivasi belajar dan strategi belajar, dimana nilai  $R^2=0,433$  dan  $p=0,000$ .

Hasil analisis regresi linier terhadap rerata nilai blok menunjukkan bahwa variabel yang memberikan pengaruh lebih kuat adalah strategi belajar dengan nilai  $R^2=0,251$  dan  $p=0,000$ , maka disimpulkan bahwa variabel rerata nilai blok sebanyak 25,1% dipengaruhi oleh strategi belajar. Berikutnya, variabel motivasi belajar memberikan pengaruh sekitar 17%. Variabel tes seleksi akademik memberikan pengaruh sebesar 11,9% dan hasil psikotes memberikan pengaruh sebesar 5,1%. Hasil analisis tersebut juga menunjukkan ketika menggabungkan variabel nilai tes seleksi akademik, hasil psikotes, motivasi belajar, dan strategi belajar

meningkatkan kekuatan pengaruh terhadap nilai blok dengan nilai  $R^2=0,347$  dan  $p=0,000$ .

Tabel 4. Hasil Uji Regresi Linier Faktor Kognitif dan Non-kognitif secara Bersama-sama terhadap IPK S.Ked

Variabel	R <sup>2</sup>	Sig. (p)
Hasil Psikotes Nilai tes seleksi akademik Rerata nilai rapor	0,161	0,005*
Nilai tes seleksi akademik Motivasi belajar Strategi belajar	0,420	0,000*
Nilai tes seleksi akademik Usia	0,211	0,000*
Usia Nilai tes seleksi akademik Motivasi belajar Strategi belajar	0,433	0,000*

Sebuah revidu sistematik terhadap faktor pada seleksi akademik yang menjadi prediktor keberhasilan dalam pendidikan, meliputi faktor kognitif yaitu kemampuan akademik sebelumnya, dan nonkognitif meliputi kepribadian, gaya belajar, wawancara, dll. Dalam penelitian tersebut kemampuan akademik sebelumnya diukur melalui suatu tes seleksi akademik, dan hasilnya menunjukkan bahwa tes seleksi akademik mempengaruhi sekitar 9% dari prestasi akademik mahasiswa (Ferguson *et al.*, 2002).

Tes yang terstandar untuk mengukur kemampuan kognitif, contohnya tes seleksi akademik, merupakan salah satu prediktor yang kuat dan konsisten terhadap performa baik dalam pendidikan maupun pekerjaan. Skor tes menunjukkan kemampuan yang telah berkembang, dan fungsi dari potensi, pengetahuan dan kemampuan sebelumnya, serta faktor lingkungan yang mempengaruhi perkembangan pengetahuan dan kemampuan, misalnya kesempatan untuk memperoleh pendidikan sebelumnya. Hasil beberapa penelitian menunjukkan nilai korelasi yang rendah (0,20) antara tes seleksi akademik dengan prestasi akademik, meskipun demikian nilai tersebut akan memberikan

informasi yang bernilai. Demi kepentingan akademik maupun pekerjaan, kekuatan prediksi dari skor tes seleksi akademik lebih kuat dan memiliki akurasi yang lebih baik dibandingkan alat ukur lain seperti nilai rapor atau wawancara (Kuncell & Hezlett, 2011).

Kekuatan prediksi dari tes seleksi akademik telah banyak diteliti, namun dari mana asalnya kekuatan prediksi ini masih diperdebatkan. Teori dan penelitian mengindikasikan bahwa tes adalah alat ukur yang berharga karena pengukuran terhadap kemampuan pengetahuan dan keterampilan dapat memprediksi cara orang tersebut akan belajar dan berkembang di masa depan.

Hasil penelitian metaanalisis terhadap hasil pelatihan menunjukkan bahwa kemampuan kognitif mempengaruhi akuisisi pengetahuan dan keterampilan selama pelatihan, sehingga orang yang belajar lebih baik selama pelatihan akan melakukan tugas lebih baik dengan menerapkan hasil selama pelatihan pada waktu melakukan tugas. Faktor situasional seperti pengalaman sebelumnya, pendidikan, dan pelatihan juga akan mempengaruhi tingkat pengetahuan dan kemampuan seseorang saat ini. Berdasarkan hal ini, maka tes kemampuan kognitif seperti tes seleksi akademik dapat memprediksi prestasi karena dapat meramalkan sejauh mana kemampuan seseorang saat ini dan kemampuannya untuk terus berkembang mendapatkan pengetahuan dan keterampilan spesifik yang dibutuhkan untuk melakukan performa yang efektif dalam pendidikan dan pekerjaan (Arnold *et al.*, 2011). Namun, meskipun tes kemampuan kognitif dapat memprediksikan keberhasilan mahasiswa dari segi akademik, ia gagal memprediksikan keberhasilan dalam keterampilan klinik, karena yang diukur hanya kemampuan intelektual (Adebayo, 2008).

Menurut beberapa penelitian terakhir, faktor yang memberikan kontribusi terbesar bagi keberhasilan mahasiswa adalah faktor kognitif, termasuk nilai rapor SMA. Menurut Tinto (1993, *cit.* Adebayo, 2008), performa dan keberhasilan akademik mahasiswa dipengaruhi oleh karakteristik maha-

siswa yaitu tingkat persiapan yang dilalui ketika SMA dan tes seleksi akademik (Parry *et al.*, 2006). Namun, hasil yang didapatkan dalam penelitian ini tidak sejalan dengan pernyataan tersebut karena nilai rerata rapor SMA tidak memiliki korelasi yang bermakna dengan prestasi akademik mahasiswa. Hasil penelitian Adebayo (2008) menunjukkan bahwa nilai rapor SMA menunjukkan korelasi dengan IPK sebesar  $r = 0,21$ ,  $p = 0,01$ , serta mempengaruhi IPK semester pertama mahasiswa sebesar 37%. Penelitian tersebut juga menunjukkan lemahnya korelasi antara nilai rapor SMA dengan tes seleksi akademik ( $r = 0,10$ ) karena fakta bahwa nilai rapor SMA juga tergantung pada kualitas SMA asal mahasiswa (Parry *et al.*, 2006). Hal ini juga dijumpai pada penelitian ini dimana asal mahasiswa tersebar dari berbagai daerah di Indonesia, sehingga kualitas asal SMA dapat mempengaruhi nilai rerata rapor mahasiswa.

Kemampuan kognitif bukan satu-satunya penentuan prestasi akademik dan performa dalam pekerjaan. Kekuatan prediksi terhadap prestasi berdasarkan skor tes kognitif dapat ditingkatkan dengan menambahkan kepribadian, nilai, minat, dan kebiasaan, tetapi hanya bila hal-hal tersebut dipilih dan dikembangkan secara hati-hati. Sedikit sekali literatur yang memberikan panduan tentang cara terbaik untuk mengukur karakteristik tersebut. Sebagian besar institusi menggunakan kombinasi dari beberapa metode untuk mengukur karakteristik psikologi, meskipun tidak ada pendekatan yang terstandar untuk hal tersebut (Adam *et al.*, 2012).

Dalam penelitian ini pengukuran terhadap karakteristik psikologi menggunakan kombinasi dari berbagai tes yaitu tes yang mengukur potensi kecerdasan Test Intelligensi Kolektif Indonesia–Pendidikan Tinggi (TIKI-T), serta beberapa instrumen yang mengukur penyesuaian diri, kemampuan bersosialisasi, kemampuan kerja mengukur motivasi, daya tahan terhadap stres, dan ketelitian, yaitu Instruksi WZT, Baum, Pauli dan EPPS. Hasil pengukuran ini dirangkum dalam psikogram yang penentuan kriteria-

nya dilakukan sendiri oleh psikolog yang menyelenggarakan tes-tes tersebut. Hasil penelitian menunjukkan bahwa berdasarkan hasil psikotes hanya 4 yang disarankan dan 21 yang dipertimbangkan untuk mengikuti pendidikan kedokteran, sedangkan 53 orang sisanya tidak disarankan. Sementara itu, hasil analisis menunjukkan korelasi yang lemah antara hasil psikotes dengan prestasi akademik. Hal ini kemungkinan dapat dipengaruhi oleh beberapa hal yaitu: besar sampel yang kurang serta sebagian besar yang diterima masuk FK adalah yang tidak disarankan, tes – tes yang dilakukan hanya mengukur potensi dimana manifestasinya dipengaruhi berbagai hal lain baik dari dalam maupun dari luar diri mahasiswa itu sendiri. Hasil wawancara terhadap psikolog yang melakukan tes ini didapatkan informasi bahwa sebelum tes dilakukan, pelaksana tidak diberikan pemahaman mengenai kriteria calon mahasiswa yang diharapkan, sehingga pendekatan yang digunakan pada waktu menentukan batasan kriteria dan melakukan interpretasi mungkin kurang sesuai dengan kriteria yang seharusnya dimiliki oleh mahasiswa kedokteran. Namun, meskipun hasil psikotes secara individu memiliki korelasi yang lemah dengan prestasi akademik tetapi ketika dilakukan regresi ternyata dapat memperkuat prediksi dari faktor kognitif terhadap prestasi akademik, sehingga pengukuran terhadap karakteristik psikologi tetap memberikan informasi yang berharga.

Penelitian ini juga menunjukkan tidak adanya korelasi yang bermakna antara karakteristik psikologi dengan nilai OSCE komprehensif. Dalam penelitian yang dilakukan oleh Adam *et al.* (2012) menunjukkan bahwa tes non-kognitif dapat memprediksi perilaku yang penting dimiliki oleh seorang dokter, seperti bekerja sama dalam kelompok, menyadari kekurangan dan menerima umpan balik, dan mengidentifikasi perilaku lainnya yang dapat mengurangi kemampuan dokter untuk melakukan tugasnya. Namun, ujian yang diselenggarakan selama proses pendidikan jarang mengukur kemampuan yang berhubungan dengan kualitas yang di-

ukur dalam tes non-kognitif. Kemampuan non-kognitif seperti komunikasi dan hubungan dokter-pasien, sering kali diukur secara informal, dan hasilnya dicatat dalam bentuk lulus/ tidak lulus. Dalam tes OSCE komprehensif yang dilakukan tampaknya masih perlu perbaikan struktur dan rubrik yang digunakan sehingga ada kecocokan antara faktor non-kognitif yang diukur dalam OSCE dan karakteristik psikologi yang diukur dalam psikotes.

Hasil pengukuran terhadap motivasi belajar mahasiswa menunjukkan sekitar 19% IPK S.ked dan 17% nilai rerata blok dipengaruhi oleh motivasi belajar mahasiswa. Selain itu, strategi belajar pun memiliki pengaruh yang cukup signifikan yaitu sebesar 30% terhadap IPK S.ked dan 25% terhadap nilai rerata blok. Umumnya, prestasi akademik yang tinggi berhubungan secara positif dengan *strategic* dan *deep approach*, serta berhubungan secara negatif dengan *surface approach*. Sebuah penelitian yang menganalisis hubungan antara gaya belajar (asimilator-eksplorer), strategi belajar (*deep, strategic, dan surface*), motivasi untuk berprestasi, dan *need for cognition* menunjukkan hasil yang positif dan signifikan (Diseth & Martinsen, 2003). Hubungan antara motivasi dan prestasi akademik dimediasi oleh pendekatan *deep approach* (Bruinsma, 2003).

Gaya belajar meliputi motivasi belajar dan strategi yang dilakukan oleh mahasiswa untuk melakukan tugasnya. Hasil penelitian menunjukkan gaya belajar tidak memiliki korelasi secara bermakna dengan prestasi akademik. Hal ini dimungkinkan karena gaya belajar dapat berubah, dipengaruhi kestabilan kepribadian, kesesuaian dengan konteks dan tugas. Oleh karena itu, akan sangat berguna bagi institusi pendidikan kedokteran mengajarkan cara untuk mengaplikasikan gaya belajar agar sukses dalam belajar (Ferguson *et al.*, 2002).

Pada penelitian ini gender tidak menunjukkan perbedaan yang bermakna terhadap prestasi akademik. Beberapa penelitian menunjukkan perempuan cenderung memiliki prestasi lebih baik dibandingkan laki-laki dalam proses pendidikan kedokteran. Pe-

rempuan dinilai cenderung lebih baik dalam penilaian klinik. Namun, perbedaan ini sangat kecil dan mencapai nilai signifikan bila sampel yang digunakan dalam penelitian cukup besar (Ferguson *et al.*, 2002).

Suatu penelitian tentang prediksi karakteristik non-kognitif terhadap prestasi akademik mahasiswa kedokteran di Sri Lanka, didapatkan bahwa yang mencapai prestasi tinggi secara signifikan adalah yang berusia muda (Ranasinghe *et al.*, 2012). Hasil analisis dalam penelitian ini menunjukkan hal yang serupa dimana usia mempengaruhi 4,8% dari IPK S.ked dan yang memiliki prestasi tinggi secara signifikan adalah yang berusia muda juga. Hubungan antara *gender*, usia dan prestasi akademik tidak tampak konsisten dalam berbagai penelitian. Hubungan antara *gender*, usia dan prestasi akademik dipengaruhi oleh berbagai variabel lain misalnya disiplin ilmu, manajemen waktu belajar, motivasi dan *deep approach* (Bruinsma, 2003).

Keterbatasan pada penelitian ini adalah tidak dilakukan analisis dari masing-masing komponen karakteristik psikologi yang diukur, bagaimana pengaruhnya terhadap prestasi akademik, karena pada pengukuran motivasi untuk berprestasi didapatkan hubungan yang bermakna. Selain itu, belum dilakukan analisis terhadap rubrik OSCE komprehensif yang digunakan sehingga dapat mempengaruhi hubungan yang tidak bermakna antara tes seleksi akademik, nilai rapor mahasiswa dan karakteristik psikologi dengan nilai OSCE.

## Simpulan

Tes seleksi akademik memiliki *predictive validity* terhadap prestasi akademik berupa IPK mahasiswa dan rerata nilai blok, tetapi tidak dapat memprediksi nilai OSCE komprehensif. Nilai rapor SMA tidak memiliki *predictive validity* terhadap prestasi akademik. Karakteristik psikologi yang diukur pada penerimaan mahasiswa baru tidak berhubungan secara bermakna dengan prestasi akademik, tetapi dapat meningkatkan kekuatan prediksi terhadap prestasi akademik apabila digunakan bersama dengan tes

seleksi akademik. Selain itu, hasil analisis menunjukkan faktor non-kognitif berupa motivasi dan strategi belajar memiliki pengaruh yang signifikan terhadap prestasi akademik.

Oleh karena itu, pada penerimaan mahasiswa baru perlu memperhatikan aspek kognitif dan non-kognitif, di antaranya adalah faktor motivasi, disertai dengan pengembangan tes pada seleksi mahasiswa, terutama pada pengukuran karakteristik psikologi yang dapat menunjang keberhasilan proses pembelajaran. Perlu dilakukan pengukuran strategi dan motivasi belajar mahasiswa pada akhir tahun pertama sehingga dapat menjadi umpan balik bagi mahasiswa dan institusi. Hubungan antara faktor kognitif dan non-kognitif dengan nilai OSCE komprehensif harus dibuktikan kembali karena belum dilakukan pengujian reliabilitas pada rubrik OSCE komprehensif.

## Persetujuan Etika

Penelitian ini telah mendapatkan persetujuan etik dari Komite Etik Kedokteran dan Kesehatan Fakultas Kedokteran Universitas Gadjah Mada, serta ijin melakukan penelitian yang diberikan oleh pimpinan Fakultas Kedokteran Universitas Swadaya Gunung Jati.

## Daftar Pustaka

- Adam, J., Bore, M., Mckendree, J., Munro, D., Powis, D., 2012. Can personal qualities of medical students predict in – course examination success and professional behaviour? An exploratory prospective cohort study. BMC medical education. Available at: <http://www.biomedcentral.com/1472-6920/12/69>.
- Adebayo, B., 2008. Cognitive and non-cognitive factors affecting the academic performance and retention of conditionally admitted freshmen. Journal of college admission.
- Arnold, WH., Gonzales, P., Gaengler, P., 2011. The predictive value of criteria

- for student admission to dentistry. *European journal of dental education*.
- Bruinsma, M. 2003. Effectiveness of higher education: factors that determine outcomes of university education. <http://dissertations.ub.rug.nl/faculties/ppsw/2003/m.bruinsma>
- Cassidy, S. & Eachus, P., 2000. Learning style, academic belief systems, self – report student proficiency and academic achievement in higher education. *Educational psychology*. <http://search.ebscohost.com>
- Disseth, A. & Martinsen, O., 2003. Approaches to learning, cognitive style, and motives as predictors of academic achievement. *Educational Psychology*. Vol. 23(2) pp.195-207.
- Faroki-Khajeh-Pasha, Y., Nedjat, S., Mohammadi, A., Rad, EM., Monajemi, F., Jamali, E., Yazdani, S., 2012. The validity of Iran's national university entrance examination (Konkour) for predicting medical student's academic performance. *BMC medical education*. Available at: <http://www.biomedcentral.com/1472-6920/12/60>.
- Ferguson, E., James, D., Madeley L., 2002. Factors associated with success in medical school: systematic review of the literature. *BMJ* vol 324. Available at: <http://www.bmj.com>
- Frischenschlager, O., Haidinger, G., Mitterauer, L., 2005. Factors associated with academic success at Vienna Medical School: prospective survey. *Croatian medical journal*, 46(1), pp. 58-65.
- Honey, P., Mumford, A. 2000. *The learning style helper's guide*. Maidenhead: Peter Honey Publication Ltd.
- Kumara, A. Hubungan antara motivasi berprestasi dan kecemasan dengan prestasi belajar para mahasiswa fakultas psikologi UGM Tahun ajaran 1989/1990. Laporan hasil penelitian.
- Kuncel, NR., Hezlett, SA., 2011. Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Association for psychological science*.
- Parry, J., Mathers, J., Stevens, A., Parsons, A., Lilford, R., Spurgeon, P., Thomas, H., 2006. Admissions processes for five year medical courses at English school: review. *BMJ*.
- Ranasinghe, P., Ellawela, A., Gunatilake, SB., 2012. Non– cognitive characteristics predicting academic success among medical students in Sri Lanka. *BMC medical education*. Available at: <http://www.biomedcentral.com/1472-6920/12/66>.
- Wilkinson, D., Zhang, J., Byrne, GJ., Luke, H., Ozolin, IZ., Parker, MH., Peterson, RF., 2008. Medical school selection criteria and the prediction of academic performance. *Medical journal of australia*, 188(6), pp. 349 – 354.

**EKSPLORASI PERMASALAHAN  
DALAM IMPLEMENTASI *COMMUNITY-BASED EDUCATION*  
DI FAKULTAS KEDOKTERAN**

<sup>1)</sup>Vivi Meidianawaty, <sup>2)</sup>Widyandana, <sup>3)</sup>Tri Nur Kristina

<sup>1)</sup>Fakultas Kedokteran Universitas Swadaya Gunung Jati, Cirebon

<sup>2)</sup>Fakultas Kedokteran Universitas Gadjah Mada, Yogyakarta

<sup>3)</sup>Fakultas Kedokteran Universitas Diponegoro, Semarang

<sup>1)</sup>vivi.meidianawaty@gmail.com

**Abstrak**

Penelitian ini bertujuan untuk mengeksplorasi permasalahan-permasalahan yang dapat ditemukan saat implementasi *Community-based Education* (CBE) di fakultas kedokteran. Pendekatan kualitatif eksplorasi dilakukan untuk mengidentifikasi permasalahan yang sering ditemui dalam implementasi program CBE. Pengumpulan data dilakukan di fakultas kedokteran negeri dan swasta dengan evaluasi dokumen, wawancara semistruktur, dan observasi. Kesimpulan dari penelitian ini adalah program CBE yang tidak menetapkan tujuan pembelajaran bersifat generik, kurang melibatkan peran aktif masyarakat dalam kegiatan pembelajaran mahasiswa, kebutuhan masyarakat yang jarang teridentifikasi, dan tidak adanya program yang berkelanjutan dapat menyebabkan kejenuhan masyarakat atau kegagalan program CBE mencapai tujuannya.

**Kata kunci:** *community-based education, implementasi, masalah, metode berkelanjutan*

**EXPLORATION OF PROBLEMS IN COMMUNITY-BASED EDUCATION IN  
THE FACULTY OF MEDICINE**

<sup>1)</sup>Vivi Meidianawaty, <sup>2)</sup>Widyandana, <sup>3)</sup>Tri Nur Kristina

<sup>1)</sup>Fakultas Kedokteran Universitas Swadaya Gunung Jati, Cirebon

<sup>2)</sup>Fakultas Kedokteran Universitas Gadjah Mada, Yogyakarta

<sup>3)</sup>Fakultas Kedokteran Universitas Diponegoro, Semarang

<sup>1)</sup>vivi.meidianawaty@gmail.com

**Abstract**

There were several problems in the implementation of Community-based Education CBE that need to be further explored. Exploratory qualitative approach was carried out to identify problems that were often encountered. Data collection was conducted in Public Medical School and Private Medical School by using document evaluation, semi-structured interviews, and observation. The conclusion of this study is CBE program that does not set the generic learning objectives, less involvement of the community active participation, the needs of community who are rarely identified, and the absence of a sustainable program can lead to saturation of the CBE program or failure to achieve its objectives.

**Keywords:** *community-based education, implementation, problems, sustainable methods*

## **Pendahuluan**

*Community-Based Education* (CBE) adalah suatu bentuk kegiatan belajar dimana mahasiswa mempelajari kompetensi seorang dokter di masyarakat sekaligus membantu mahasiswa membangun hubungan dengan masyarakatnya (WHO Regional Office for the Eastern Mediteranean, 2014). Konsep CBE ini perlu diimplementasikan sebagai upaya meningkatkan kualitas layanan kesehatan dokter di masyarakat.

Implementasi CBE pada beberapa negara bervariasi tergantung dari kebutuhan kesehatan masyarakatnya, sistem pelayanan kesehatan yang digunakan, dan tujuan belajar program CBE (Magzoub & Schmidt, 2000). Implementasi CBE di Eropa dan Amerika umumnya dilakukan dengan mengirimkan mahasiswa untuk magang di tempat praktik dokter keluarga atau spesialis, dan atau melakukan pelayanan sosial (Cauley et al., 2001). Chastonay et al. (2012) dalam penelitiannya memberikan contoh sebuah program pengenalan pada pelayanan pasien rawat jalan. Tujuannya adalah untuk memperkenalkan mahasiswa pada berbagai aspek pelayanan kesehatan primer. Lebih khusus mahasiswa berkenalan dengan tugas-tugas profesional dari dokter umum dan belajar untuk mengintegrasikan pengelolaan pasien yang berkelanjutan dalam praktik sehari-hari. Implementasi CBE di negara-negara berkembang seperti di Afrika dan Asia, dilaksanakan dengan mengirimkan mahasiswa ke daerah-daerah pedesaan, sarana pelayanan kesehatan masyarakat primer atau ke komunitas tertentu. Kristina, et al., (2006) melakukan penelitian terhadap kegiatan yang dilakukan mahasiswa meliputi mengidentifikasi masalah, mencari solusi intervensi bersama masyarakat, melaksanakan sampai mengevaluasi hasil intervensi.

Meskipun dalam penerapannya terdapat berbagai variasi, namun masalah yang ditemukan dalam implementasi CBE hampir sama. Masalah utama program CBE meliputi perancangan program, perencanaan program CBE, lokasi/komunitas, dan implementasi program. Permasalahan yang berhubungan dengan kurikulum dan sum-

ber daya sampai pengorganisasian permasalahan dapat menjadi kelemahan dari implementasi program CBE (Hamad, 1991).

Mengingat belum banyak penelitian mengenai implementasi CBE di Indonesia, maka belum banyak diketahui tentang permasalahan yang ditemukan saat CBE diimplementasikan di fakultas kedokteran. Oleh karena itu diperlukan penelitian yang dapat memberikan informasi mengenai bagaimana implementasi CBE dan permasalahan yang ditemukan pada saat implemenasinya.

## **Metode Penelitian**

Penelitian ini menggunakan metodologi kualitatif eksplorasi untuk mengidentifikasi permasalahan yang dapat muncul saat implementasi program CBE. Penelitian dilakukan di satu fakultas kedokteran negeri (Institusi A) dan satu fakultas kedokteran swasta (Institusi B). Subjek pada penelitian ini adalah staf pengajar pada masing-masing institusi yang merupakan penanggung jawab dan pelaksanaan program CBE. Proses pemilihan informan dipilih yaitu penanggung jawab program dan staf yang terlibat dalam program CBE (pembimbing CBE di lapangan) minimal 3 tahun.

Pengumpulan data dilakukan dengan cara mengevaluasi dokumen, wawancara semi-terstruktur dan observasi. Pada evaluasi dokumen digunakan panduan yang peneliti kembangkan sendiri berdasarkan Buku Desain & Implementasi Pembelajaran Berbasis Masyarakat di Fakultas Kedokteran yang ditulis oleh Kristina (2011). Dokumen yang dianalisis adalah buku panduan pelaksanaan CBE, buku pedoman KKN, dan laporan kegiatan. Aspek-aspek yang dinilai pada buku panduan pelaksanaan dan buku pedoman KKN meliputi tujuan pembelajaran, dasar teori, kegiatan belajar, penilaian untuk mahasiswa, dan tata tertib kegiatan. Pada laporan kegiatan dinilai kesesuaian kegiatan yang dilaksanakan dengan perencanaan yang ada di buku panduan pelaksanaan. Wawancara semi-terstruktur dilakukan dengan menggunakan daftar pertanyaan yang sudah disiapkan sebagai panduan yang

telah disusun dengan merujuk pada artikel *Twelve Tips for Community-based Medical Education* oleh Howe dan Buku Desain & Implementasi Pembelajaran Berbasis Masyarakat di Fakultas Kedokteran oleh Kristina. Wawancara dilakukan sendiri oleh peneliti. Wawancara dilakukan sebanyak 1 kali selama 30 menit -1 jam pada setiap informan. Pada saat wawancara, dilakukan pencatatan oleh peneliti dan perekaman dengan menggunakan audio recorder. Observasi dilakukan oleh peneliti dengan mengamati kegiatan program di lokasi atau tempat pembelajaran menggunakan panduan yang peneliti kembangkan berdasarkan Buku Desain & Implementasi Pembelajaran Berbasis Masyarakat di Fakultas Kedokteran yang ditulis oleh Kristina (2011). Pada saat observasi dilakukan pencatatan dan direkam dengan menggunakan kamera. Aspek-aspek yang diamati meliputi lokasi atau tempat belajar, kegiatan belajar, dosen pembimbing, mahasiswa, dan masyarakat. Kegiatan di Institusi A yang diobservasi adalah kegiatan survei kesehatan dan kegiatan edukasi pada siswa sekolah dasar. Observasi dilakukan sebanyak 1 kali. Observasi kegiatan di Institusi B dilakukan pada blok 3.1, 5.3, 6.1, 6.3. Observasi dilaksanakan selama program CBE berjalan yaitu 1 minggu.

Peneliti melakukan analisis data dengan mereduksi data yang diperoleh menjadi informasi yang memiliki makna namun lebih ringkas. Pada tahap pertama dilakukan dengan membaca keseluruhan data untuk membangun pemahaman umum dari informasi yang diperoleh. Selanjutnya, pada langkah dilakukan dengan menganalisis lebih detail dengan meng-*coding* data. Langkah ini melibatkan beberapa tahap yaitu mengambil data tulisan atau gambar yang telah dikumpulkan, mensegmentasi kalimat kemudian dikelompokkan ke dalam kategori. Setelah itu mencari hubungan antarkategori-kategori. Tahap yang terakhir pada analisis data adalah menetapkan tema untuk menjadi sebuah deksripsi umum.

Upaya keabsahaan data pada penelitian ini dilakukan dengan merujuk pada strategi-strategi keabsahan data kualitatif

yang disarankan oleh Creswell (2003) yaitu sebagai berikut.

Pertama adalah *Intercoder agreement*. Persetujuan ini didasarkan pada 2 atau lebih pemeriksa kata kunci yang telah sepakat mengenai kata kunci-kata kunci yang ada pada data. Pada penelitian ini dilakukan pencarian kata kunci oleh peneliti dan 1 orang dosen dengan latar belakang pendidikan kedokteran. Pada saat terjadi ketidaksepakatan maka akan dibandingkan dengan teori atau penelitian sebelumnya.

Kedua adalah *Triangulate*. Data dikumpulkan dari berapa sumber yaitu dokumen, wawancara semi-terstruktur, dan observasi. Ketiga adalah *Member checking*. Proses ini dilakukan dengan membawa kembali laporan deksripsi atau tema kepada informan untuk mengecek apakah tema atau deskripsi yang telah dibuat oleh peneliti sudah akurat. Bagi informan yang berasal dari Institusi A hasil analisis dikirimkan melalui email.

Penelitian ini telah mendapatkan persetujuan dari Komisi Etik Penelitian Fakultas Kedokteran Universitas Gadjah Mada.

## Hasil Penelitian dan Pembahasan

Gambaran secara umum dari implementasi CBE di ke 2 institusi adalah sebagai berikut: Sejak 2013, Institusi A mulai melaksanakan program CBE bersamaan dengan program Kuliah Kerja Nyata (KKN) yang diselenggarakan oleh pihak universitas. Program ini dilaksanakan di semester akhir atau semester 7. Mahasiswa ditempatkan di desa-desa yang telah ditentukan oleh universitas. Pada 1 desa ditempatkan 1-2 orang mahasiswa dari fakultas kedokteran. Setiap mahasiswa mendapat tugas untuk melakukan survei kesehatan, kemudian menentukan prioritas masalah dan menentukan intervensinya bersama dengan masyarakat. Setiap mahasiswa dibekali buku pedoman KKN. Untuk pendampingan selama mahasiswa menjalani program tersebut disediakan 2 orang dosen pembimbing lapangan untuk 1 kecamatan, dan diakhir program mahasiswa membuat laporan untuk setiap kegiatan yang telah dilaksanakan dilengkapi dengan *log book*.

Program CBE diimplementasikan sejak tahun 2010 di Institusi B. Untuk mempersiapkan program ini, fakultas terlebih dahulu mengadakan MOU dengan Dinas Kesehatan setempat untuk penyediaan lahan pendidikan dan instruktur untuk program CBE. Pemetaan program CBE dalam kurikulum tersebar pada setiap tingkatan disesuaikan dengan level kompetensi yang ingin dicapai. Kegiatan CBE sebagian besar dilakukan oleh bagian keterampilan klinik dengan tujuan kegiatan yang terintegrasi dengan blok yang berjalan. Tempat belajar dilakukan di Puskesmas atau langsung di masyarakat. Jangka waktu kegiatan CBE ini sekitar 1 minggu. Selain kegiatan CBE yang dilakukan oleh bagian keterampilan klinik ada juga yang dilakukan pada Blok *implementation of family medicine and community health*, dimana mahasiswa menghabiskan waktu selama 4 minggu di lapangan, baik di puskesmas atau di masyarakat. Tujuan dari blok ini mahasiswa mampu mengaplikasikan pengetahuan yang telah didapat pada tahap sebelumnya terutama yang terkait dengan kedokteran keluarga dan kesehatan masyarakat. Selama kegiatan CBE mahasiswa didampingi oleh dosen pembimbing yang berasal dari puskesmas.

Hasil dari penelitian menunjukkan bahwa permasalahan-permasalahan yang muncul saat implementasi program CBE dapat berasal dari perancangan program yaitu dengan penentuan tujuan pembelajaran yang belum bersifat generik, tujuan belajar kurang dipahami oleh pembimbing dan mahasiswa, lamanya program, tidak adanya program yang berkelanjutan, kurang dilibatkannya peran aktif dari masyarakat saat kegiatan belajar mahasiswa, komitmen dari pembimbing di lapangan dan munculnya kejenuhan masyarakat terhadap program CBE.

Tema-tema yang muncul pada penelitian ini adalah sebagai berikut.

#### Kurang Dipahaminya Tujuan Program CBE

Pada Institusi B, tujuan pembelajaran sudah tercantum pada buku pedoman pelaksanaannya. Meskipun demikian tujuan

program CBE ini dirasakan masih belum jelas dipahami oleh para pembimbing dan mahasiswa terutama di awal penerapan program CBE.

*“.....modul yang disiapkan kurang jelas, tujuan pembelajaran kurang jelas sehingga membingungkan untuk mahasiswa.”*  
(Responden 2)

*“Terutama pada saat di awal-awal ya, yang pada saat kita, saya di puskesmas sini juga masih belajar, mahasiswa juga yang memang baru turun ke puskesmas. Kadang mereka bingung mau cari apa di puskesmas.”*  
(Responden 3)

*“ ....biasanya di hari pertama mahasiswa kurang persiapan. Mahasiswa masih belum paham apa yang akan mereka pelajari di puskesmas..”* (Responden 4)

Hasil evaluasi dokumen di kedua institusi menunjukkan bahwa harapan agar mahasiswa melibatkan peran aktif dari masyarakat dalam kegiatannya masih belum optimal. Pada buku pedoman kegiatan di institusi A belum tercantum tujuan belajar untuk mahasiswa agar melibatkan masyarakat dalam kegiatannya. Untuk Institusi B tujuan belajar yang mengharapkan mahasiswa untuk melibatkan partisipasi masyarakat dalam kegiatannya baru ada di blok terakhir

#### Implementasi Program CBE dalam Kurikulum

Program CBE di Institusi A hanya dilaksanakan satu kali yaitu di semester akhir atau semester 7 bersamaan dengan program KKN.

*“ Untuk program CBE itu kami punya program sebelum dia dinyatakan lulus jadi sarjana kedokteran yaitu pada semester 7 “.*  
(Responden 1)

Program CBE dalam kurikulum Institusi B tersebar pada setiap tingkatan disesuaikan dengan level kompetensi yang ingin dicapai. Kegiatan CBE sebagian besar dilakukan oleh bagian keterampilan klinik dengan tujuan kegiatan yang terintegrasi dengan blok yang berjalan. Di semester 7 ma-

hasiswa pada Institusi B akan melaksanakan kegiatan CBE secara lebih terintegrasi. Meskipun demikian masih belum terlihat adanya program yang berkelanjutan. Kegiatan yang dilaksanakan banyak menguntungkan dari segi mahasiswa karena mahasiswa dapat belajar keterampilan klinis secara langsung pada pasien.

Kegiatan belajar yang tidak sesuai tujuan belajar juga disadari oleh pembimbing sebagai kelemahan implementasi program CBE pada institusi B. Contohnya untuk mahasiswa semester 6 diharapkan mahasiswa mampu mengidentifikasi permasalahan kesehatan masyarakat melalui survei kesehatan, tetapi pada praktiknya mahasiswa baru mengidentifikasi di tingkat perorangan, seperti komentar responden di bawah ini.

*“Praktik pembelajaran yang telah dilakukan selama ini masih berfokus pada usaha kesehatan perorangan dalam bentuk wawancara pasien yang datang ke puskesmas. Praktik tersebut tidak mengambil sebagian dari masyarakat untuk di wawancara. Sehingga permasalahan yang dikajipun adalah permasalahan perorangan dan bukan masyarakat.” (Responden 4)*

Pembimbing juga merasakan bahwa waktu yang dihabiskan mahasiswa semester 1 di puskesmas selama 1 hari untuk melihat tugas dan peran dokter puskesmas masih sangat kurang.

*“Community based-nya kurang optimal. Waktunya hanya 1 hari, aktifitasnya hanya observasi kegiatan di dalam gedung. Alokasi waktu tersebut kurang memadai untuk memotret gambaran umum kegiatan dalam gedung. Sedangkan kegiatan kesehatan masyarakat yang fokusnya luar gedung (kunjungan rumah, posyandu, posbindu, musyawarah warga, dsb.) tidak terekam”. (Responden 4)*

Dari hasil pengamatan mengenai aktivitas belajar mahasiswa pada kedua institusi didapatkan variasi kegiatan belajar di beberapa tempat seperti puskesmas, posyandu, posyandu lansia, sekolah dasar, dan keluarga. Kegiatan-kegiatan tersebut memberikan kesempatan kepada mahasiswa untuk belajar langsung di masyarakat. Meskipun de-

mikian, harapan agar mahasiswa melibatkan masyarakat secara aktif dalam kegiatannya belum terlihat. Masyarakat baru sebatas sebagai pasien dalam pembelajaran keterampilan klinik, sebagai responden dalam kegiatan survei kesehatan atau sebagai peserta dalam kegiatan penyuluhan kesehatan. Contohnya adalah saat mahasiswa melakukan kegiatan edukasi di sekolah dasar, mahasiswa memberitahukan pihak sekolah mengenai kegiatan tersebut. Pemilihan materi edukasi yang akan diberikan ditentukan sendiri oleh mahasiswa belum berdasarkan kebutuhan yang dirasakan oleh masyarakat.

#### Peran dan Tugas Pembimbing Program CBE

Pelaksanaan program CBE mengambil tempat di luar institusi seperti di puskesmas atau posyandu, ini berarti institusi membagi tugas dan tanggung jawab pendidikan bersama institusi lain. Pembimbing harus membagi peran dan tugasnya antara sebagai pelayan kesehatan dan pendidikan. Keterbatasan waktu untuk melaksanakan tugas pendidikan disadari oleh para pembimbing di puskesmas.

*“Karena sebagian besar yang dijadikan pembimbing adalah kepala puskesmas dan sering diundang rapat sehingga waktu untuk membimbing mahasiswa menjadi terbatas.” (Responden 2)*

Kegiatan belajar yang mengharuskan mahasiswa untuk ke luar gedung puskesmas juga menjadi kendala bagi para pembimbing untuk dapat mendampingi mahasiswa saat kegiatannya. Hal ini membuat pembimbing harus mendelegasikan tugasnya kepada tenaga medis lain yang belum dilatih sebagai pembimbing mahasiswa untuk program CBE.

*“Nanti kalau di lapangan, misalnya imunisasi saya akan kasih tau ke petugas imunisasi. Kalau ada kegiatan di bagian gizi nanti akan saya kasih tau ke petugas gizi. Kemudian pada saat nanti di dalam gedung ada beberapa pos yang harus dimasuki ya saya koordinasi dengan petugas atau programmer yang akan membimbing mahasiswa”. (Responden 3)*

“ Karena keterbatasan saya sebagai pembimbing harus menilai, tetap inti-inti dari penilaian itu saya delegasikan ke tiap-tiap pos yang meihat langsung kegiatan dari adik-adik mahasiswa. Jadi nanti mereka akan memberikan masukan ke saya”. (Responden 3)

#### Kejenuhan terhadap program CBE

Kejenuhan masyarakat terhadap program CBE yang beberapa kali diterapkan pada suatu komunitas akan menjadi hambatan terhadap kesuksesan program CBE.

“Pada awalnya masyarakat masyarakat akan merasa senang, welcome, namun setelah beberapa siklus begitu akan ada kejenuhan.” (Responden 1)

Institusi A mengimplementasi program CBE pada akhir semester 7 dengan mengikuti mahasiswa pada program KKN yang diselenggarakan oleh universitas. Program CBE yang hanya ada di semester akhir sebagaimana yang diimplementasikan di Institusi A kurang sesuai dengan tujuan program CBE pada umumnya yaitu untuk meningkatkan derajat kesehatan masyarakat. Dengan program CBE yang hanya ada di akhir semester tentunya tidak ada suatu program yang berkelanjutan yang dapat menghubungkan program CBE yang satu dengan yang lainnya sehingga manfaat dari program CBE dapat dirasakan oleh masyarakat.

Tujuan program CBE di Institusi A sebenarnya sudah mencerminkan tujuan pembelajaran yang bersifat generik, namun tujuan ini tidak tercantum di buku pedoman KKN. Tujuan pembelajaran yang tidak tercantum dalam buku pedoman seperti yang ada di Institusi A dapat membuat mahasiswa tidak bisa memahami apa yang harus dicapai selama mengikuti program tersebut. Tujuan pembelajaran ini harus diketahui oleh pembimbing dan mahasiswa termasuk kegiatan yang harus dilakukan oleh mahasiswa agar tujuan pembelajaran tersebut dapat tercapai.

Kegiatan belajar yang dilaksanakan pun masih belum terlihat melibatkan peran aktif dari masyarakat atau berdasarkan kebutuhan dari masyarakat tersebut. *Communi-*

*nity-based education* memiliki implikasi yang luas pada seluruh proses pendidikan dan keterlibatan masyarakat merupakan bagian tak terpisahkan dari proses perkembangan pendidikannya.

Hasil penelitian di Institusi A ditemukan bahwa masyarakat mengalami kejenuhan terhadap kegiatan mahasiswa saat melaksanakan program CBE. Hal ini juga dialami oleh Christian Medical College di Vellore India saat melaksanakan program CBE mereka. Kegagalan program CBE untuk melibatkan peran aktif masyarakat dalam kegiatan pembelajaran mahasiswa, kebutuhan masyarakat yang jarang teridentifikasi, dan tidak adanya program yang berkelanjutan dapat menyebabkan kejenuhan atau keengangan masyarakat terhadap program CBE.

Institusi B memetakan program CBE dalam kurikulumnya tersebar dari semester awal hingga semester akhir tahap sarjana kedokteran. Namun, hal tersebut baru menguntungkan dari aspek mahasiswa dan belum terlihat adanya program yang berkelanjutan berdasarkan permasalahan kesehatan masyarakat yang menguntungkan dari aspek masyarakatnya. Dalam pengembangan kurikulum program CBE disarankan untuk dapat diterapkan dengan menggunakan metode berkelanjutan dari semester awal sampai akhir sehingga aspek tingkat kesulitan dalam belajar dapat disesuaikan dengan level mahasiswa. Disamping itu, metode berkelanjutan dapat menghubungkan antara program CBE yang satu dengan yang lainnya dan manfaat dari program CBE dapat dirasakan oleh masyarakat. Durasi mahasiswa mengikuti program CBE dirasakan oleh pembimbing masih belum cukup untuk mendapatkan gambaran yang utuh mengenai puskesmas. Memang belum ada kesepakatan mengenai berapa lama sebaiknya program CBE dilaksanakan, namun menurut Talaat & Ladhani bahwa rata-rata beberapa fakultas menerapkan program CBE selama 1 bulan dalam setahun di tempat yang sama.

Program CBE adalah proses pembelajaran mahasiswa yang mengambil tempat di luar institusi. Kurangnya pengawasan da-

pat terjadi selama masa pelaksanaannya dan akan membuat program CBE tidak berjalan sesuai dengan yang sudah direncanakan. Tujuan pembelajaran ini harus diketahui oleh pembimbing dan mahasiswa termasuk kegiatan yang harus dilakukan oleh mahasiswa agar tujuan pembelajaran tersebut dapat tercapai (6). Pada penelitian ini didapatkan bahwa meskipun tujuan pembelajaran sudah tercantum dalam buku pedoman di Institusi B, namun tujuan pembelajaran tersebut dirasakan oleh pembimbing masih kurang jelas terutama pada masa-masa awal pelaksanaan program CBE walaupun persiapan untuk pembimbing sudah dilaksanakan sebelum program dimulai. Permasalahan seperti tersebut juga pernah ditemukan oleh Kaye et al., (2011) saat melakukan evaluasi program CBE di Uganda bahwa in-struktur masih belum memahami kurikulum, terutama mengenai kompetensi yang diharapkan termasuk bagaimana kurikulum diimplementasikan.

Dalam menyusun kegiatan belajar untuk program CBE Institusi B sudah melaksanakan dengan baik yaitu terjun langsung ke masyarakat baik di puskesmas, lingkungan RW maupun di tingkat keluarga. Apapun tujuan belajar yang telah ditetapkan oleh institusi, kegiatan belajar yang diselenggarakan harus dapat memberikan pengalaman belajar kepada mahasiswa. Namun demikian, saat melaksanakan kegiatan peran aktif dari masyarakat kurang dilibatkan. Disebutkan oleh Kristina bahwa kegiatan-kegiatan yang dilaksanakan harus berdasarkan kebutuhan masyarakat dan harus ada peran aktif dari masyarakat untuk ikut terlibat karena salah satu faktor penting yang menentukan keberlanjutan dari suatu program CBE adalah adanya peran aktif dari masyarakat, tanpa adanya keterlibatan masyarakat maka program CBE tidak akan memberikan manfaat.

Peneliti menyadari bahwa dalam penelitian ini masih banyak beberapa kekurangan diantaranya adalah: (1) penelitian baru dilakukan pada 2 institusi, sehingga tidak cukup menggambarkan implementasi

program CBE pada fakultas kedokteran di Indonesia; (2) penelitian belum menggali data dari mahasiswa dan masyarakat sebagai bagian dari komponen program CBE; (3) penelitian belum mencakup outcome program CBE sehingga belum dapat mengevaluasi program CBE secara keseluruhan. Untuk itu diperlukan penelitian lanjutan agar didapatkan gambaran secara utuh pelaksanaan program CBE pada fakultas kedokteran di Indonesia. Selain itu persepsi peneliti terhadap implementasi program CBE terbentuk dari pengalaman pribadi peneliti tentu membawa bias-bias tersendiri ke dalam penelitian ini. Akan tetapi, bias-bias ini membantu bagaimana peneliti memandang dan memahami data yang dikumpulkan.

## **Simpulan**

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, didapatkan beberapa permasalahan saat implementasi program CBE di kedua institusi yaitu dengan penentuan tujuan pembelajaran yang belum bersifat generik, tujuan belajar kurang dipahami oleh pembimbing dan mahasiswa, lamanya program, tidak adanya program yang berkelanjutan, kurang dilibatkannya peran aktif dari masyarakat saat kegiatan belajar mahasiswa, komitmen dari pembimbing di lapangan dan munculnya kejenuhan masyarakat terhadap program CBE.

Beberapa hal yang dapat peneliti sarankan berdasarkan temuan pada penelitian ini yaitu menetapkan tujuan program CBE yang bersifat generik yang responsif terhadap kebutuhan masyarakat, menerapkan metode berkelanjutan pada program CBE, mensosialisasikan tujuan pembelajaran dan kegiatan mahasiswa, melakukan pelatihan untuk semua yang dapat dijadikan pembimbing, mempertimbangkan beberapa lokasi yang berbeda untuk mencegah kejenuhan dari masyarakat, dan melibatkan peran aktif masyarakat dalam kegiatan belajar.

## **Daftar Pustaka**

Cauley, K., Canfield, A., Clasen, C., Dobbins, J., Hemphill, S., Jaballas, E.,

- & Walbroehl, G. (2001). Service Learning: Integrating Student Learning and Community Service. *Education for Health (Abingdon, England)*, 14, 173–181.
- Habbick, B. F., & Leeder, S. R. (1996). Orienting medical education to community need: a review. *Medical Education*, 30(3), 163–71. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8949549>
- Hamad, B. (1991). Community-oriented medical education: what is it? *Medical Education*, 25(1), 16–22. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1997823>
- Howe, A. (2002). Twelve tips for community-based medical education. *Medical Teacher*, 24(1), 100–2. <http://doi.org/10.1080/00034980120103423>
- Kaye, D. K., Muhwezi, W. W., Kasozi, A. N., Kijjambu, S., Mbalinda, S. N., Okullo, I., ... Mwanika, A. (2011). Lessons learnt from comprehensive evaluation of community-based education in Uganda: a proposal for an ideal model community-based education for health professional training institutions. *BMC Medical Education*, 11(1), 7. <http://doi.org/10.1186/1472-6920-11-7>
- Kristina, T. N. (2011). *Desain dan Implementasi Pembelajaran Berbasis Masyarakat di Fakultas Kedokteran*. (T. N. Kristina, Ed.). Yogyakarta: Bagian Pendidikan Kedokteran FK UGM.
- Kristina, T. N., Majoor, G., & Van Der Vleuten, C. (2006). Does Community-Based Education come close to what it should be? A case study from the developing world: students' opinions. *Education for Health (Abingdon, England)*, 19(2), 179–88. <http://doi.org/10.1080/13576280600783596>
- Magzoub, M. E., & Schmidt, H. G. (2000). A taxonomy of community-based medical education. *Academic Medicine: Journal of the Association of American Medical Colleges*, 75(7), 699–707. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10926020>
- Omotara, B. a, Yahya, S. J., Shehu, U., Bello, H. S., & Bassi, a P. (2006). Communities' awareness, perception and participation in the Community-Based Medical Education of the University of Maiduguri. *Education for Health (Abingdon, England)*, 19(2), 147–54. <http://doi.org/10.1080/13576280600783661>
- Talaat, W., & Ladhani, Z. (2014). *Community Based Education in Health Professions: Global Perspectives*. Regional Office for Eastern Mediterranean of World Health Organization.
- WHO Regional Office for the Eastern Mediteranean. (2014). *Community-based Education in Health Professions: Global Perspectives*. (W. Talaal & Z. Ladhani, Eds.). World Helath Organization.
- WHO Study Group Report. (1987). *Community based education of health personel*. Geneva.

## PENGEMBANGAN INSTRUMEN BAKAT KEGURUAN

<sup>1)Wasidi, 2)Djemari Mardapi</sup>

<sup>1)Universitas Bengkulu, 2)Universitas Negeri Yogyakarta</sup>

<sup>1wasidirma@gmail.com, 2djemarimardapi@gmail.com</sup>

### Abstrak

Tujuan penelitian ini adalah untuk mengembangkan instrumen bakat keguruan yang valid dan reliabel. Penelitian ini dilakukan dengan tiga tahap yaitu: pra-pengembangan, pengembangan konseptual, dan uji coba instrumen. Data dianalisis dengan *item response theory partial credit model*, analisis faktor konfirmatori, validitas konkuren, validitas konvergen, dan koefisien reliabilitas. Hasil penelitian menunjukkan bahwa instrumen bakat keguruan terdiri atas tiga dimensi yaitu kreativitas pedagogi, komitmen pedagogi, dan kecerdasan emosi. Ketiga dimensi instrumen bakat keguruan memenuhi syarat IRT PCM. Hasil *confirmatory factor analysis* menunjukkan bahwa instrumen bakat keguruan fit. Koefisien reliabilitas gabungan tinggi. Analisis *multitrait multimethod* menunjukkan bahwa korelasi antara skor kreativitas pedagogi dengan skor IQ rendah. Korelasi antara skor komitmen pedagogi dengan skor *Edwards Personal Preference Schedule* adalah cukup. Korelasi antara skor kecerdasan emosi dengan skor EPPS adalah cukup. Validitas konvergen dimensi komitmen pedagogi termasuk tinggi, dan validitas konvergen kecerdasan emosi termasuk tinggi. Dengan demikian instrumen bakat keguruan mempunyai validitas isi, validitas konstruk, dan validitas konvergen yang baik, sedangkan dimensi komitmen pedagogi dan kecerdasan emosi mempunyai validitas konkuren yang termasuk kategori cukup. Koefisien reliabilitas gabungan instrumen bakat keguruan memenuhi persyaratan minimal. Dengan demikian instrumen bakat keguruan dapat digunakan oleh LPTK sebagai tes bakat calon mahasiswa.

**Kata kunci:** *bakat keguruan, validitas isi, validitas konstruk, validitas konkuren, validitas konvergen, koefisien reliabilitas gabungan*

## DEVELOPING A TEACHER APTITUDE INSTRUMENT

<sup>1)Wasidi, 2)Djemari Mardapi</sup>

<sup>1)Universitas Bengkulu, 2)Universitas Negeri Yogyakarta</sup>

<sup>1wasidirma@gmail.com, 2djemarimardapi@gmail.com</sup>

### Abstract

The purpose of this study is to develop an instrument of teacher aptitude which is valid and reliable. This research was carried out in three phases: pre-development, conceptual development, and instrument try out. The data were analyzed using the item response theory partial credit model, confirmatory factor analysis, concurrent validity, convergent validity, and reliability coefficient. The result of this research is a teacher aptitude instrument that consists of three dimensions: pedagogical creativity, pedagogical commitment, and emotional intelligence. The three dimensions of the instrument have IRT PCM qualification. The result of CFA shows that the teacher aptitude instrument is fit. The coefficient of the reliability is high. The correlation between pedagogical creativity score and intelligence quotient score is low. The correlation between pedagogical commitment and Edwards Personal Preference Schedule score is sufficient. The correlation between emotional intelligence score and EPPS score is sufficient. The convergent validity of pedagogical commitment is high, while the convergent validity of emotional intelligence is high. Therefore the teacher aptitude instrument has a good content validity, construct validity, and convergent validity, while dimension of pedagogical commitment and emotional intelligence has sufficient concurrent validity. The coefficient of composite reliability of the instrument meets the minimum requirements. Therefore the teacher aptitude instrument could be used by LPTK as an aptitude test for student candidates.

**Keywords:** *teacher aptitude, content validity, construct validity, convergent validity, coefficient of composite reliability*



## Pendahuluan

Undang-Undang Nomor 20 Tahun 2003, Undang-Undang tentang Guru dan dosen No. 14 Tahun 2005 pasal 7, dan Peraturan Pemerintah No.19 Tahun 2005 pasal 28 mensyaratkan guru ideal akademik dan non-akademik. Keberadaan guru dalam proses pendidikan sangat penting yang tidak dapat digantikan oleh apapun (Block, 2008; Sudarnoto, 2009). Banyak penelitian yang mendukung pentingnya guru dalam proses pendidikan dan pembelajaran, namun ada kesenjangan antara harapan dan kenyataan (Alkharusi, Kazem, & Musawai 2011; Dicky, 2011; Mardiana, 2006; Rowe, 2003; Sugiharto, 2003). Pengaruh karakter (aspek non-kognitif) guru sangat besar terhadap perkembangan peserta didik serta keberhasilan pendidikan (Açikgöz, 2005; Dorman, 2008; Garcia, Kupczynski, & Holland, 2011; Ishak, Piet, & Ramli, 2010; Jacobs & Mitchell 2008; Justice & Espinoza, 2007; Kokkinos, Charalambous, & Davazoglou 2010). Namun, dalam seleksi calon guru belum digunakan instrumen non-kognitif sebagai bahan pertimbangan kelulusan, karena belum tersedianya instrumen karakter/non-kognitif. Untuk mengisi kekosongan tersebut maka dikembangkan instrumen bakat keguruan yang mencakup bakat non-kognitif yang memenuhi syarat validitas dan reliabilitas. Masalah pokok dalam penelitian ini adalah bagaimana konstruk bakat keguruan, dan berapa koefisien reliabilitasnya.

Bakat pada awalnya mengacu pada pengertian kecerdasan (*intelligence*). Bakat adalah kondisi. *Aptitude* adalah kondisi atau cirikhas dianggap sebagai gejala kemampuan individu untuk memperoleh pelatihan, biasanya berupa pengetahuan, keterampilan atau tanggapan, seperti kemampuan untuk berbicara bahasa, untuk menghasilkan musik (Bingham, 1948), sedangkan Martinson (1974) menyebutkan bahwa anak-anak berbakat dan berbakat diidentifikasi sebagai individu yang berkualitas, mampu berkinerja tinggi. Salkind (2008) menyebutkan bahwa bakat adalah perbedaan individu yang berkaitan dengan pembelajaran berikutnya dalam jangka waktu yang sama. Perbedaan in-

dividu dapat terjadi pada ranah *cognitive* dan *non-cognitive*. *Aptitude* mengacu pada kualitas karakteristik perilaku orang yang menunjukkan seberapa baik ia dapat belajar untuk memenuhi dan memecahkan masalah tertentu. Ada prasyarat individu untuk dapat menguasai keahlian tertentu, prasyarat itulah yang dimaksud dengan bakat.

Teori bakat telah berubah, Renzulli (1986, pp. 53-93) yang didukung oleh Sternberg & Davidson (2005, pp.246-273) mengemukakan bahwa keberbakatan adalah multidimensi dan dapat diletakkan di setiap area kemampuan manusia. Bakat merupakan interaksi antara tiga kelompok dasar sifat manusia yaitu kemampuan umum di atas rata-rata, tingkat komitmen tugas dan tingkat. Ketiga dimensi tersebut dikenal dengan teori *the three ring concepts*. Renzulli dalam Baum, Reis, & Maxfield (1998) mengatakan bahwa kemampuan umum adalah kemampuan verbal, numerik, spasial, ingatan, yang biasanya diukur dengan kecerdasan umum. Namun Gardner (2006, pp. 8-24) mengatakan bahwa kecerdasan ada 8 jenis, dua di antaranya adalah kecerdasan antarpersonal dan interpersonal. Kecerdasan emosional meliputi interpersonal dan intrapersonal yang merujuk berbagai keterampilan yang terkait dengan emosi diri sendiri dan emosi orang lain, mempelajari perilaku dan tanggapan terhadap emosi orang lain, mampu untuk berbagi pengalaman positif yang bermakna kepada orang lain, dan kemampuan untuk berhubungan dengan mereka (Sigmar, Hynes, & Cooper, 2010). Kecerdasan emosi diperlukan oleh guru dalam melaksanakan tugas mendidik, guru yang baik adalah seorang yang dapat sebagai pamong yang mempunyai kecerdasan emosi yang bagus, yang dapat mengendalikan dan mengelola emosinya dengan baik. (Anonim, 1977, p.16; Sudiyat, 1987, p.3). Kemampuan di atas rata-rata (*above average ability*) dalam konsep bakat keguruan dapat digantikan dengan kecerdasan emosi, dengan demikian bakat keguruan terdiri atas kreativitas pedagogi, komitmen pedagogi, dan kecerdasan emosi. Bakat keguruan adalah potensi kemampuan individu dapat berkembang dengan pendidikan untuk

melaksanakan proses pendidikan dan pembelajaran dengan baik.

Kreativitas dalam pendidikan terdiri atas aspek: orisinalitas, imajinasi, pemecahan masalah, adaptasi, jaringan dan komunikasi (Dawson Tan, & McWilliam 2011). Mengajar kreatif adalah pembelajaran yang dilakukan dengan tertib, dan dengan improvisasi dan dalam cakupan kinerja yang luas (Sawyer, 2004). Merencanakan pembelajaran merupakan sebuah pekerjaan kreatif, ibaratnya dipandang sebagai menulis sebuah skenario pembelajaran (Beetlestones, 2011, p.172). Guru yang kreatif dapat dilihat dengan ciri-ciri rasa ingin tahu, berpikir orisinal, mandiri, berani mengambil risiko, energik, mempunyai rasa humor, memecahkan suatu masalah yang kompleks, artistik, berpikiran terbuka, dan intuitif (Lilly & Bramwell-Rejskind 2004). Guru yang kreatif akan meningkatkan motivasi siswa dan dirinya, selain itu dalam mengajar akan memanfaatkan kecerdasan personal dalam memilih kegiatan yang sesuai dengan minat dan kebutuhan siswa dan guru itu sendiri (Bramwell, Reilly, Lilly, Kronish, & Chennabathni, 2011). Pengalaman kerja akan membantu guru menemukan kreativitas sendiri, dengan kreativitasnya tersebut guru dapat membantu kesulitan siswa sesuai dengan kendala yang terjadi di kelas (Dawson, Tan, & McWilliam 2011).

Seorang guru harus paham kondisi pergaulan peserta didik. Dalam pergaulan antarpeserta didik kadang bahkan sering terjadi *bullying*. Guru di samping sebagai pendidik juga sebagai psikolog, sebagai psikolog dapat mengidentifikasi anak-anak yang agresif dan pasif yang dapat digunakan sebagai bahan menentukan langkah-langkah yang tepat untuk membantu anak-anak tersebut (Leff, Kupersmidt, & Power, 1999). Pelaksanaan aturan sanksi tegas terhadap kekerasan atau intimidasi, adanya pujian bagi siswa yang berprestasi, peningkatan pengawasan selama istirahat akan tercipta kondisi iklim sekolah yang baik (Roth, Kanat-Maymon, & Bibi, 2010). Selanjutnya, guru dapat menciptakan suasana yang menyenangkan akan membuat pembelajaran lebih efektif (Strean,

2011). Kedalaman pemahaman guru terhadap kemampuan setiap siswa pada gilirannya akan mempengaruhi kualitas penilaian siswa (Stiggins, 1992).

Pada umumnya guru di samping mempunyai pekerjaan yang berat sebagai pendidik, juga dengan sukarela akan memberikan waktu ekstra untuk membantu dalam menangani berbagai masalah siswanya (Twemlow, et al, 2006). Guru memberikan tambahan waktu, pikiran dan tenaga bagi siswa, sekolah, dan pelajaran (Choi, 2009). Guru mengembangkan pembelajaran di sekolah, mengajar merupakan tuntutan jiwa pendidik, mengajar sebagai karir, dan berinteraksi dengan siswa dengan *attitude* yang baik (Kwok-wai, 2006). Guru yang mempunyai komitmen akan berperilaku disiplin, bekerja keras-ulet, bertanggung jawab, bermotivasi kerja tinggi (Razak, Darmawan, & Keeves, 2010). Guru berkualitas dapat dilihat pada kualitas pengajaran dan pembelajaran walaupun masih tradisional yang diperolehnya selama bertahun-tahun (Gu & Day, 2013). Komunikasi guru dan keluarga secara berkala akan meningkatkan penyelesaian tugas di luar sekolah, meningkatkan kualitas perilaku dan partisipasi peserta didik di kelas (Kraft & Dougherty, 2013).

Kecerdasan emosi merujuk pada kemampuan seseorang untuk memahami arti emosi diri sendiri dan hubungan dengan orang lain, dan dapat mengatasi masalah berdasarkan problemnya (Mayer, Caruso, & Salovey, 2000). Guru adalah orang empati yang dapat memahami perasaan siswa dan menanggapi perasaan mereka dengan tepat (Olson & Wyett, 2000). Guru yang cerdas emosinya lebih sesuai sebagai pemimpin. Guru dapat mengelola secara efektif rasa marah atau frustrasi, dan dapat berperilaku adaptif sesuai dengan kondisi situasi yang berbeda. Jika menghadapi siswa yang bermasalah akan menanggapi siswanya dengan tepat (Iordanoglou, 2007). Ketika ada siswa mengganggu siswa lain, guru yang bijaksana segera menangani masalah tersebut dengan baik tanpa menimbulkan masalah baru (Twemlow, et al, 2006). Kepedulian guru yang baik terhadap siswa akan

mempengaruhi perilaku siswa kepada guru tersebut (Teven, 2007). Kecerdasan emosi (Patton, 1977; Sundari, 2013) terdiri atas lima komponen yaitu mengenali emosi sendiri, mengelola emosi, empati, dan keterampilan sosial.

Instrumen bakat keguruan terdiri atas kreativitas pedagogi dengan indikator-indikator kecepatan berpikir, keluwesan berpikir, keaslian berpikir, dan elaborasi berpikir. Kecepatan berpikir adalah tingkat kemampuan memproduksi alternatif pemecahan pendidikan dalam batas waktu yang ditentukan. Keluwesan berpikir adalah kemampuan mencari alternatif pemecahan masalah pendidikan dalam waktu yang ditentukan. Keaslian berpikir adalah kemampuan memproduksi alternatif masalah yang baru atau bersifat sintesa dalam waktu yang ditentukan. Elaborasi berpikir adalah kemampuan mencari alternatif lain yang lebih luas pada pemecahan masalah yang sudah ada sebelumnya pada waktu yang ditentukan.

Komitmen pedagogi terdiri atas empat faktor yaitu motivasi terhadap tugas, disiplin terhadap tugas, tanggung jawab terhadap tugas, dan keuletan menjalankan tugas. Motivasi terhadap tugas adalah dorongan dari dalam dan luar untuk menyelesaikan tugas yang diembannya. Disiplin terhadap pelaksanaan tugas adalah tingkat ketepatan waktu yang diperlukan untuk melaksanakan tugas yang diembannya. Tanggung jawab terhadap tugas adalah tingkat keberanian menanggung beban yang diembannya. Keuletan dalam menjalankan tugas adalah tingkat kegigihan pelaksanaan tugas yang diembannya.

Kecerdasan emosi adalah kemampuan mengenali, mengelola emosi diri sendiri sehingga tercipta keharmonisan hubungan dengan orang lain. Indikator kecerdasan emosi adalah emosi sendiri, mengelola emosi, memotivasi, empati terhadap orang lain, dan membina hubungan dengan orang lain.

Terkait dengan hal tersebut, penelitian ini bertujuan untuk mengembangkan instrumen bakat keguruan yang valid dan reliabel.

## Metode Penelitian

Pengembangan instrumen ini mengacu pada metode yang dikembangkan oleh Gable (1986, pp. 170-177), terdiri atas lima belas tahapan yang dapat dikelompokkan tiga bagian yaitu pra-pengembangan, pengembangan konsep, dan uji coba. Tahap pra pengembangan dilakukan kajian pustaka, sehingga diperoleh kerangka konsep bakat keguruan. Tahap pengembangan konsep bakat keguruan dikaji melalui *focus group discussion* dan validasi ahli yang terdiri atas ahli pengukuran, pendidikan, bahasa, dan psikologi. Tahap uji coba dilakukan untuk menguji kesesuaian konsep bakat keguruan dengan data lapangan. Uji terbatas dilakukan pada 3 mahasiswa Universitas Negeri Yogyakarta untuk mengetahui keterbacaan instrumen, waktu yang dibutuhkan untuk menjawab pertanyaan dan pernyataan instrumen, sedangkan uji coba lapangan dilakukan di FKIP Universitas Bengkulu dengan jumlah responden 630 mewakili seluruh program studi yang ada.

Kreativitas pedagogi diukur dengan Tes Kreativitas Verbal (TKV) seperti yang dilakukan oleh Munandar (1999, p.81). Butir-butir pertanyaan kreativitas pedagogi disusun model *what if not* yaitu metode untuk mengungkap kemampuan siswa dalam pemecahan masalah jika disajikan satu alternatif pemecahan masalah (Shriki, 2013; Silver, 1997). Komitmen pedagogi dan kecerdasan emosi diukur dengan inventori model Guttman (Nunnally, 1978, pp.45-46; Rankin, Knezek, Wallace & Zhang, 2004, p.99) dengan tiga pilihan berjenjang dengan penskoran 1 sampai dengan 3, dengan penempatan urutan skor secara acak, agar tidak mudah ditebak skornya.

Analisis validitas isi menggunakan rumus *scale content validity index* (S-CVI) dari Linn (1990) dan Polit & Beck (2006). Kualitas butir instrumen setiap dimensi bakat keguruan dianalisis dengan pendekatan *item response theory partial credit model* atau IRT PCM (Adams & Khoo, 1996, p.26; Andrich, 1978). Untuk menguji konstruk bakat keguruan dilakukan dengan *confirmatory factor analysis* (CFA) dari Kirsch &

Guthrie (1980) dan Shavelon & Stanton (1975). Model konstruk fit jika memenuhi minimal 3 kriteria, seperti yang dikemukakan oleh Joreskog & Sorbom (2003) dan Schermelleh-Engel, Moosbrugger & Mueller (2003). Reliabilitas instrumen bakat keguruan diperoleh dengan menggunakan reliabilitas gabungan (Mardapi, 2012, p.93; Nunnally, 1978, p. 248). Bukti validitas instrumen diperkuat dengan pendekatan validitas eksternal yaitu validitas konkuren dan konvergen. Kriteria eksternal instrumen kreativitas pedagogi menggunakan tes kecerdasan umum (IQ), instrumen komitmen pedagogi, kecerdasan emosi menggunakan instrumen *Edward Preference Personality Schedule* (EPPS). Validitas konvergen diperoleh dengan pendekatan *multitrait multimethod matrix* (Campbell & Fiske, 1959) dengan kriteria eksternal pengamatan (observasi) komitmen pedagogi dan kecerdasan emosi.

### Hasil Penelitian dan Pembahasan

Hasil pengembangan instrumen diperoleh konstruk bakat keguruan yang terdiri atas tiga dimensi, yaitu kreativitas pedagogi, komitmen pedagogi, dan kecerdasan emosi. Konstruk bakat keguruan mempunyai validitas isi sebesar 0,93. Kreativitas pedagogi terdiri atas 4 faktor yaitu kecepatan berpikir, keluwesan berpikir, keaslian berpikir, dan elaborasi berpikir. Masing-masing faktor terdiri atas 8 butir pertanyaan uraian terbatas, sehingga jumlah seluruhnya 32 butir pertanyaan. Dimensi kreativitas pedagogi mempunyai validitas isi sebesar 0,92. Komitmen pedagogi terdiri atas motivasi terhadap tugas, penyelesaian tugas, tanggung jawab terhadap tugas, dan keuletan menjalankan tugas. Setiap faktor terdiri atas 8 butir pernyataan yang diikuti 3 pilihan jawaban bertingkat. Dimensi komitmen pedagogi mempunyai validitas isi sebesar 0,94. Kecerdasan emosi terdiri atas 5 faktor yaitu mengenali emosi diri sendiri, mengelola emosi, motivasi diri, empati, dan hubungan sosial. Setiap faktor terdiri atas 8 butir pertanyaan, sehingga jumlah seluruhnya 40 butir. Kecerdasan emosi mempunyai validitas isi sebesar 0,97.

Uji coba lapangan untuk membuktikan kesesuaian antara konsep dan data empirik. Hasil analisis butir dengan pendekatan IRT PCM. Hasil analisis menunjukkan bahwa nilai infit instrumen kreativitas pedagogi bergerak dari 0,88 sampai dengan 1,23, masih dalam rentang batas *infit mean square* (MNSQ) untuk butir yaitu 0,77 sampai dengan 1,3. *Infit mean square* diperoleh mean 1,01, standar deviasi 0,07. *Outfit mean square* diperoleh mean 1,01 dan standar deviasi 0,06. Infit t dengan rata-rata 0,09 dan standar deviasi 1,05. Hasil estimasi testi *infit mean square* diperoleh 1,02 dengan standar deviasi 0,04. Infit t dengan rata-rata -0,04 dan standar deviasi 1,0. *Outfit mean square* -0,08 dan standar deviasi 1,06. Koefisien reliabilitas sebesar 0,94. Nilai informasi kemampuan kreativitas pedagogi antara -2,05 sampai dengan +1,4.

Nilai infit komitmen pedagogi bergerak dari nilai infit terkecil 0,77 dan terbesar 1,18 masih di dalam interval syarat *infit* mulai dari 0,77 sampai dengan 1,3. Bahwa estimasi butir *infit mean square* diperoleh mean 0,99 dan SD sebesar 0,08, sedangkan *outfit mean square* diperoleh mean 0,98 dan SD sebesar 0,07. *Infit t* dengan mean sebesar 0,15 dan SD sebesar 1,05, sedangkan *outfit t* dengan mean sebesar 0,08 dan SD sebesar 0,17. Koefisien reliabilitas instrumen komitmen pedagogi sebesar 0,707. Nilai informasi kemampuan (*ability*) komitmen pedagogi sampai +0,01.

Kecerdasan emosi mempunyai nilai infit terkecil 0,86 dan terbesar 1,28 masih di dalam interval syarat infit 0,77 sampai dengan 1,3, sedangkan estimasi butir *infit mean square* diperoleh rata-rata sebesar 0,99 dan *standard deviation* sebesar 0,08, sedangkan *outfit mean square* diperoleh mean 0,97 dan SD sebesar 0,16. *Infit t* dengan mean sebesar 0,19 dan SD sebesar 1,01, sedangkan *outfit t* dengan mean sebesar -0,08 dan SD sebesar 1,06. Hasil estimasi testi menunjukkan *infit means square* 1,00 dan standar deviasi 0,04, *outfit means square* 0,97 dan standar deviasi 0,03. Kecerdasan emosi mempunyai nilai informasi responden dengan kemampuan (*ability*) sampai +0,4. Ketiga dimensi bakat

keguruan menurut Adams & Koo (1996, p. 91) semua instrument dimensi bakat keguruan baik butir maupun testi model cocok dengan IRT PCM.

Tabel 1. Hasil analisis CFA

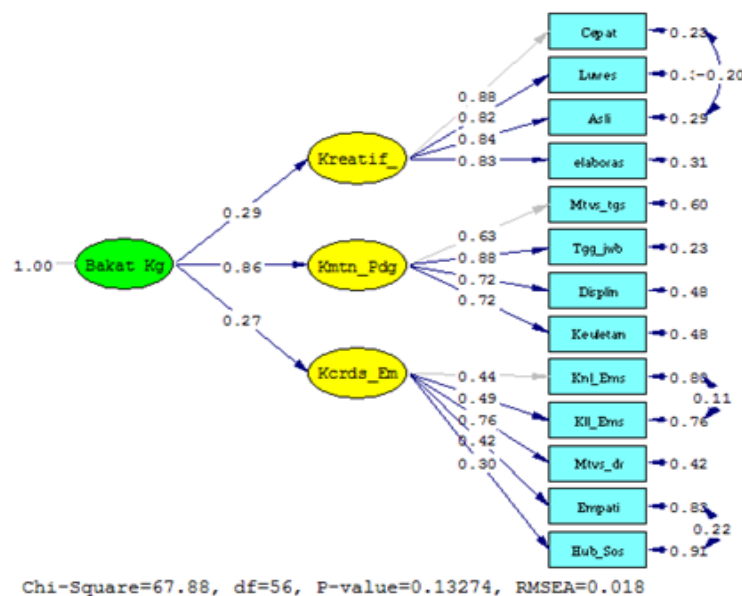
No	Kriteria GOF	Cut off value	Hasil penelitian	kecocokan
1.	df	-	56	-
2.	Chi-Square	≤ 112	67,88	baik
3.	p-value	≤ 0,05	0,13274	baik
4.	RMSEA	≤ 0.08	0,018	baik
5.	RMR	≤ 0,05	0,0048	baik
6.	NFI	≥ 0,90	0,98	baik
7.	NNFI	≥ 0,90	1,00	baik
8.	IFI	≥ 0,90	1,00	baik
9.	GFI	≥ 0,90	0,98	baik
10.	AGFI	≥ 0,90	0,97	baik

Hasil pengujian *second order* CFA konstruk bakat keguruan diperoleh 13 variabel manifes yang memiliki nilai t hitung lebih besar dari 1,96 dan nilai koefisien faktor lebih besar dari 0,5, namun ada empat variabel manifes dengan muatan faktornya kurang dari 0,5. Variabel tersebut adalah mengenali emosi diri sendiri dengan muatan faktor 0,44, mengelola emosi diri sendiri dengan muatan faktor 0,49 dan nilai t sebesar 8,00, empati terhadap orang lain dengan muatan faktor 0,42 dan nilai t sebesar 6,50,

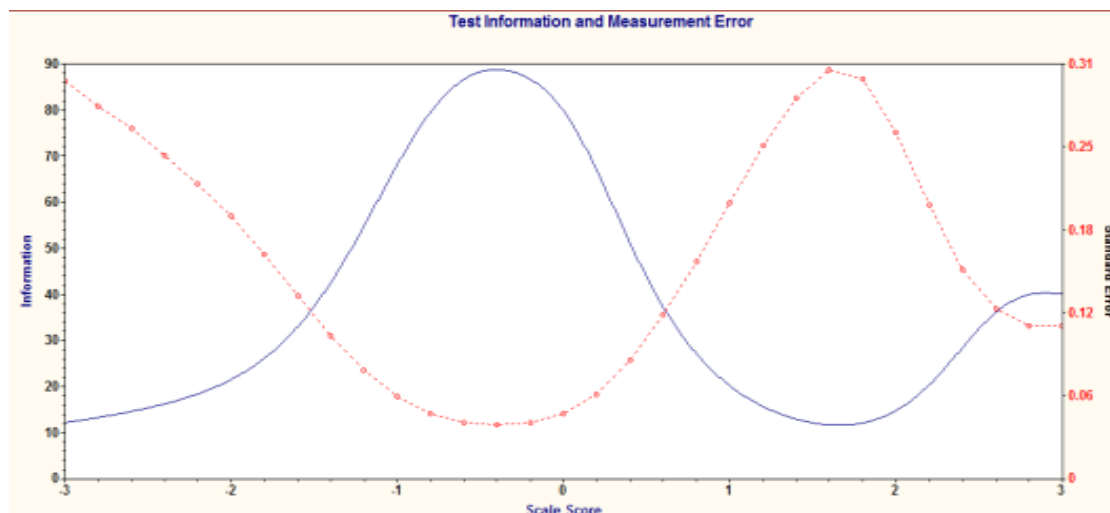
serta hubungan sosial dengan muatan faktor 0,30 dan nilai t sebesar 5,31.

Analisis model CFA menghasilkan Chi-square sebesar 67,88, df sebesar 56, p sebesar 0,13274, dan RMSEA sebesar 0,018. Hasil pengujian konstruk bakat keguruan memenuhi syarat *goodness of fit statistics* (GOF). Hubungan antar variabel manifes dengan variabel laten instrumen bakat keguruan disajikan dalam Gambar 1.

Muatan faktor variabel laten kreativitas pedagogi sebesar 0,29 dengan nilai t sebesar 3,06 dengan  $p < 0,05$ . Muatan faktor komitmen pedagogi sebesar 0,86 dengan nilai t sebesar 3,27 dengan  $p < 0,05$ . Muatan faktor kecerdasan emosi sebesar 0,27 dengan nilai t sebesar 2,72 dengan  $p < 0,05$ . Konstruk penyusun bakat keguruan hanya komitmen pedagogi dengan muatan faktor lebih besar 0,30, namun kecerdasan emosi dan kreativitas pedagogi kedudukan sebagai penyusun konstruk bakat keguruan signifikan karena  $P < 0,05$ . Dengan demikian, konstruk bakat keguruan terdiri atas kerativitas pedagogi, komitmen pedagogi, dan kecerdasan emosi didukung oleh data empirik. Dari ketiga dimensi bakat keguruan ternyata komitmen pedagogi mempunyai muatan faktor paling besar, berikutnya dimensi kecerdasan emosi, dan terakhir kreativitas pedagogi. Hasil analisis ini ternyata bakat keguruan didominasi oleh dimensi non-kognitif.



Gambar 1. Hasil Uji CFA Konstruk Bakat Keguruan



Gambar 2. Nilai Total Informasi

Berdasarkan Gambar 2, nilai fungsi informasi bakat keguruan antara -2,5 sampai dengan +0,7, dan di atas 2,58. Hal ini berarti instrumen bakat keguruan sesuai digunakan pada individu dengan kemampuan (*ability*) antara -2,5 sampai +0,7 dan di atas 2,58. Nilai informasi antara -2,5 sampai +0,7 sekitar 0,88. Nilai informasi antara +0,7 sampai dengan +2,5 terjadi penurunan informasi, hal ini berarti individu yang mempunyai kemampuan pada interval ini tidak tepat mengerjakan soal-soal instrumen ini. Fungsi informasi di atas 2,5 berkisar 0,09.

Tabel 2. Validitas Konkuren

No Variabel	SPM	EPPS
1. Kreativitas pedagogi	0,163	-
p	0,379	
2. Komitmen pedagogi	-	0,355
p		0,050*
3. Kecerdasan emosi	-	0,459
p		0,009*

Tabel 2 menunjukkan bahwa uji validitas konkuren antara kreativitas pedagogi dengan tes *SPM* sebesar 0,163. Hubungan antara skor kreativitas pedagogi dengan skor ini tidak signifikan, hal ini ditunjukkan besarnya *p* yang muncul 0,379. Korelasi sebesar 0,163 ini menunjukkan bahwa antara kreativitas dan pedagogi merupakan konsep yang berbeda. Korelasi antara kreativitas

pedagogi dengan kecerdasan (*IQ*) sebesar 0,163 membuktikan bahwa korelasi ini sangat kecil, dibanding dengan kriteria validitas eksternal yang mensyaratkan minimal 0,350. Instrumen bakat keguruan yang terdiri dari kreativitas pedagogi, komitmen pedagogi, dan kecerdasan emosi.

Berdasarkan Tabel 3 korelasi antara hasil tes kreativitas dengan kecerdasan umum *IQ* sebesar 0,163.

Tabel 3. Rangkuman Multitrait Multimethod

Dimensi	IQ	Obs KP	Obs KE
Kreativitas Pedagogi	0,163	-	-
Komitmen Pedagogi (KP)		0,832	-
Kecerdasan Emosi (KE)			0,806

Banyak penelitian tentang hubungan antara inteligensi dan kreativitas yang menunjukkan bahwa kreativitas belum tentu merupakan fungsi dari inteligensi (Renzulli, 1978). Kim (2005) melakukan meta penelitian hubungan antara kreativitas dan kecerdasan. Sampel 447 koefisien korelasi dari 21 studi dengan sampel 45.880 orang. Koefisien korelasi rata-rata yang kecil  $r = 0,174$ . Terdapat korelasi yang rendah antara kreativitas dengan kecerdasan (Silvia, 2008). Korelasi antara kreativitas pedagogi dengan kecerdasan yang rendah ini sesuai dengan penelitian yang dilakukan oleh Naderi & Abullah (2010) bahwa kreativitas menjelas-

kan *error variance* kecerdasan sebesar 0,135. Hal ini berarti hubungan antara kreativitas dan kecerdasan sangat rendah. Hasil penelitian yang dilakukan oleh Cho et al. (2010) menunjukkan bahwa antara proses operasi berpikir kreatif berbeda dengan proses berpikir cerdas. Fuchs-Beauchamp, Karnes, & Johnson (1993) menyebutkan bahwa ada korelasi antara kecerdasan dan kreativitas. Anak-anak dengan IQ kurang dari 120 korelasinya moderat, sedangkan anak-anak dengan IQ di atas 120 korelasi dengan kreativitas menurun drastis sebesar 0,12. Jauk et al. (2013) menemukan bahwa kreativitas sangat dipengaruhi oleh personalitas, tidak ditemukan bahwa kreativitas yang tinggi ditentukan oleh kecerdasan yang tinggi. Penelitian pada masyarakat Korea yang dilakukan oleh Kim (2009) menyebutkan bahwa kreativitas dipengaruhi oleh budaya. Korelasi yang rendah ini menunjukkan bahwa secara empirik kreativitas pedagogi berbeda dengan konsep kecerdasan (IQ). Dari berbagai penelitian dan pendapat tersebut, ternyata konsep kreativitas berbeda dengan konsep kecerdasan IQ. Hal ini menunjukkan bahwa kreativitas dan kecerdasan menunjukkan konsep yang berbeda, temuan ini mendukung konsep keberbakatan *the three ring concept* dari Renzulli.

Korelasi hasil pengukuran komitmen pedagogi dengan inventori dan hasil observasi komitmen pedagogi sebesar 0,832, hal ini menunjukkan bahwa antara pengukuran inventori dan observasi sesuai *convergent*. Korelasi sebesar 0,832 tersebut menunjukkan bahwa instrumen dimensi komitmen pedagogi mempunyai validitas konvergen. Korelasi hasil pengukuran kecerdasan emosi dengan inventori dengan hasil observasi sebesar 0,806, hal ini menunjukkan bahwa kedua pengukuran tersebut mengukur satu konsep yang sama *convergent*. Dengan demikian, kecerdasan emosi memiliki validitas konvergen.

Penelitian yang dilakukan oleh Kadyschuk (1997) menyebutkan bahwa komitmen guru terdiri atas pengalaman kerja, karakteristik kerja, dan kepemimpinan. Hal ini berbeda dengan komitmen pedagogi

yang dikembangkan dalam penelitian ini. Perbedaannya terletak pada cakupan masalah yang diteliti, yaitu motivasi terhadap tugas yang diembannya, tanggung jawab terhadap tugas, disiplin terhadap tugas, dan keuletan/ketangguhan dalam menjalankan tugas yang berkaitan dengan tugas guru sebagai pendidik. Perbedaan lainnya adalah bentuk pernyataannya. Kadyschuk mengembangkan instrumen dalam bentuk model Likert, sedangkan dalam penelitian ini bentuk butir-butirnya berupa pernyataan yang diikuti dengan pilihan bergradasi seperti yang dikembangkan oleh Guttman (Rankin, Knezek, Wallace & Zhang, 2004).

Penelitian yang dilakukan oleh Justice & Espinoza (2007) menyebutkan bahwa kecerdasan emosi berkaitan dengan keterampilan calon guru dalam hal kenyamanan, empati, pengambilan keputusan, mendorong kekuatan, manajemen waktu, komitmen etika, harga diri, manajemen stress, dan mengubah orientasi. Ada hal yang berbeda dengan penelitian ini yaitu komitmen etika masuk dalam komitmen pedagogi, dan keterampilan hubungan sosial yang ada dalam penelitian ini. Di samping itu, butir-butir yang dikembangkan dalam penelitian ini berbentuk pernyataan yang diikuti dengan pilihan bergradasi.

Penelitian yang dilakukan oleh Corcorana & Tormey (2013) tentang hubungan antara kecerdasan emosi dengan kinerja praktikum mengajar. Tidak semua faktor kecerdasan emosi yang berhubungan signifikan dengan praktik mengajar, namun faktor *perception appraisal and expression of emotion* (PEIQ) mempunyai hubungan yang signifikan dengan kinerja praktik mengajar. Hal ini sesuai dengan hasil penelitian ini bahwa kecerdasan emosi merupakan dimensi dari bakat keguruan.

Berdasarkan uraian tersebut, ada perbedaan pada antara komitmen pedagogi dengan komitmen guru yang dikembangkan oleh Kadyschuk (1997), dan antara kecerdasan emosi dengan kecerdasan emosi yang dikembangkan oleh Justice & Espinoza (2007). Kreativitas pedagogi yang dikembangkan dalam penelitian ini belum di-

kembangkan oleh peneliti lain. Dengan demikian ada kebaruan dalam penelitian ini.

Hasil pengukuran bakat keguruan (BKg) mahasiswa FKIP Universitas Bengkulu tahun akademik 2014/2015 disajikan pada Tabel 4.

Tabel 4. Distribusi Frekuensi BKg

No	Interval	f
1.	170 -181	5
2.	182 -193	13
3.	194 -205	51
4.	206 -217	91
5.	218 -229	126
6.	230 -241	123
7.	242 -253	96
8.	254 -265	76
9.	266 -277	37
10.	278 -289	12
	Jumlah	630

Berdasarkan Tabel 4, dapat diketahui bahwa skor 104 sampai 184 atau nilai 1 sampai dengan 3,9 (kurang sekali berbakat) terdapat 8 mahasiswa atau 1,27%. Skor 185 sampai dengan 228 atau nilai 4 sampai dengan 5,9 (kurang berbakat) terdapat 261 mahasiswa atau 41,58%. Skor 229 sampai 269 atau nilai 6 sampai dengan 7,9 (cukup berbakat) sebanyak 262 mahasiswa atau 41,58%. Skor 270-312 atau nilai 8 sampai dengan 10 (berbakat) sebanyak 36 mahasiswa atau 5,72%.

### Simpulan dan Saran

Berdasarkan hasil analisis dan pembahasan, pengembangan instrumen bakat keguruan dapat disimpulkan berikut ini. Instrumen bakat keguruan yang dikembangkan tergolong baik, karena mempunyai indeks validitas isi yang lebih besar daripada batas minimal yang ditentukan, didukung dengan data yang memenuhi syarat *goodness of fit statistics*, dan mempunyai koefisien reliabilitas yang tinggi.

Instrumen kreativitas pedagogi, komitmen pedagogi, dan kecerdasan emosi merupakan dimensi bagian instrumen bakat

keguruan tergolong baik, karena mempunyai indeks validitas isi dan koefisien reliabilitas yang tinggi.

Instrumen kreativitas pedagogi mempunyai validitas konkuren dengan *standard progresive matrices (SPM)* yang rendah, sedangkan dimensi komitmen pedagogi dan kecerdasan emosi mempunyai indeks validitas konkuren dengan EPPS yang cukup. Instrumen komitmen pedagogi dan instrumen kecerdasan emosi mempunyai validitas konvergen yang tinggi.

Kondisi bakat keguruan mahasiswa FKIP Universitas Bengkulu mempunyai distribusi frekuensi normal, jumlah mahasiswa yang mempunyai bakat keguruan kategori rendah dan katagori tinggi jumlahnya sedikit, sedangkan jumlah mahasiswa berbakat katagori sedang jumlahnya paling banyak.

Penelitian belum mencakup batas skor potensi bakat keguruan, oleh karena itu perlu dilakukan penelitian lanjutan tentang *cut score* bakat keguruan.

### Daftar Pustaka

- Açikgöz, F. (2005). A studi on teacher characteristics and their effects on students attitudes. *The Reading Matrix*, 5(5), 103-115.
- Adams, R. J., & Khoo Siek Toon. (1996). *QUEST. The Interactive Test Analysis System*. Camberwell, Victoria: The Australian Council for Educational Research.
- Alkharusi, H., Kazem, A. M., & Musawai, A. A. (2011). Knowledge, skills, and attitudes of preservice and in-service teachers in educational measurement. *Asia-Pacific Journal of Teacher Educational* 39(2), 113–123.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psichometrika*, 43(4), 561-573.
- Anonim. (1977). *Piagam dan Peraturan besar Persatuan Taman Siswa*. Yogyakarta: Majelis Luhur Persatuan Taman Siswa.

- Beetlestones, F. (2011). *Creative Learning*. Bandung: Nusa Media.
- Bingham, W. V. (1948). What is an aptitude? *Test Bulletin Service*, 36, 1-3.
- Block, A. A. (2008). Why should I be a teacher? *Journal of Teacher Education*, 59, 416-427.
- Bramwell, G., Reilly, R., Lilly, F. R., Kronish, N., & Chennabathni, R. (2011). Creative teachers. *ProQuest*, 33(4), 228-237.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105.
- Cho, S. H., te Nijenhuis, J, van Vianen, A. E. M., Kim, H. B., & Lee, K. H. (2010). The relationship between diverse components of intelligence and creativity. *Journal Creative Behavior*, 44(2), 125-137.
- Choi, P. L., & Tang, S. Y. F. (2009). Teacher commitment trends: Cases of Hong Kong teachers from 1997 to 2007. *Teaching and Teacher Education* 25, 767-777.
- Corcorana, R. P., & Tormey, R. (2013). Does emotional intelligence predict student teachers' performance? *Teaching and Teacher Education*, 35, 34-42.
- Dawson, S., Tan, J. P. L., & McWilliam, E. (2011). Measuring creative potential: Using social network analysis to monitor a learners' creative capacity. *Australasian Journal of Educational Technology*, 27(6), 924-942.
- Dicky, N. (2011). Indonesian primary teachers' mathematical knowledge for teaching geometry: implications for educational policy and teacher preparation programs. *Asia-Pacific Journal of Teacher Educational*, 39(2), 151-164.
- Dorman, J. (2008). Testing a model for teacher burnout. *Australian Journal of Educational & Developmental Psychology*, 3, 35-47.
- Fuchs-Beauchamp, K. D., Karnes, M. B., & Johnson, L. J. (1993). Creativity and Intelligence in Preschoolers. *Gifted Child Quarterly*, 37(3), 113-117.
- Gable, R. K. (1986). *Instrument development in affective domain*. Boston: Kluwer-Nijhoff Publishing.
- Garcia, P., Kupczynski, L., & Holland, G. (2011). Impact of teacher personality styles on academic excellence of secondary students. *National Forum of Teacher Educational Journal*, 21(2).
- Gardner, H. (2006). *Multiple Intelligences*. New York: Basic Book.
- Gu, Q., & Day, C. (2013). Challenges to teacher resilience: conditions count. *British Educational Research Journal*, 39(1), 22-44.
- Iordanoglou, D. (2007). The relationship between emotional intelligence and leadership effectiveness, commitment, and satisfaction. *Journal of Leadership Studies*, 1(3), 57-66.
- Ishak, N. M., Piet, I. I. & Ramli, R. (2010). Emotional intelligence of Malaysian teacher : A comparative study on teacher in daily and residential school. *Procedia- Social and Behavior Sciences*, 9, 604-612.
- Jacobs, K., & Mitchell, J. (2008). Emotional intelligence: hidden ingredient for emotional health of teachers. *Africa Education Review* 5(11), 131-143.
- Jauk, E., Benedek, M., Dunst, B., & Neubauer, A. C. (2013). The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. *Intelligence*, 41, 212-221.

- Joreskorg, K., & Sorbom, D. (2003). *Lisrel 8: User's reference guide*. Chicago: Scientific software international.
- Justice, M., & Espinoza, S. (2007). Emotional intelligence and beginning teacher candidates. *Education; ProQuest*, 127(4), 456-461.
- Kadyschuk, R. (1997). *Teacher commitment: a study of organizational commitment, professional commitment, and union commitment of teachers in public school in Saskatchewan*. University of Saskatchewan, Canada. Diunduh pada tanggal 21 mei 2012 dari <http://ecommons.usask.ca/handle/10388/etd-10212004-000011>
- Kim, K. H. (2005). Can Only Intelligent People Be Creative? *The Journal of Secondary Gifted Education XVI* (2/3), 57-66.
- Kim, K. H. (2009). The relationship between Asian culture and creativity. *Journal of Creative Behavior*, 43 ( 2 ), 73-93.
- Kirsch, I. S., & Guthrie, J. T. (1980). Construct validity of functional reading tests. *Journal of educational measurement*, 1(7), 281-293.
- Kokkinos, C. M., Charalambous, K., & Davazoglou, A. (2010). Primary school teacher interpersonal behavior through the lens of students' Eysenckian personality traits. *Soc Psychol Educ.*, 13, 63-349
- Kraft, M. A., & Dougherty, S.M. (2013). The effect of teacher-family communication on student engagement: Evidence from a randomized field experiment. *Journal of Research on Educational Effectiveness*, 6, 199-222.
- Kwok-wai, C. (2006). In-service teachers' motives and commitment in teaching. *Hong Kong Teachers' Centre Journal*, 5, 112-128.
- Leff, S. S., Kupersmidt, J. B., Patterson, C. J., & Power, T. J. (1999). Factors influencing teacher identification of peer bullies and victims. *School Psychology Review*, 28(3), 505-517.
- Lilly, F. R., & Bramwell-Rejskind, G. (2004). The dynamics of creativity teaching. *Journal of Creative Behavior*, 38, 102-124.
- Linn, R. L. (1990). *Measurement and evaluation in teaching*. New York: McMilan Publishing Company.
- Mardapi, D. (2012). *Pengukuran penilaian dan evaluasi pendidikan*. Yogyakarta: Nuha Litera.
- Mardiana. (2006). Komitmen guru dalam melaksanakan tugas mengajar di SD Kecamatan Padang Panjang Timur. *Jurnal guru*, 2(59-77).
- Martinson, R. A. (1974). *The Identification of The Gifted and Tallent*. California: Office of The Ventura County Superintended of Schools.
- Mayer, J. D., Caruso, D. R., & Salovey, P. (2000). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 27(4), 267-298.
- Munandar, S. C. U. (1999). *Kreativitas dan keberbakatan: Strategi mewujudkan potensi kreatif dan bakat*. Jakarta: Gramedia Pustaka Utama.
- Nunnally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill Book Company.
- Olson, C. O., & Wyett, J. L. (2000). Teachers need affective competencies. *Education*, 120(4), 471-625..
- Patton, P. (1977). *Emotional Intelligence*. Singapore: SNP Publishing Pte Ltd.
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29, 489-497.
- Rankin, P. D., Knezek, G. A., Wallace, S., & Zhang, S. (2004). *Scaling method*. New

- Jersey: Lawrence Erlbaum Associates Publisher.
- Razak, N. A., I., Darmawan, I. G. N., Keeves, J. P. (2010). The influence of culture on teacher commitment. *Soc Psychol Educ.*, 13, 185-205.
- Renzulli, J. S. (1978). What makes giftedness?: Reexamining a Definition. *Psi Delta Kappan*, 60(3), 180-184.
- Renzulli, J. S. (1986). *The three-ring conception of giftedness: A developmental model for creative productivity*. New York: Cambridge University Press.
- Renzulli, J. S. (1998). The three-ring conception of giftedness. In S. M. Baum, Reis, S. M., & Maxfield, L. R. (Ed.), *Nurturing the gifts and talents of primary grade students*. Mansfield Center, CT: Creative Learning Press.
- Roth, G., Kanat-Maymon, Y., & Bibi., U (2010). Prevention of school bullying: The important role of autonomy-supportive teaching and internalization of pro-social values. *British Journal of Educational Psychology* 81(4), 656-666.
- Rowe, K. J. (2003). *The Importance of Teacher Quality as a Key Determinant of Students' Experiences and Outcomes of Schooling*. Paper presented at the ACER Improving Learning, Melbourne.
- Salkind, N. J. (2008). Aptitude *Encyclopedia of educational psychology*. New York: Sage Publisher.
- Sawyer, R. K. (2004). Creative teaching: collaborative discussion as disciplined improvisation. *Educational Researcher*, 33(2), 12-20.
- Schermelleh-Engel, K., Moosbrugger, H., & Mueler, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research* 8(2), 23-74.
- Shavelon, R. J., & Stanton, G. C. (1975). Construct validation : methodology and application to three measures of cognitive struktur. *Journal Educational Measurement*, 12(2), 67- 85.
- Shriki, A. (2013). A model for assessing the development of students' creativity in the context of problem posing. *Creative Education*, 4(7), 430-439.
- Sigmar, L., Hynes, G., & Cooper, T. (2010). Emotional intelligence: Pedagogical considerations for skills-based learning in business communication courses. *Journal of Instructional Pedagogies*, 3(June), 1-13.
- Silver, E. A. (1997). Fostering creativity through Instruction rich in Mathematical problem solving and problem posing. *ZDM- The International Journal on Mathematics Education*, 29, 75-80.
- Silvia, P. J. (2008). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences*, 44, 1012-1021.
- Sternberg, R. J., & Davidson, J. E. (2005). *Conceptions of Giftedness*. Cambridge: Cambridge University Press.
- Stiggins, R. J. (1992). High quality classroom assessment: What does it really mean? *Educational Measurement: Issues and Practice*, 11(2), 35-39.
- Strean, W. B. (2011). Creating student engagement? HMM: Teaching and learning with humor, music, and movement. *Creative Education*, 2(3), 189-192.
- Sudarnoto, L. F. N. (2009). *Memberdayakan profesi pendidik untuk mencapai peningkatan kualitas kehidupan kerja*. Universitas Atma Jaya. Yogyakarta.
- Sudiyat, I. (1987). *Pamong yang berwatak satriya pandhita dan pinandhita sanatriyo & belajar menjadi siswa ke hadjar dewantara*. Yogyakarta: Sarjanawiyata Tamansiswa.
- Sugiharto. (2003). Kemampuan guru mengelola pembelajaran Fisika di SLTP

N Kota Banjarmasin. *Jurnal Penelitian dan Evaluasi*, 5(6), 116-127.

Sundari, A. (2013). Emotional intelligence for professional excellence. *International journal of the frontiers of english literature and the patterns of ELT* 1(2), 1-13.

Teven, J. S. (2007). Teacher Caring and Classroom Behavior: Relationships with Student Affect and Perceptions

of Teacher Competence and Trustworthiness. *Communication Quarterly*, 55(4), 433-450.

Twemlow, S. W., Fogarty, P., Sacco, F. C. & Brethour Jr, J.R. (2006). Teacher who bully students: A hidden trauma. *International Journal of Social Psychiatry*, 52(3), 187-198.

## PENGEMBANGAN INSTRUMEN DIAGNOSTIK KOGNITIF PADA MATA PELAJARAN IPA DI SMP

<sup>1)</sup>Yuli Prihatni, <sup>2)</sup>Kumaidi, <sup>3)</sup>Mundilarto

<sup>1)</sup>Universitas Sarjanawiyata Tamansiswa, <sup>2)</sup>Universitas Muhammadiyah Surakarta,

<sup>3)</sup>Universitas Negeri Yogyakarta

<sup>1)</sup>yuliku7781@gmail.com, <sup>2)</sup>kuma\_426@yahoo.com, <sup>3)</sup>mundilarto@uny.ac.id

### Abstrak

Penelitian ini bertujuan untuk: (1) menghasilkan instrumen diagnostik kognitif pada mata pelajaran IPA materi kalor berdasarkan *learning continuum*, (2) menemukan karakteristik instrumen diagnostik kognitif pada mata pelajaran IPA materi kalor, berdasarkan *learning continuum*. Metode yang digunakan dalam penelitian ini adalah metode pengembangan tes diagnostik. Penelitian didahului dengan penyusunan *learning continuum*, kisi-kisi tes, hierarki materi prasyarat dan spesifikasi *item* pada materi IPA konsep kalor yang ditelaah melalui *Focus Group Discussion* (FGD). Uji coba dilaksanakan dengan subjek uji sebanyak 484 siswa kelas VII SMP di Kabupaten Sleman. Seleksi butir tes untuk mendapatkan fit tes menggunakan program *Quest* dan Program M Plus untuk uji kecocokan/kesesuaian model atribut dengan matriks Q. Hasil penelitian menunjukkan: (1) instrumen yang berhasil dikembangkan berbentuk tes pilihan ganda dengan alasan, berjumlah 28 butir soal dan menghasilkan tujuh Q matrik; (2) analisis dengan program *Quest* diperoleh butir fit dengan model dan program M Plus menemukan besarnya probabilitas pada setiap *latent class* pada tujuh Q matriks yang tersusun.

**Kata kunci:** *instrumen, diagnostik kognitif, IPA*

## DEVELOPING THE INSTRUMENTS OF COGNITIVE DIAGNOSTIC FOR SCIENCE SUBJECT IN JUNIOR HIGH SCHOOL

<sup>1)</sup>Yuli Prihatni, <sup>2)</sup>Kumaidi, <sup>3)</sup>Mundilarto

<sup>1)</sup>Universitas Sarjanawiyata Tamansiswa, <sup>2)</sup>Universitas Muhammadiyah Surakarta,

<sup>3)</sup>Universitas Negeri Yogyakarta

<sup>1)</sup>yuliku7781@gmail.com, <sup>2)</sup>kuma\_426@yahoo.com, <sup>3)</sup>mundilarto@uny.ac.id

### Abstract

The study aimed at: (1) producing cognitive diagnostic instruments for science subjects of heat topic based on learning continuum (2) finding the characteristics of the instrument developed. The method used in this research is the development of diagnostic tests. The research was preceded by the preparation of learning continuum, test blueprint, hierarchy of prerequisite materials and test specification on materials of science concept which were examined through Focus Group Discussion (FGD). The test trial was implemented with 484 students of class VII SMP in Sleman as subjects. Selection of test items to get fit tests used Quest program and M Plus program to test the suitability/fitness model with the attribute matrix Q. The results show: (1) the instruments developed in the form of a multiple choice test with reason, totaling 28 items and seven Q matrix; (2) analysis using Quest program finds item fit with the model, and program M Plus find the probability in each latent class on seven Q matrix composed.

**Keywords:** *instruments, cognitive diagnostic, science*

## Pendahuluan

Pelaksanaan pembelajaran Ilmu Pengetahuan Alam (IPA) di sekolah tidak lepas dari hakikat IPA itu sendiri. Fisika merupakan bagian dari IPA yang mempelajari gejala-gejala alam dan interaksinya. Pembelajaran IPA fisika adalah pembelajaran yang tidak mengabaikan hakikat IPA fisika sebagai sains (Taufik, Sukmadinata, Abdulkhak, & Tumbelaka 2010, p.33). Pelajaran fisika apabila melalui proses yang benar dapat melatih siswa mengembangkan daya pikir kritis dan sikap ilmiah. Sebagai ilmu pengetahuan yang mendasarkan pada pengamatan dan pengukuran, fisika dibawa kepada kecermatan, ketepatan, dan ketakterdugaan. Dalam mempelajari fisika, banyak siswa yang mengalami salah konsep (Suparno, 1998, p.96). Siswa cenderung hafal rumus fisika tanpa memahami konsepnya.

Rata-rata skor prestasi sains siswa Indonesia pada TIMSS 2007 adalah 433. Dengan skor tersebut siswa Indonesia menempati peringkat 35 dari 49 negara. Rata-rata skor siswa Indonesia pada TIMSS 2007 di bawah skor rata-rata yaitu 500, dan hanya mencapai *Low International Benchmark*. Capaian tersebut menunjukkan rata-rata siswa Indonesia hanya mampu mengenali sejumlah fakta dasar tetapi belum mampu mengkomunikasikan dan mengaitkan berbagai topik sains, apalagi menerapkan konsep-konsep yang kompleks dan abstrak (Efendi, 2010, p.385).

Kesulitan dan kelemahan siswa dalam bidang sains penting untuk dipecahkan. IPA merupakan ilmu pengetahuan alam yang berkaitan dengan cara mencari tahu tentang alam secara sistematis, sehingga IPA bukan hanya penguasaan kumpulan pengetahuan yang berupa fakta-fakta, konsep-konsep atau prinsip-prinsip saja tetapi juga merupakan suatu proses penemuan. Kemampuan guru dalam mendeteksi kesulitan dan kelemahan siswa merupakan tantangan dalam kegiatan pembelajaran IPA untuk mengantarkan siswa mencapai keberhasilan dalam belajar. Setiap siswa dalam kelas mempunyai karakteristik yang berbeda-beda. Perbedaan individual siswa dapat disebabkan pada cara ber-

pikirnya, oleh karena itu perbedaan individu perlu diperhatikan oleh guru dalam kegiatan pembelajaran. Pelaksanaan pembelajaran di kelas umumnya dilaksanakan dengan melihat siswa sebagai individu dengan kemampuan rata-rata dan kebiasaan yang kurang lebih sama. Kemampuan individual siswa kurang diperhatikan, sehingga siswa yang memiliki kelemahan akan terus tertinggal dalam pembelajaran. Proses belajar antara siswa yang memiliki kemampuan tinggi tentu tidak sama dengan siswa yang memiliki kemampuan sedang, bahkan dengan siswa yang kemampuannya kurang. Guru perlu memperhatikan individu siswa, kelemahan dan kesulitan siswa perlu diperhatikan. Pembelajaran IPA saat ini lebih menekankan pada runtutan materi pelajaran bukan pada proses berpikir dan psikologi kognitif siswa, sehingga dalam pembelajaran siswa banyak yang mengalami kesulitan dan salah konsepsi. Guru memegang peranan penting untuk mengatasi kesulitan siswa dan memperbaiki proses pembelajaran.

Kesulitan siswa dalam belajar erat kaitannya dengan perkembangan kognitif siswa. Menurut teori kognitif, belajar menunjukkan adanya jiwa yang sangat aktif, jiwa mengolah informasi yang diterima, tidak sekedar menyimpannya saja tanpa mengadakan transformasi (Gage & Berliner, 1984, p.267). Menurut teori ini anak memiliki sifat aktif, konstruktif, dan mampu merencanakan sesuatu. Anak mampu mencari, menemukan fakta, menganalisis, menafsirkan dan menarik kesimpulan. Berpikir pada umumnya diasumsikan sebagai proses kognitif, tindakan mental untuk memperoleh pengetahuan. Produk-produk berpikir seperti pikiran, pengetahuan, alasan serta proses yang lebih tinggi seperti penilaian dapat juga dihasilkan. Kaitan-kaitan kompleks dikembangkan melalui berpikir ketika digunakan sebagai bukti dari waktu ke waktu. Kaitan-kaitan ini dapat dihubungkan pada struktur yang terorganisasi dan diekspresikan oleh pemikir dalam beragam cara. Membangun model kognitif untuk mendesain ciri soal merupakan hal penting dalam pendekatan desain kognitif (Mislevy, 1994, pp.439-483).

Oleh karena itu, dalam penelitian ini penyusunan instrumen diagnostik kognitif memperhatikan proses berpikir siswa.

Pengembangan instrumen diagnostik berkaitan dengan kegiatan penilaian. Pada Permendiknas No 20 tahun 2007 tentang standar penilaian dijelaskan bahwa penilaian adalah proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar siswa. Penilaian tidak sekedar pengumpulan data siswa, tetapi juga pengolahannya untuk memperoleh gambaran proses dan hasil belajar siswa. Penilaian tidak sekedar memberi soal siswa kemudian selesai, tetapi guru harus menindaklanjutinya untuk kepentingan pembelajaran. Dunn, Morgan, Parry, & Reilly (2004, p.16) menyatakan bahwa tujuan dan peranan penilaian (*assessment*) dalam pembelajaran yaitu: (1) mendiagnosis kesulitan belajar siswa; (2) mengukur peningkatan dari waktu ke waktu; (3) menentukan penguasaan siswa terhadap pengetahuan dan keterampilan tertentu; (4) menentukan ranking siswa dari keseluruhan siswa dalam kelas; (5) mengevaluasi metode pembelajaran; (6) mengevaluasi efektivitas program pembelajaran.

Penilaian merupakan bagian terpenting dari proses pembelajaran. Tujuan penilaian memberikan masukan yang berupa informasi secara komprehensif tentang hasil belajar siswa, baik dilihat ketika saat kegiatan pembelajaran berlangsung maupun dilihat dari hasil akhirnya, dengan menggunakan berbagai cara penilaian sesuai dengan kompetensi yang diharapkan dapat dicapai siswa. Untuk menilai sejauhmana siswa telah menguasai beragam kompetensi, diperlukan berbagai jenis penilaian. Dengan demikian, penilaian merupakan rangkaian kegiatan untuk memperoleh data, menganalisis, dan menafsirkan data tentang proses dan hasil belajar siswa yang dilakukan secara sistematis dan berkesinambungan, sehingga menjadi informasi yang bermakna dalam pengambilan keputusan

Penilaian di kelas dilakukan oleh guru bertujuan untuk mengetahui kemajuan dan hasil belajar siswa, membandingkan kemampuan siswa, mendiagnosa kesulitan belajar,

memberikan umpan balik/perbaiki proses belajar mengajar, dan penentuan kenaikan kelas. Melalui penilaian dapat diperoleh informasi yang akurat tentang penyelenggaraan pembelajaran dan keberhasilan belajar siswa, guru, serta proses pembelajaran itu sendiri. Berdasarkan informasi itu, dapat dibuat keputusan tentang pembelajaran, kesulitan siswa dan upaya bimbingan yang diperlukan serta keberadaan kurikulum itu sendiri. Penilaian memiliki tujuan yang sangat penting dalam pembelajaran, diantaranya untuk *grading*, seleksi, mengetahui tingkat penguasaan kompetensi, bimbingan, diagnosis, dan prediksi.

Seorang guru harus memberikan peluang bagi siswa dan guru untuk mendapatkan dan menggunakan informasi tentang kemajuan menuju tujuan pembelajaran. Banyak kegiatan atau aktivitas yang terjadi di dalam kelas dilakukan guru dan siswa yang dapat digambarkan sebagai penilaian. Artinya, tugas dan pertanyaan siswa dapat digunakan untuk menunjukkan pengetahuan mereka, pemahaman dan keterampilan yang dimiliki siswa. Apa yang dilakukan siswa kemudian diamati dan diinterpretasikan dan dapat digunakan untuk memperbaiki proses pembelajaran. Proses penilaian ini merupakan bagian penting dari praktek di kelas yang dilaksanakan sehari-hari dan melibatkan guru dan siswa dalam refleksi, dialog dan pengambilan keputusan (Leung & Mohan 2004, pp.335-359).

Menurut Nitko (1996, pp.284-285) pada praktiknya penilaian diagnostik di kelas memiliki dua kawasan yaitu: (1) untuk mengidentifikasi target pembelajaran yang belum dikuasai siswa; dan (2) untuk menemukan penyebab-penyebab atau alasan-alasan yang membuat siswa belum dapat menguasai target-target pembelajaran. Penilaian yang dilakukan dalam penelitian ini menggunakan instrumen tes.

Penilaian dalam penelitian ini dilakukan dengan menggunakan instrumen yang berupa tes diagnostik. Dalam penelitian ini penilaian dikhususkan sebagai alat diagnosis kognitif siswa pada mata pelajaran IPA SMP, penilaian digunakan untuk menunjukkan

letak kelemahan dan kesulitan belajar yang dialami siswa dan kemungkinan prestasi yang bisa dikembangkan. Hal ini akan membantu guru untuk memberikan pembelajaran yang sesuai dengan proses berpikir siswa.

Mengamati kesalahpahaman dalam pembelajaran memainkan peran penting dalam pengembangan kemampuan mental anak-anak. Oleh karena itu, meningkatkan kemampuan belajar anak merupakan aspek penting dari mendidik anak (Afaq & Al-Mashari, 2010, p.76). Dengan demikian, perlu dirancang tes diagnostik yang dapat mengidentifikasi kelemahan dan kekuatan siswa dalam belajar. Tes diagnostik menurut Ebel (1979, p.375) adalah rancangan untuk mengetahui kekurangan-kekurangan yang khusus atau kegagalan dalam belajar dalam beberapa subjek atau pelajaran seperti membaca dan aritmatika. Pendapat senada dinyatakan oleh Gronlund yang menyatakan bahwa tes diagnostik adalah tes yang dirancang untuk mengetahui sebab kegagalan siswa dalam mengajar. Tes diagnostik memiliki dua fungsi utama, yaitu, mengidentifikasi masalah atau kesulitan yang dialami siswa, dan merencanakan tindak lanjut berupa upaya-upaya pemecahan sesuai masalah atau kesulitan yang telah teridentifikasi.

Weeden, Winter & Brounfoot (2002, p.20) menyatakan bahwa tes diagnostik adalah tes yang digunakan untuk mengidentifikasi masalah-masalah spesifik yang dialami siswa. Tes diagnostik dapat menunjukkan dan memetakan kelemahan-kelemahan siswa dan ketidakajegan pengetahuan. Tes diagnostik seharusnya mampu mengungkap mengapa siswa memberikan respon seperti yang mereka lakukan (Gorin, 2007, p.174). Sesungguhnya tes diagnostik juga dapat membantu penilaian guru dalam memperdalam pemahaman konseptual siswa mereka dari topik tertentu (Russell, 2009, p.414). Dengan tes diagnostik, kelemahan siswa dapat dideteksi dan guru diharapkan dapat memberi bantuan yang tepat sesuai dengan apa yang dibutuhkan oleh siswa bila siswa mengalami kesulitan.

Kesulitan siswa dalam pembelajaran penting untuk dipecahkan. Dengan demikian, belajar dapat menghasilkan perubahan-perubahan dalam memperoleh kemajuan pe-ngetahuan pemahaman, ketrampilan dan nilai. Belajar menurut psikologi kognitif berarti belajar dengan memperhatikan tahapan-tahap proses berpikir yang mengacu pada psikologi kognitif. Anderson & Krathwohl (2001, p.39) menetapkan empat jenis pengetahuan yakni pengetahuan faktual, konseptual, prosedural dan metakognitif. Pengetahuan faktual meliputi elemen dasar yang harus diketahui siswa untuk mempelajari disiplin ilmu atau untuk menyelesaikan masalah-masalah dalam disiplin ilmu tersebut. Pengetahuan konseptual meliputi hubungan-hubungan antar elemen dalam sebuah struktur besar yang memungkinkan elemen-elemennya berfungsi secara bersama-sama. Pengetahuan prosedural meliputi bagaimana melakukan sesuatu, mempraktikkan metode-metode penelitian dan kriteria-kriteria untuk menggunakan ketrampilan, algoritma, teknik dan metode. Pengetahuan metakognitif meliputi pengetahuan tentang kognisi secara umum dan kesadaran dan pengetahuan tentang kognisi diri sendiri.

Pengembangan tes diagnostik harus memenuhi dua asumsi yaitu (1) tes mampu menganalisa keterampilan atau pengetahuan untuk dimasukkan ke dalam komponen sub-keterampilan atau sub-pengetahuan, (2) butir tes yang dikembangkan mampu mengukur sub-ketrampilan atau subpengetahuan tersebut (Mehrens & Lehmann, 1984, p.410). Dengan tes diagnostik, kemampuan siswa dapat dilihat pada level/tahapan pengetahuannya. Tahapan pengetahuan siswa dapat diketahui apabila *learning continuum* sudah dirumuskan. *Learning continuum* yaitu urutan logis yang menunjukkan keterkaitan materi secara vertikal dalam pelajaran. Pada kompetensi-kompetensi yang akan diukur diidentifikasi sejumlah kompetensi yang harus dikuasainya terlebih dahulu, kompetensi yang harus dikuasai adalah kemampuan pra-syarat.

Penelitian tentang tes diagnostik di Indonesia sudah mulai berkembang, Kusaeri (2012) mampu mengembangkan instrumen

tes diagnostik dengan model DINA menggunakan studi simulasi, membangun *syntak* dengan menggunakan CDM. Proses membangkitkan data dan analisis dalam penelitian ini menggunakan paket *software* Mplus dari Muthen & Muthen. Hasil studi simulasi menunjukkan bentuk matriks ideal. Suatu matriks dikatakan ideal bila matriks tersebut menghasilkan struktur *latent class* yang mak-simum. Penelitian Duskri, Kumaidi & Suryanto (2014, p.44) mengembangkan tes diagnostik kesulitan belajar matematika di SD. Penelitian ini dapat memunculkan informasi hasil tes secara klasikal dan individual, grafik ketuntasan belajar, profil individual, analisis salah konsepsi dan saran remedial. Namun, keterbatasan penelitian ini adalah *software* komputer yang dirancang untuk melakukan analisis hasil tes diagnostik kesulitan belajar masih berbasis *Microsoft excel* 2007 dan dapat dikembangkan dengan program lain yang lebih interaktif. Penelitian Isgiyanto (2011) menggunakan atribut dalam penyusunan instrumen, namun atribut yang tersusun tidak diidentifikasi *latent class* nya.

Tes diagnostik yang dikembangkan dalam penelitian ini adalah tes diagnostik kognitif pada mata pelajaran IPA SMP materi kalor. Pengembangan tes berdasar pada *Learning Continuum*, dan menggunakan atribut dalam penyusunan matriks Q dan diidentifikasi *latent class* nya. Kegunaan *Learning Continuum* menurut *Northwest Evaluation Association* (2003, p.4) adalah sebagai panduan untuk mengetahui perbedaan tingkat prestasi siswa. Pendidik harus menyadari bahwa perbedaan prestasi dan tingkat pencapaian siswa saat ini harus menjadi pertimbangan dominan ketika pengelompokan dalam memberikan instruksi dan memilih bahan/materi. *Learning Continuum* akan membantu guru mengetahui kelemahan siswa yang disebabkan karena siswa belum menguasai materi prasyarat dan akan mengembangkan potensi siswa apabila siswa telah menguasai kompetensi yang ditetapkan. Atribut didefinisikan sebagai suatu prosedur proses atau kemampuan siswa untuk menyelesaikan suatu *item* (Gierl, 2007,

p.327). Atribut adalah karakteristik laten dari peserta tes/responden (Rupp, Templin & Henson, 2010, p.315).

Tes diagnostik memiliki karakteristik yang berbeda dengan butir soal tes yang lain. Pada tes diagnostik jawaban atau respons yang diberikan oleh siswa harus memberikan informasi yang cukup untuk menduga masalah atau kesulitan yang dialaminya (memiliki fungsi diagnosis). Tes diagnostik dirancang dengan memperhatikan kognitif siswa. Tes diagnosis kognitif yang dikembangkan mengacu pada proses kognitif Bloom yang dikembangkan Anderson & Krathwohl (2001, p.39) dengan menggunakan tahapan vertikal dan tahapan horizontal. Tahapan vertikal yaitu dengan memperhatikan dimensi pengetahuan yang meliputi pengetahuan faktual, pengetahuan konseptual dan pengetahuan prosedural. Pengetahuan metakognitif tidak dikembangkan dalam penelitian ini dikarenakan merupakan pengetahuan diri. Dalam memecahkan masalah sains beberapa pengetahuan dapat berjalan bersama-sama pada proses kognitif yang berbeda. Proses kognitif merupakan tahapan secara horizontal, dalam tahapan ini terdiri dari: mengingat, memahami, mengaplikasikan, dan menganalisis.

Tujuan penelitian ini adalah (1) menghasilkan tes diagnostik kognitif pada mata pelajaran IPA materi kalor berdasarkan *learning continuum*; (2) menemukan karakteristik tes diagnostik kognitif pada mata pelajaran IPA materi kalor, berdasarkan *learning continuum*.

## Metode Penelitian

Penelitian ini menggunakan metode penelitian pengembangan tes diagnostik yang bertujuan untuk penilaian kognitif menurut Nichols (1994, p.587) yaitu: (1) mengkaji konstruksi teori yang substansif sebagai hal yang mendasar dalam pengembangan tes berdasar penelitian atau *review* penelitian; (2) melakukan seleksi desain pengukuran untuk mengkonstruksi butir yang dapat direspon oleh peserta tes berdasarkan pengetahuan, ketrampilan yang spesifik atau karakteristik lain sesuai teori; (3) melakukan administrasi

tes yang meliputi beberapa aspek diantaranya menyusun bentuk tes, teknologi yang digunakan untuk membuat tes, analisis situasi lingkungan pada waktu pengetesan, membuat kisi-kisi tes, menulis tes dan sebagainya; (4) membuat skoring hasil tes yaitu penentuan nilai tes yang telah dilakukan melalui proses analisis butir soal; (5) melakukan revisi yaitu proses penyesuaian dan perbaikan sesuai dengan tujuan tes diagnostik dan standar yang diharapkan.

Tes diagnostik kognitif dikembangkan sesuai dengan fokus penelitian melalui tiga tahapan yaitu (1) tahap perencanaan tes, (2) tahap uji coba, dan (3) tahapan pengembangan *prototype software computer based testing* agar penelitian ini dapat langsung diaplikasikan di sekolah. Penelitian pengembangan secara umum, prosesnya mencakup: membangun kerangka konseptual, diikuti dengan merancang sistem, dan berakhir dengan membangun prototipe untuk pengujian dan evaluasi (Ellis & Levy, 2010). Tahap perencanaan dimulai dengan studi literatur (kaji pustaka dan hasil penelitian terdahulu) dan studi lapangan. Studi literatur dilakukan untuk memahami lebih jauh tentang tes diagnostik kognitif. Diperlukan teori-teori baru yang menjelaskan penyebab yang mendasari masalah (Ellis & Levy, 2010, p.108). Pada studi literatur ini peneliti mengumpulkan temuan riset dan informasi yang mendukung terhadap produk yang direncanakan. Penyusunan instrumen tes diagnostik kognitif pada mata pelajaran IPA Fisika dimulai dengan kegiatan *workshop* yang diikuti oleh 6 orang guru yang tergabung dalam MGMP IPA kabupaten Sleman, pengawas dan pakar pengukuran. Pada tahap ini peneliti dan peserta *workshop* merumuskan *learning continuum* penguasaan konsep materi IPA SMP. *Learning continuum* yang dibuat diharapkan dapat menggambarkan perkembangan dan peningkatan kemampuan siswa untuk menguasai konsep IPA Fisika mulai dari konsep dasar menuju konsep lanjut. *Learning continuum* dapat digunakan untuk mengidentifikasi kompetensi dasar dan merumuskan indikator materi IPA SMP yaitu materi Kalor.

Tes yang dibuat dalam penelitian ini adalah tes diagnostik kognitif yang bertujuan untuk mengetahui kelemahan-kelemahan siswa dalam mata pelajaran IPA fisika materi kalor. Hasil tes diagnostik kognitif dapat digunakan sebagai dasar untuk memberikan tindak lanjut berupa perlakuan yang tepat dan sesuai dengan kelemahan yang dimiliki siswa. Agar tujuan tes dapat tercapai maka pada tahapan ini juga disusun atribut tes diagnostik kognitif pada materi kalor. Atribut adalah proses kognitif atau ketrampilan dasar yang dibutuhkan untuk memecahkan tes. Atribut dapat dilihat sebagai sumber dari kompleksitas kognitif pada hasil tes. Penyusunan atribut penting dilakukan sebagai landasan untuk menggambarkan kemampuan kognitif peserta tes. Atribut pada instrumen tes diagnostik kognitif ini dibuat berdasarkan pada *learning continuum* dan hierarki materi prasyarat. Pemilihan atribut dilakukan secara hati-hati karena berimplikasi pada proses mengkonstruksi Q matriks.

Bentuk tes yang digunakan dalam penelitian ini adalah bentuk tes objektif dengan beralasan. Dengan dilengkapi alasan, akan diketahui cara berfikir siswa dan langkah-langkah siswa dalam menyelesaikan soal. Siswa yang benar-benar menguasai akan dapat memberikan alasan jawabannya. Kelebihan tes obyektif adalah sebagai berikut: (a) jumlah materi yang dapat ditanyakan relatif lebih luas dibandingkan dengan materi yang dapat dicakup soal bentuk lainnya. Jumlah soal yang ditanyakan umumnya relatif banyak; (b) dapat mengukur berbagai jenjang; (c) penskorannya mudah, cepat, objektif, dan dapat mencakup ruang lingkup bahan dan materi yang luas dalam satu tes untuk suatu kelas atau jenjang; (d) sangat tepat digunakan apabila pesertanya banyak; (e) reliabilitas soal pilihan ganda relatif lebih tinggi dibandingkan dengan soal uraian. Namun tes obyektif juga mempunyai kelemahan salah satunya yaitu hanya dapat mengukur sampai dengan tahap analisis. Oleh karena itu dalam penelitian ini proses kognitif yang diukur sampai pada tahap analisis.

Uji coba penelitian pengembangan tes diagnostik kognitif pada mata pelajaran IPA

materi kalor dilaksanakan pada 484 siswa di Tujuh SMP yang mempunyai karakteristik berbeda di Kabupaten Sleman yaitu SMPN 1 Sleman, SMPN 1 Kalasan, SMPN 1 Turi, SMPN 4 Tempel, SMPN 4 Gamping, SMPN 2 Cangkringan dan SMP Muhammadiyah Pakem. Sampel uji coba harus mempunyai karakteristik yang kurang lebih sama dengan karakteristik peserta tes yang sesungguhnya. Ujicoba ini di gunakan sebagai sarana untuk memperoleh data empirik tentang tingkat kebaikan soal yang telah disusun. Melalui ujicoba data diperoleh tentang: realibilitas, validitas, tingkat kesukuran, pola jawaban, efektivitas pengecoh, daya beda dan lain-lain (Mardapi, 2008, p.95).

Analisis data pengembangan instrumen tes, pada tahap pertama menggunakan validasi isi produk oleh para ahli (*experts*) melalui proses FGD untuk mengetahui sejauh mana naskah tes disagnostik kognitif mata pelajaran IPA SMP mencerminkan tujuan pengembangan. Analisis data secara kuantitatif dengan bantuan program *Quest*, CDM dan menggunakan Program *M Plus*. Program *Quest* digunakan untuk mendapatkan parameter butir soal yaitu nilai tingkat kesulitan butir untuk melengkapi informasi diagnostik kognitif pada instrumen yang telah disusun yaitu dengan mengkaitkan tahap *learning continuum* dan dimensi kognitif. Program *Quest* dapat membandingkan *ability* dan *difficulty* yang tidak dapat diketahui apabila dianalisis menggunakan analisis secara klasik. Dengan demikian, hasil analisis program *Quest* dapat membantu memberikan informasi diagnostik tentang kelemahan dan kesulitan siswa. Program *Quest* juga dapat digunakan untuk menentukan *fit item* dengan model. Dengan demikian program *Quest* dapat membuat *item* yang terstandar. Batas suatu *item* dinyatakan *fit* dengan model jika memiliki *infit* MNSQ diantara 0,77 sampai dengan 1,30 (Adam & Khoo, 1996, p.30). Persyaratan *Fit* statistik pada program *Quest*, yakni jika nilai rata-rata *infit* MNSQ mendekati 1,0 dengan simpangan baku 0,00.

Program CDM dalam penelitian ini digunakan untuk membangkitkan *syntak* M Plus. Program M Plus digunakan untuk me-

nemukan informasi yang lebih banyak tentang *latent class*, dengan cara menyimulasikan ukuran matriks Q dan banyaknya peserta tes. Penelitian Simulasi menggunakan studi simulasi *Monte Carlo (M Plus)*. Pengklasifikasian profil atribut dilakukan untuk melihat kecenderungan pola jawaban peserta tes diagnostik kognitif sehingga dapat diketahui atribut-atribut yang telah dikuasai oleh kebanyakan peserta tes dan atribut-atribut yang belum dikuasai oleh peserta tes. Hal ini disampaikan oleh Rupp, Templin & Henson (2010, p.232) bahwa mengestimasi parameter peserta tes sama halnya dengan mengklasifikasi profil atribut peserta tes.

### Hasil Penelitian dan Pembahasan

Model Instrumen diagnostik kognitif yang dikembangkan berupa (1) *Learning Continuum* pada materi kalor mata pelajaran IPA SMP yang menjadi landasan dalam penyusunan kisi-kisi tes, (2) hierarki materi prasyarat mata pelajaran IPA SMP pada materi kalor, (3) *Item Specification* yang dikembangkan berisi tentang uraian yang menunjukkan karakteristik diagnostik kognitif yang harus dimiliki setiap butir tes, (4) atribut soal dan Q matrik, (5) perangkat tes diagnostik kognitif Mata Pelajaran IPA SMP yang terdiri atas 28 soal pilihan ganda dengan alasan dan (6) *software* yang dilengkapi panduan Tes Diagnostik Kognitif pada Mata Pelajaran IPA SMP khususnya pada materi kalor. *Software* komputer yang dikembangkan merupakan program aplikasi berbasis *website* yang menggunakan pemograman HTML dipadu dengan PHP dan menggunakan *database* jenis *MySQL*.

Secara keseluruhan instrumen hasil validasi para pakar mempunyai nilai baik dan sangat baik. Data yang diperoleh dari hasil FGD kemudian disusun dalam Q matriks dan atribut tes diagnostik kognitif mata Pelajaran IPA SMP. Atribut bersifat hierarkis, hal ini menggambarkan kemampuan prasyarat secara langsung antar atribut (Gierl, 2007, p.328). Pada penelitian ini atribut yang disusun berjumlah 17 dan disusun dalam 7 Q matrik.

Data yang sudah terkumpul di validasi secara empirik menggunakan *Rasch model* melalui program *Quest*. Seleksi butir dengan program *Quest* digunakan untuk mendapatkan nilai tingkat kesulitan butir dan fit tes. Analisis dengan *Rasch Model* digunakan untuk mengukur satu konstruks teori pada satu waktu pada hierarki skala logit. Batas suatu *item* dinyatakan fit dengan model jika memiliki infit MNSQ diantara 0,77 sampai dengan 1,30 (Adam & Khoo, 1996, p.30). Persyaratan Fit statistik pada program *Quest*, yakni jika nilai rata-rata infit MNSQ mendekati 1,0 dengan simpangan baku 0,00. Instrumen tes diagnostik kognitif pada mata pelajaran IPA memiliki jumlah butir soal yang dianalisis sebanyak 28 butir.

Pada penelitian ini 27 *item* berada di dalam garis vertikal atau memiliki infit MNSQ diantara 0,77 sampai dengan 1,30. Butir nomor 7 berada pada infit MNSQ 1,31. Namun karena butir soal nomor 7 tidak berada jauh dari garis vertikal, dan mempunyai selisih infit MNSQ 0,01 maka pada penelitian ini butir 7 tetap digunakan dengan direvisi. Instrumen penelitian ini telah memenuhi persyaratan Fit statistik pada program *Quest*, yakni dengan nilai rata-rata infit MNSQ 1,00 dengan simpangan baku 0,11. Nilai reliabilitas sampel pada penelitian ini sebesar 0,99 Menurut Wright & Masters (1982, p.106) Nilai reliabilitas berdasarkan estimasi *item* disebut reliabilitas sampel. Dengan demikian, semakin tinggi nilainya, semakin banyak *item* yang fit dengan model. Sedangkan reliabilitas tes dalam penelitian ini diperoleh sebesar 0,72 yaitu dengan melihat nilai reliabilitas berdasarkan estimasi *case* atau testi. Nilai reliabilitas berdasarkan estimasi *case/teste* disebut reliabilitas tes (Wright & Masters, 1982). Reliabilitas tes sebesar 0,72 memberikan informasi bahwa pengukuran menggunakan instrumen tes diagnostik kognitif pada pelajaran IPA memberikan hasil yang konsisten. Gronlund (1982) menyatakan besarnya indeks reliabilitas suatu alat ukur khususnya pada pengukuran pendidikan seperti tes buatan guru sekitar 0,6-0,85. Semakin tinggi nilai reliabilitas tes, semakin banyak sampel untuk uji coba yang memberikan informasi

yang diharapkan, begitu juga sebaliknya. Semakin rendah nilai reliabilitas tes maka semakin sedikit sampel untuk ujicoba yang memberikan informasi yang diharapkan.

Uji kecocokan fit kesesuaian model atribut dengan Q matriks dengan menggunakan program M Plus. Templin (2009) mengelompokan uji fit model dalam tiga kelompok yaitu; Fit mutlak, fit relatif dan fit *item*. Uji Fit perlu dilakukan sebelum menginterpretasikan dan melakukan desiminasi hasil (Rupp, Templin & Henson, 2010, p.265). Uji fit model atribut dengan Q matriks pada penelitian ini menggunakan nilai *entropy*. Nilai *entropy* 1 berarti seluruh peserta tes diklasifikasikan secara tepat (*Good fit*) ke dalam kelasnya. Nilai *entropy* 0 berarti semua peserta tes memiliki peluang sama diklasifikasikan ke dalam semua kelas (*poor fit*). Idealnya dalam mendesain Q matriks harus berkorespondensi dengan struktur *latent class* (de la Torre & Karelitz, 2009, p.233). Nilai *entropy* pada Q matriks yang telah tersusun dalam penelitian ini terendah adalah pada Q matriks 7 sebesar 0,472 hal ini berarti sebesar 47,2,6 % peserta tes diklasifikasikan secara tepat (*good fit*) ke dalam kelasnya. Nilai *entropy* terbesar terdapat pada Q matriks 6 yaitu sebesar 0,812 berarti sebanyak 81,2% peserta tes diklasifikasikan secara tepat (*good fit*) ke dalam kelasnya.

$$Q6 = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} \quad Q7 = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix}$$

Untuk menentukan Fit relatif menggunakan *information criteria*. Informasi ini digunakan untuk membandingkan dua Q matriks yang memiliki atribut yang berbeda. Dengan melihat nilai *Akaike Information Criteria* (AIC), *Bayesian Information Criteria* (BIC) dan *sample size adjusted BIC*. Nilai paling rendah dari *information criteria* menunjukkan paling baik memberikan struktur *latent class*. Menurut Nylund, Asparouhov, & Muthen (2007) penggunaan nilai BIC lebih tepat jika dibandingkan dengan AIC, hal ini dikarenakan hasil AIC seringkali lebih tinggi diban-

dingkan dengan nilai sebenarnya dan BIC lebih konsisten.

Pada Tabel 1 didapat bahwa Q matriks 6 mempunyai nilai *AIC*, *BIC* dan *Sample-Size Adjusted BIC* lebih rendah jika dibandingkan dengan Q matriks yang lain. Artinya ukuran matriks 3x3 lebih ideal dibandingkan matriks lain karena mampu memberikan informasi struktur *laten class* secara maksimum. Penelitian yang dilakukan oleh Kusaeri (2012, p.127) menyebutkan bahwa nilai paling rendah pada ketiga *information criteria*, menunjukkan bahwa Q matriks itu paling ideal karena mampu memberikan informasi struktur *latent class* secara maksimum. Pada penelitian ini Q matriks 6 dan Q matriks 7 adalah matriks dengan ukuran 3x3. Namun mempunyai susunan yang berbeda. Ukuran matriks 3x3 bukan merupakan satu-satunya yang menentukan baik tidaknya sebuah model namun juga dipengaruhi oleh pola dari matriks tersebut. Seperti yang disampaikan De la Torre, & Karelitz (2009, p.233) bahwa dengan memperhatikan aspek kesesuaian, maka dalam mendesain Q matrik harus berkorespondensi dengan struktur *latent class*.

Tabel 1. *Information Criteria* pada Q matrik

Matrik	<i>Information Criteria</i>		
	<i>Akaike (AIC)</i>	<i>Bayesian (BIC)</i>	<i>Sample-Size Adjusted BIC</i>
Q1	2387,752	2450,484	2402,875
Q2	2415,299	2478,031	2430,422
Q3	2910,025	3014,577	2935,229
Q4	1414,467	1477,199	1429,590
Q5	2710,149	2814,701	2735,353
Q6	1356,752	1411,119	1369,858
Q7	1951,507	2005,874	1964,613

Pada penelitian ini Q1 terdiri atas 3 atribut dan 4 *item*, Q2 terdiri atas 3 atribut dan 4 *item*, Q3 terdiri atas 4 atribut dan 5 *item*, Q4 terdiri atas 3 atribut dan 4 *item*, Q5 terdiri atas 4 atribut dan 5 *item*, Q6 terdiri atas 3 atribut dan 3 *item* dan pada Q7 terdiri atas 3 atribut dan 3 *item*. De La Torre & Karelitz (2009) menyarankan agar banyak-

nya *item* di dalam Q matriks lebih dari banyaknya atribut. Rupp & Templin (2008, p.83) mengungkapkan bahwa setiap penambahan sejumlah atribut pada Q matriks seharusnya panjang tes juga ditambah agar dapat memberikan informasi yang reliabel tentang semua variabel laten.

Idealnya Q matriks dijadikan sebagai dasar dalam menyusun kisi-kisi selama proses penyusunan *item*. Seperti penelitian yang telah dilakukan Kusaeri (2012, p. 150) bahwa *item* tes yang tidak dibangun berdasarkan Q matrik, kurang ada kesesuaian antara Q matriks dengan struktur *latent class* sehingga akan mempengaruhi rendahnya nilai *entropy*. Pada penelitian ini Q matriks dikonstruksi berdasarkan hierarki materi prasyarat dan *learning continuum* sehingga terdapat Q matriks dengan *entropy* rendah.

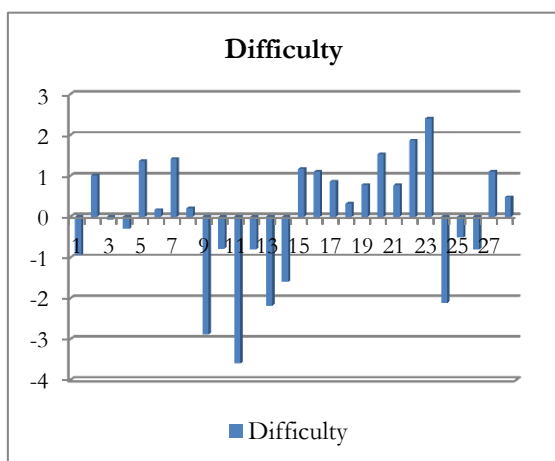
Untuk mengetahui Fit *item* dengan menggunakan program M Plus yaitu didasarkan pada dua informasi, yakni *univariate information* dan *bivariate information*. Sebuah *item* dikatakan fit apabila informasi yang diperoleh dari *univariate information*, data hasil observasi dengan data yang diprediksi model bernilai sama. Sementara *bivariate information* menghasilkan statistik fit dari pasangan *item*. Berdasarkan hasil analisis M Plus diketahui bahwa semua informasi yang diperoleh dari *univariate information*, data hasil observasi dengan data yang diprediksi model bernilai sama.

Hasil tes diagnostik kognitif yang telah diujicobakan di tujuh SMP di Kabupaten Sleman, diperoleh informasi bahwa pada soal jenis pengetahuan faktual memiliki rata-rata >70% pada seluruh sekolah yang menjadi subjek penelitian. Pada sekolah berkategori baik rata-rata siswa yang dapat mengerjakan soal jenis pengetahuan faktual >90% namun pada pengetahuan koseptual <65% dan pada pengetahuan prosedural <50%. Sekolah dengan kategori sedang, rata-rata siswa yang dapat mengerjakan soal pada jenis pengetahuan faktual >80%, namun pada pengetahuan koseptual <40% dan pada pengetahuan prosedural <30%. Pada sekolah dengan kategori kurang, rata-rata siswa yang dapat mengerjakan soal pada

jenis pengetahuan faktual sebesar >75% namun pada pengetahuan koseptual <40% dan pada pengetahuan prosedural <30%.

Pada jenis pengetahuan faktual diperoleh informasi diagnostik yaitu adanya siswa yang masih belum menguasai kemampuan prasyarat kalor (suhu dan zat), dan tidak dapat membedakan antara menyublim dan mengkristal. Pada jenis pengetahuan konseptual diperoleh informasi diagnostik bahwa masih banyak siswa yang belum dapat menjelaskan konsep perpindahan kalor, menganalisis perubahan suhu dan wujud zat menggunakan gambar, dan belum dapat menganalisis penerapan manfaat perpindahan kalor. Pada jenis dimensi prosedural diperoleh informasi diagnostik bahwa masih banyak siswa yang belum dapat menentukan jumlah kalor pada zat yang diketahui kalor jenisnya, menentukan kapasitas kalor pada zat yang diketahui massa dan kalor jenisnya, mengukur kenaikan suhu pada zat yang dipanaskan dengan kalor tertentu, menganalisis proses perubahan wujud zat menggunakan grafik dan menganalisis hukum kekekalan energi pada persamaan kalor (Azas Black).

Tingkat kesukaran butir soal dapat menambah informasi diagnostik. Dari 28 soal pilihan ganda dengan alasan yang dikerjakan oleh siswa diperoleh tingkat kesukaran butir yang disajikan pada Gambar 1.



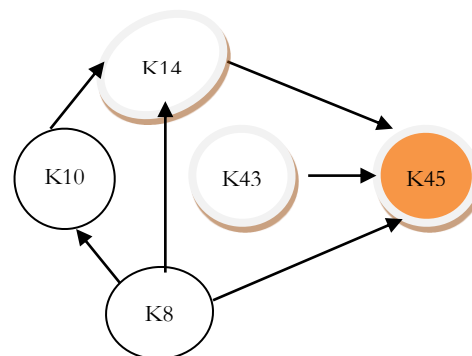
Gambar 1. *Item Difficulty* Tes Diagnostik

Pada Gambar 1 dapat dilihat bahwa soal yang paling sulit adalah *item* soal nomor 23 dan *item* soal paling mudah adalah *item*

soal nomor 11. Apabila dianalisis menggunakan Q matrik, maka *item* soal nomor 23 berada pada Q matriks 5. Atribut yang menyusun Q matriks 5 adalah atribut A3, A8 dan A12. Atribut A3 yaitu satuan dan besaran-besaran kalor, Atribut A8 yaitu tentang kalor laten dan A12 dan atribut A12 yaitu tentang Azas Black. Matriks Q 5 berukuran 5x4. Besarnya kelas laten pada matriks Q5 adalah  $2^4 = 16$ .

$$Q_5 = \begin{array}{c|cccc} & A_3 & A_5 & A_8 & A_{12} \\ \hline I_{17} & 1 & 0 & 0 & 0 \\ I_{20} & 1 & 1 & 1 & 0 \\ I_{21} & 1 & 1 & 1 & 0 \\ I_{22} & 1 & 0 & 0 & 1 \\ I_{23} & 1 & 0 & 1 & 1 \end{array}$$

Untuk dapat menjawab soal nomor 23 maka siswa harus menguasai materi prasyarat seperti yang ditampilkan pada Gambar 2.



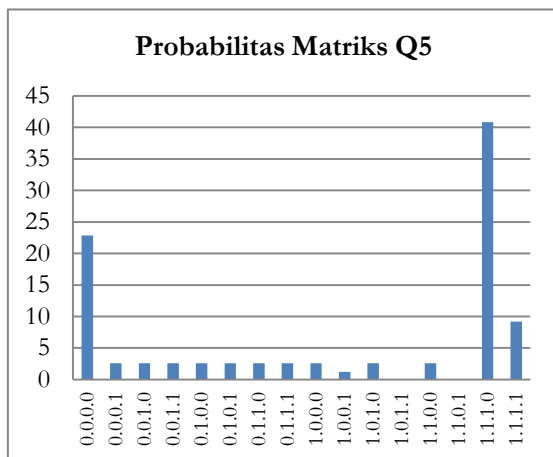
Gambar 2. Hierarki Materi prasyarat *item* soal nomor 23

Keterangan:

- K8 : Menjelaskan proses perubahan suhu pada zat
- K10: Menjelaskan konsep perpindahan kalor
- K14: Menentukan konversi satuan pada besaran kalor
- K43: Menjelaskan hukum kekekalan energi pada persamaan kalor (Azas Black)
- K45: Menentukan suhu campuran pada zat yang berbeda.

Kesulitan yang dialami siswa terutama dalam menentukan konversi dan menghitung besaran-besaran kalor dan kalor la-

ten. Selain itu, siswa juga belum paham terhadap konsep Azas Black. Hasil diagnostik ini digunakan oleh guru untuk menentukan tindakan sesuai dengan yang dibutuhkan siswa. *Item* pada matriks Q5 adalah *item* dengan jenis pengetahuan prosedural yang mempunyai tingkat kesukaran tinggi. Untuk dapat menyelesaikan soal dalam bentuk prosedural, siswa harus memahami bagaimana menyelesaikan soal dengan langkah-langkah dan prosedur yang tepat, mengkonversi satuan, menggunakan rumus sesuai dengan konsep yang benar. Oleh karena itu, diperlukan pengetahuan prasyarat untuk dapat menyelesaikan soal ini, siswa tidak cukup hanya menghafal rumusnya, namun harus paham terhadap prosedur dan teknik penyelesaiannya.



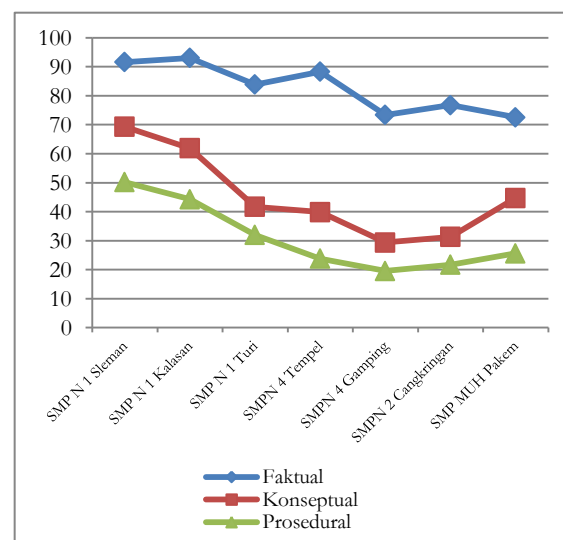
Gambar 3. Probabilitas Q Matriks 5

Guru harus membiasakan dan melatih siswa mengerjakan soal-soal dengan jenis pengetahuan prosedural. Untuk mendapatkan pengetahuan prosedural yang terkoneksi secara baik dengan pengetahuan konseptualnya tentu saja kepada siswa perlu ditanamkan konsep-konsep secara baik, dikaitkannya konsep tersebut dengan konsep yang lain, dan dilatihkan bagaimana merepresentasikan konsep dengan simbol, serta dilatihkan bagaimana menggunakan aturan atau prosedur untuk menyelesaikan permasalahan fisika.

Soal yang mempunyai *item difficulty* terendah adalah *item* soal nomor 11. Soal ini merupakan soal ingatan dan jenis soal faktual. Soal faktual cenderung mudah karena

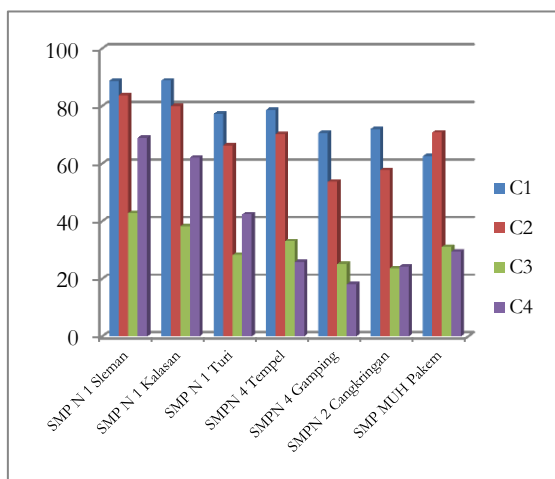
dalam kehidupan sehari-hari siswa sering menjumpainya bahkan mempraktikannya. Pada soal nomor 11 sebesar 97,7% siswa dapat menyelesaikan soal dan dapat menyampaikan alasannya dengan benar. Pada pengetahuan faktual cenderung lebih dikuasai siswa dari pada pengetahuan jenis konseptual dan prosedural. Pengetahuan faktual berkaitan dengan pengalaman yang dialami siswa dalam kehidupan sehari-hari. Apa yang dialami oleh siswa akan lebih mudah tersimpan dalam otak siswa. Oleh karena itu, dalam pembelajaran IPA, guru seharusnya senantiasa mengkaitkan materi pelajaran dengan peristiwa sehari-hari.

Pada penelitian ini soal dengan jenis pengetahuan faktual lebih banyak dikuasai oleh siswa dibandingkan dengan jenis pengetahuan prosedural. Siswa lebih menguasai soal yang berbentuk ingatan dan hafalan tanpa memahami konsepnya. Siswa cenderung hafal dengan rumus-rumus fisika dan tahapan-tahapan penyelesaiannya tanpa memahami konsep dan prosedur yang benar. Dengan demikian, apabila diberikan soal yang berbeda, siswa akan mengalami kesulitan. Hal ini sejalan dengan penelitian Oksuz (2007, p.13) menyatakan bahwa siswa sulit menghubungkan pengetahuan konseptual yang dimiliki ke dalam pengetahuan prosedural sehingga siswa cenderung menghafalkan tahap pengerjaan soal, tidak berdasarkan pemahaman terhadap konsep terkait.



Gambar 4. Hasil Tes Diagnostik Dilihat dari Jenis Pengetahuan

Hasil tes diagnostik kognitif yang telah diujicobakan di beberapa sekolah menunjukkan bahwa siswa pada setiap sekolah masih mengalami kesulitan pada tiap dimensi kognitif. Pada pelajaran IPA Fisika siswa disekolah cenderung hafal rumus namun tidak paham konsepnya. Sehingga pada soal-soal dengan jenis prosedural yang disusun dengan tipe yang berbeda maka siswa akan mengalami kesulitan. Dengan diketemukannya kesulitan siswa pada setiap sekolah berkaitan dengan pengetahuan konseptual dan prosedural, maka guru menjadi salah satu faktor yang mempunyai peranan, artinya kesalahan dan kesulitan yang dialami siswa kemungkinan dipengaruhi oleh cara guru dalam mengajar dan cara guru dalam memberikan evaluasi.



Gambar 5. Grafik Hasil Tes Diagnostik Berdasar Proses Pengetahuan

Grafik menunjukkan bahwa pada tiap sekolah masih mengalami kesulitan pada aspek aplikasi dan analisis. Hal ini seperti penelitian yang dilakukan Fauzan (2010, p.227) bahwa hampir semua siswa mengalami kesulitan dalam mengerjakan butir-butir soal yang termasuk analisis. Pada mata pelajaran IPA Fisika seharusnya soal aplikasi dan analisis sudah dibiasakan diberikan oleh guru sehingga siswa tidak mengalami kesulitan, namun pada penelitian ini aspek analisis dan aplikasi siswa masih sangat lemah. Dengan demikian, guru pada setiap sekolah perlu memperkaya dan melatih siswa dengan menggunakan soal-soal yang mampu

menjadikan siswa berpikir kritis dengan memperhatikan *learning continuum* agar dapat meningkatkan kualitas pembelajaran.

Pengembangan tes diagnostik kognitif akan lebih bermakna apabila dapat digunakan oleh guru. Pada penelitian ini guru dapat memanfaatkan hasil tes diagnostik kognitif sebagai dasar untuk memberikan tindak lanjut berupa perlakuan yang tepat dan sesuai dengan kelemahan yang dimiliki siswa melalui hasil pengembangan *software* komputer yang dilengkapi dengan *password*. Hasil tes diagnostik kognitif yang telah dikembangkan dengan *software* komputer terdiri atas dua informasi yaitu (1) banyaknya *item* yang dijawab dengan benar dan *item* yang dapat dijawab siswa berdasarkan jenis pengetahuannya, (2) informasi tentang materi yang sudah dikuasai dan belum dikuasai siswa berdasarkan *learning continuum* dan materi prasyarat yang belum diketahui siswa.

Penelitian ini dapat menjawab pertanyaan mengapa peserta tes tidak menguasai suatu materi sekaligus dapat menginterpretasikan hasil tes diagnostik kognitif pada mata pelajaran IPA SMP sehingga mampu mendiagnosis kesulitan belajar siswa. Kesulitan siswa dalam pembelajaran IPA Fisika materi kalor akan dapat dipecahkan, karena melalui *latent class* diketahui materi apa saja yang tidak dikuasai siswa. Guru dapat memberikan penanganan terutama letak kelemahan dan kesulitan belajar yang dialami siswa dan kemungkinan prestasi yang bisa dikembangkan. Hal ini sesuai dengan yang disampaikan Maddalena (2008) bahwa pengajar tidak hanya mampu mengidentifikasi dimana dan bagaimana siswa sering salah paham terhadap konsep yang sangat rumit, tetapi juga bahwa mereka dapat mengidentifikasi cara terbaik untuk memperbaikinya.

Tes diagnostik kognitif yang dikembangkan dalam penelitian ini dipadu *software* komputer yang dapat membantu guru untuk memberikan pembelajaran yang sesuai dengan proses berpikir siswa sehingga kualitas pembelajaran IPA meningkat. Penelitian ini sejalan dengan hasil penelitian Afaq & Al-Mashari (2010, p.85) yang telah mengembangkan alat diagnostik bertujuan untuk

membantu siswa dengan mendiagnosis kesulitan belajar. Alat yang dikembangkan memiliki kemampuan menguji kemampuan pelajar untuk menjawab pertanyaan dengan benar dengan mengevaluasi respon otomatis dan kemudian menghasilkan daftar kelemahan atau kesalahan yang ditemukan. Umpan balik yang diberikan dapat digunakan oleh instruktur/guru untuk meningkatkan keterampilan pelajar dengan berfokus pada kelemahan yang dilaporkan, dan menghasilkan tes diagnostik yang terstandar. Tes diagnostik kognitif yang tersusun dilengkapi dengan *software* komputer, sehingga tes diagnostik ini dapat langsung digunakan oleh guru untuk menghasilkan informasi yang dapat mendiagnosis kelemahan dan kekuatan siswa sehingga guru dapat memberikan tindak lanjut dan penanganan yang sesuai agar dapat meningkatkan kualitas pembelajaran.

### Simpulan dan Saran

Tahapan pengembangan instrumen diagnostik kognitif pada mata pelajaran IPA SMP materi kalor dimulai dengan perumusan *Learning Continuum*. Tahapan berikutnya adalah penyusunan Kisi-kisi tes, Hierarki Materi Prasyarat, Spesifikasi Tes, Soal dan Q matriks yang diturunkan dari Atribut. Karakteristik instrumen diagnostik kognitif yang tersusun telah memenuhi validitas isi melalui *expert judgement* dengan FGD.

Hasil pengembangan pada revisi akhir terdiri atas 28 soal yang berbentuk pilihan ganda dengan alasan. Instrumen diagnostik kognitif pada mata pelajaran IPA materi kalor telah memenuhi persyaratan yaitu *fit* dengan model, dengan menggunakan program *Quest*. Instrumen penelitian ini memiliki *mean infit MNSQ 1,01* dengan simpangan baku 0,09. Reliabilitas tes dengan rumus *alpha Cronbach* menunjukkan tes memiliki reliabilitas tinggi yaitu 0,72. Uji fit kesesuaian model atribut dengan Q matriks menggunakan program M Plus berhasil menyusun 7 buah Q matriks dengan *Latent class* pada setiap matriks adalah  $2^n$  dengan  $n$  adalah jumlah atribut. Penelitian ini berhasil menemukan besarnya propabilitas pada se-

tiap *latent class* pada tujuh Q matriks yang tersusun.

Saran yang diberikan penelitian ini untuk Guru IPA SMP yaitu agar dapat menerapkan produk Instrumen Tes Diagnostik Kognitif pada materi Kalor, sehingga guru dapat mendiagnosis kesulitan belajar siswa dan dapat memberikan bantuan sesuai dengan yang dibutuhkan siswa. Tindak lanjut yang dilakukan guru dalam pembelajaran hendaknya lebih menanamkan konsep fisika dan prosedur yang tepat dan memperhatikan materi prasyarat yang harus dikuasai siswa. Dengan demikian, kualitas pembelajaran dapat meningkat. Bagi peneliti yang tertarik untuk mengembangkan model tes diagnostik kognitif lebih lanjut, perlu ketelitian dan kecermatan dalam menyusun atribut yang harus selaras dengan *learning continuum* dan hierarki materi prasyarat.

Pemerintah khususnya Badan Standar Nasional Pendidikan (BSNP), diharapkan memperhatikan hierarki materi prasyarat dan *learning continuum* dalam menyusun standar kompetensi dan kompetensi dasar agar sesuai dengan struktur berpikir siswa. Dengan demikian, anak dapat belajar sesuai dengan tahap berpikirnya. Instrumen tes diagnostik kognitif akan lebih mudah dimanfaatkan jika dikembangkan dengan model aplikasi yang lengkap dan spesifik. Oleh karena itu, bagi yang berminat mengembangkan dapat melakukan penelitian lanjutan secara lebih mendalam dan komprehensif baik dalam bidang kajian IPA ataupun kajian pada bidang yang lain dapat mengembangkannya dengan model yang lebih lengkap dan spesifik.

### Daftar Pustaka

- Adam, R.J. & Khoo, S.T. (1996). *Quest. The Interactive Test Analysis System*. Version 2.1. Melbourne: ACER.
- Afaq, A & Al-Mashari, A. (2010). On the development of a computer based diagnostic assessment tool to help in teaching and learning process. *International Journal of Education and Development using Information and Communi-*

- cation Technology(IJEDICT). Vol 6. Issue 1.76-87.*
- Anderson, L. W., & Krathwohl, D. R. (Eds). (2001). A taxonomy for learning, teaching and assessing: *A revision of Bloom's Taxonomy of educational objectives*: Complete edition, New York: Longman.
- De La Torre, J. & Karelitz, T.M (2009). Impact of Diagnosticity on the Adequacy of Models for Cognitive Diagnosis under a Linear Attribute Structure: A Simulation Study. *Journal of Educational Measurement Winter*. 46 (4):450-469.
- Dunn, L., Morgan.C., Parry, S. R., & Reilly, M.O. (2004). *The Student assessment Handbook*. New York. Taylor & Francise. Library
- Duskri, M., Kumaidi, K., & Suryanto, S. (2014). Pengembangan Tes Diagnostik Kesulitan Belajar Mtematika di SD. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 18(1), 44-56. Retrieved from <http://journal.uny.ac.id/index.php/jp ep/article/view/2123/1768>.
- Ebel, Robert, L. (1979). *Essentials of Educational Measurement*. New Jersey: Prentice. Hall.Inc.
- Efendi, R. (2010). Kemampuan Fisika Siswa Indonesia Dalam TIMSS (Trend of International on Mathematic and Science Study). *Prosiding Seminar Nasional Fisika 2010*. ISBN: 978-979-98010-6-7.
- Ellis, T.J & Levy, Y. (2010). *A Guide for Novice Researchers: Design and Development Research Methods*. Proceedings of Informing Science & IT Education Conference (InSITE), 108-118.
- Fauzan. (2010). Pengembangan tes diagnostik kesulitan belajar matematika di SMA. *Disertasi doktor*, tidak diterbitkan, Universitas Negeri Yogyakarta, Yogyakarta.
- Gage, N.L & Berliner, D.C. (1991). *Educational Psychology*. 5<sup>th</sup> ed. Boston: Houghton Mifflin Company
- Gierl, M.J. (2007). Making Diagnostic Inferences About Cognitive Attributes Using the Rule Space Model and Attribute Hierarchy Method. *Journal of Educational Measurement*, 44(4), 325-340.
- Gorin, J.S. (2007). Test construction and diagnostic testing. Dalam Leighton, J.P. & Gierl, M.J.(eds). *Cognitive diagnostic assessment for education theory and applications*.(p.174). Cambridge. Cambridge University Press.
- Gronlund, Norman, E. (1982). *Measurement and Evaluation in Teaching*. New York: Macmilan Publishing.co.
- Isgiyanto, A. (2011). Diagnosis Kesalahan Siswa Berbasis Penskoran Politomus Model Partial Credit Pada Matematika. *Jurnal Penelitian dan Evaluasi Pendidikan*, 15 (2), 308-325.
- Kusaeri. (2012). Pengembangan tes diagnostik dengan menggunakan model DINA unuk mendapatkan informasi salah konsepsi dalam aljabar. *Disertasi doktor*, tidak diterbitkan, Universitas Negeri Yogyakarta, Yogyakarta.
- Leung, C., & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: *assessment as discourse and assessment of discourse*. <http://itj.sagepub.com>.
- Maddalena, T. (2008). Issues of power and equity in two models of self-assessment. *Teaching in Higher Education*. 13 (1) 81-92.
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan non tes*. Yogyakarta: Mitra Cendikia Offset.
- Mehrens, W.A. & Lehman, I.J. (1984). *Measurement and evaluation in education and Psychology*. New York. Holt, Rinehart and Winson, Inc.

- Mislevy, R.J. (1994). *Evidence and Inference in educational assessment*. *Psychometrica*, 59, 439-483.
- Nichols, P.D. (1994). A framework for developing cognitively diagnostic assessment. *Review of Educational*. 64 (4), 575-603.
- Nitko, J.A. (1996). *Of Educational Assessment of Student*. New Jersey: Prentice Hall. Inc.
- NWEA. (2003). *Idaho State Aligned Learning Continuum Release 1.0* diambil dari <https://id.scribd.com/doc/244834242/ID-State-Learning-Continuum>
- Nylund, K.L., Asparouhov, T. & Muthen, B.O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling A Monte Carlo simulation study. *Structural Equation Modeling*, 14, 535-569.
- Oksuz, C. (2007). Childrens understanding of equality and the equal symbol. Adnan Menderes. Diambil dari [cumalioksuz@adu.edu.tr](mailto:cumalioksuz@adu.edu.tr)
- Suparno, P. (1998). Miskonsepsi (Konsep alternatif) Siswa SMU dalam Bidang Fisika. Dalam Sumaji. *Pendidikan sains yang Humanis*. 95-111. Kanisius: Yogyakarta.
- Permendiknas Nomor 20 Tahun 2007 tentang Standar penilaian.
- Rupp, A. A., & Templin, J. L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement: Interdisciplinary Research and Perspectives*, 6(4), 219-262.
- Rupp, A.A., Templin, J. & Henson, R.A. (2010). *Diagnostic measurement: Theory, methods and applications*. New York: The Guilford Press.
- Russell, M. & O'dwyer. L.M. (2009). Diagnosing students' misconceptions in algebra: results from an experimental pilot. *Behavior Research Methods*. 41(2), 414-424.
- Taufik, M., Sukmadinata, Abdulkhak, I., Tumbelaka, B. Y. (2010). Desain model pembelajaran untuk meningkatkan kemampuan pemecahan masalah dalam pembelajaran IPA (fisika) sekolah menengah pertama di kota bandung. *Berkala Fisika, Vol 13, No 2 (2010): Edisi Kbusus, E31-E44*
- Templin, J. (2009). Estimation of diagnostic models with Mplus. Slide presentasi disampaikan dalam *National Council on Measurement in education, Diagnostic Modeling*, tanggal 12 Februari 2014. <http://jtemplin.coe.uga.edu/teaching>.
- Weeden, P. Winter, J. & Broadfoot, P. (2002). *Assessment what's in for school?*. Routledge falmer: London
- Wright, B.D & Masters, GN. (1982). *Rating Scale analysis*. Chicago: Mesa Press.