

## PENGEMBANGAN MODEL PENILAIAN AKHLAK PESERTA DIDIK MADRASAH ALIYAH

<sup>1)</sup>Septimar Prihatini, <sup>2)</sup>Djemari Mardapi, <sup>3)</sup>Sutrisno

<sup>1)</sup>Sekolah Tinggi Ilmu Tarbiyah Islamic Village Tangerang Banten

<sup>2)</sup>Universitas Negeri Yogyakarta, <sup>3)</sup>UIN Sunan Kalijaga Yogyakarta

<sup>1)</sup>septimar.uny@gmail.com, <sup>2)</sup>djemarimardapi@gmail.com, <sup>3)</sup>

### Abstrak

Tujuan penelitian ini adalah mengembangkan model penilaian akhlak peserta didik Madrasah Aliyah. Konstruk akhlak dikembangkan dengan pendekatan akhlak sebagai aspek afektif dan konsep *religiosity* dari Glock and Stark (1966). Penelitian pengembangan ini menggunakan metode lapor diri dan observasi tak langsung, melibatkan 291 siswa dan 26 guru dengan teknik *purposive sampling*. Pengujian kualitas instrumen meliputi: validasi tampang, isi, dan konstruk; reliabilitas (konsistensi internal dan reliabilitas interrater); dan uji stabilitas instrumen. Hasil penelitian ini adalah: (1) Model penilaian akhlak terdiri dari model pengukuran dan sistem penilaian. Instrumen tersusun melalui proses *expert judgment*, *Focus Group Discussion* dan pengujian konstruk. (2) Hasil analisis faktor konfirmatori (CFA) berdasarkan data empirik menunjukkan bahwa konstruk akhlak peserta didik mencakup dimensi akhlak kepada Allah, akhlak kepada Nabi Muhammad SAW., akhlak kepada orangtua, akhlak kepada diri sendiri, akhlak kepada guru, akhlak kepada teman/tetangga/masyarakat, dan akhlak kepada lingkungan. (3) Instrumen mempunyai kehandalan internal antara 0,865 – 0,921 (tinggi) dan reliabilitas interater 0,866 (tinggi) dan koefisien Cohens' Kappa 0,770 (sangat baik), serta stabilitas antara 0,715 sampai 0,858 (baik sampai sangat baik). (4) Sistem penilaian setelah disimulasikan dan dikonfirmasi menunjukkan 90 % kesesuaian dengan performansi akhlak siswa.

**Kata kunci:** *model penilaian, akhlak peserta didik, konstruk akhlak, instrumen*

## DEVELOPING A MODEL OF AN ASSESSMENT OF MADRASAH ALIYAH STUDENTS' AKHLAK

### Abstract

This study aims to develop a model of an assessment of Madrasah Aliyah students' akhlak. The constructs of students' akhlak are developed on the basis of the akhlak as affective domain and as a religiosity concept, adopting the ritualistic concept and consequences from Glock and Stark (1966). This research and development used self-report and indirect observations, involving 291 students and 26 teachers by using purposive sampling technique. The quality instrument testing consists of: validation process (face, content and construct), reliability (internal consistency and inter-rater reliability); and stability. The results of the study are as follows. (1) The model of an assessment of students' akhlak includes a measurement model and an assessment system for students' akhlak. The instrument is composed through the process of expert judgment, Focus Group Discussion and testing constructs (2) The result of Confirmatory Factor Analysis (CFA) shows that constructs of students' akhlak cover the dimensions of akhlak to Allah, Prophet Muhammad, parents, self, teachers, society, and environment. (3) The internal reliability of instrument ranges from 0.865 to 0.921 (high); the inter-rater reliability is 0.851 (high) by Product Moment Correlation and Cohens' Kappa coefficient 0.770 (very good). Instruments stability levels ranging from being good to being very good (0.715 to 0.858). (4) the assessment system model after the simulation is confirmation-yields an agreement of 90% with the respondents' performances.

**Keywords:** *assessment model, students' akhlak, akhlaks' constructs, instrument*

## Pendahuluan

Sekolah merupakan *moral community* yang berperan penting dalam pembinaan moral anak didik, di samping tempat pemberian pengetahuan, pengembangan bakat dan kecerdasan (Bruce R.Thomas in Goodlad, et.al, 1984, pp.265-266). Demikian pula madrasah, sebagai sekolah bercirikan Islam harus mampu berperan sebagai lapangan sosial bagi anak-anak, tempat pertumbuhan mental, moral dan sosial serta segala aspek kepribadian dapat berjalan dengan baik.

Untuk menilai keberhasilan pendidikan dapat ditinjau dari keberhasilan bidang kognitif, psikomotor dan afektif (Griffin 1991, p.56). Namun penilaian yang banyak dikembangkan selama ini lebih banyak penilaian kognitif, sementara penilaian afektif belum banyak dikembangkan (Levin, 2011, p.3). Salah satunya adalah penilaian akhlak peserta didik. Padahal jika memperhatikan tujuan pendidikan nasional yang tercantum dalam Undang-Undang RI NO. 20 tahun 2003 Bab I pasal 3 tentang sistem pendidikan nasional yaitu *berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri dan menjadi warga negara yang demokratis dan bertanggung jawab* (DepDiknas, 2005), maka penilaian akhlak penting dilakukan.

Begitu pula jika memperhatikan Standar Kompetensi Lulusan SMA/MA yang sebagian besar didominasi aspek non kognitif (Depdiknas, 2005) dan syarat kelulusan peserta didik dari sebuah satuan pendidikan menurut Peraturan Menteri Pendidikan No. 45/2010 tentang syarat kelulusan dari satuan pendidikan selain lulus ujian nasional, ujian sekolah, juga memiliki akhlak, budi pekerti, sikap dan perilaku minimum baik (BSNP, Standar Operasional Ujian Nasional, 2011). Dengan demikian maka penilaian akhlak peserta didik mutlak diperlukan sebagai salah satu syarat kelulusan dari satuan pendidikan, termasuk syarat kelulusan peserta didik dari Madrasah Aliyah.

Penilaian akhlak peserta didik yang selama ini umum dilakukan di Madrasah Aliyah adalah menggunakan instrumen tes untuk materi ajar Akidah/Akhlak. Hasil nilai tes pelajaran pelajaran Akidah/Akhlak aspek kognitif. Untuk penilaian aspek afektif digunakan pengamatan secara sekilas tentang kepribadian dan penampilan siswa secara umum tanpa menggunakan pedoman penilaian dan pengukuran sebagaimana bila mengukur aspek afektif.

Jika memperhatikan ketetapan Permendiknas nomor 20 tahun 2007 pasal D.8 tentang penilaian akhlak mulia yang menyebutkan bahwa penilaian akhlak merupakan aspek afektif dari kelompok mata pelajaran agama dan akhlak mulia sebagai perwujudan sikap dan perilaku beriman dan bertakwa kepada Tuhan Yang Maha Esa, dilakukan oleh guru agama dengan memanfaatkan informasi dari pendidikan mata pelajaran lain dan sumber lain yang relevan, maka seharusnya dalam menilai akhlak peserta didik selain dilakukan oleh guru akhlak juga dilakukan oleh guru lain yang memiliki kompetensi dan kewenangan menilai akhlak. Namun sampai saat ini belum dikembangkan instrumen yang dapat mengukur akhlak peserta didik di Madrasah Aliyah atau di Sekolah Menengah Atas lainnya yang secara empirik dapat dipertanggungjawabkan.

Salah satu kendala belum dikembangkannya instrumen penilaian akhlak adalah pengungkapan konstruk akhlak yang akan digunakan sebagai indikator dalam mengukur akhlak peserta didik. Untuk itulah penelitian ini dilakukan.

Dari uraian di atas, penelitian ini bertujuan (1) mengidentifikasi konstruk akhlak peserta didik Madrasah Aliyah; (2) menyusun instrumen pengukuran akhlak peserta didik Madrasah Aliyah; dan (3) menyusun model penilaian akhlak peserta didik Madrasah Aliyah.

Spesifikasi produk yang dikembangkan adalah pada materi atau butir yang disusun dalam instrumen, serta pendekatan konsep akhlak sebagai konsep *religiosity* dan akhlak sebagai aspek afektif. Di samping itu

juga pada tahapan langkah pengujian kualitas instrumen yang dilakukan yaitu proses *communication value*, *readability*, validasi, pengujian reliabilitas instrumen dan reliabilitas interrater, dan penggunaan skala instrumen.

Instrumen pengukuran akhlak peserta didik terdiri dari dua inventori yaitu guru dan siswa, serta empat set instrumen yaitu instrumen IA, IB, IIA, dan IIB. Instrumen IA berisi butir penilaian diri, instrumen IB berisi penilaian oleh teman sekelas, instrumen IIA penilaian oleh guru akhlak, dan instrumen IIB berisi penilaian oleh guru konseling dan walikelas tentang akhlak peserta didik terhadap Allah, nabi Muhammad SAW., orang tua, dirinya, gurunya, teman/tetangga/masyarakatnya, dan lingkungannya. Tampak bahwa peserta didik yang dinilai dilibatkan sebagai penilai melalui penilaian diri sendiri dan penilaian antar teman. Di samping itu, penilaian tidak hanya diberikan oleh guru akhlak tapi juga berdasarkan penilaian oleh wali kelas/guru bidang studi lain dan oleh guru konseling sehingga lebih dapat dipertanggungjawabkan.

Penentuan nilai akhir akhlak peserta didik untuk kepentingan kelulusan peserta didik dari Madrasah Aliyah berdasarkan jumlah persentase penilaian “baik” dan “kurang baik” hasil respon terhadap instrumen dengan menggunakan formulasi Lawshe (1975). Penentuan kategori “baik” dan “kurang baik” sebelumnya melalui proses penentuan skor ketuntasan minimum untuk setiap dimensi akhlak yang diukur.

Penelitian ini menghasilkan produk berupa instrumen terdiri dari dua inventori yaitu guru dan siswa. Inventori siswa terdiri dari dua set instrumen yaitu instrumen IA, IB, inventori guru terdiri dari instrumen IIA dan IIB. Instrumen IA berisi butir penilaian diri, instrumen IB berisi penilaian oleh teman sekelas, instrumen IIA penilaian oleh guru akhlak, dan instrumen IIB berisi penilaian oleh guru konseling dan wali kelas. Butir indikator dikembangkan melalui analisis kajian teori tentang akhlak dengan pendekatan konsep akhlak sebagai konsep *religiosity* dan akhlak sebagai aspek afektif.

## Kajian Teori

Akhlak merupakan salah satu konsep agama Islam, istilah yang digunakan sebagai pengganti kata akhlak dalam literatur asing adalah *Islamic ethic* atau *ethical Islam*, dan *morality* (Siddiqui, 1997, p.2; Abdalati, 1978, p.8; Swarup, 1983, pp.xv-xvi; Braswell, 1996, p.11). Untuk memahami akhlak, referensi utama adalah hadis, yang merupakan tindakan dan keputusan yang dilakukan oleh Nabi Muhammad. Tindakannya mulai dari bagaimana cara beliau tidur, makan, shalat, menikah, berkeluarga, keringanan keputusan hukum, merencanakan perjalanan sampai strategi melawan musuh dalam peperangan (Swarup, 1983, pp.xv-xvi; Brashel, 1996). Abdalati (1978, p.40) menyimpulkan bahwa dalam memahami akhlak seorang muslim harus melihat Muhammad SAW. sebagai contoh yang harus diikuti, terutama dalam penegakan lima pilar dalam Islam (rukun Islam) yaitu tata cara bersyahadat, tertib dalam gerakan shalat, hukum dan tata cara puasa, zakat dan ibadah haji.

*Islamic ethic* sering dikaitkan dengan istilah *morality*. Meskipun keduanya menurut *western vocabulary* berbeda asal katanya. *Morality* berasal dari bahasa Latin (*mores* atau *morals*), sedangkan *ethics* berasal dari bahasa Yunani (*ethos*). Keduanya mempunyai makna *habits* atau *customs* (kebiasaan). Menurut pemikiran Eropa, *morals* berarti *commonly felt and done* (kebiasaan yang dirasakan dan dilakukan). Sedangkan menurut pemikiran Islam, *ethics is appropriate and rationale* (memadai dan rasional). Secara khusus akhlak termasuk dalam *knowledge of morality* dan merupakan *science of ethics* atau ilmu tentang akhlak atau *ilm-ul-akhlak* (Siddiqui, 1997, p.423). Moral secara umum didefinisikan sebagai sebuah “ajaran” tentang baik buruk yang dapat diterima umum mengenai perbuatan, sikap, dan kewajiban (Cahn & Markie, 2006, p.569). Dapat disimpulkan bahwa akhlak secara umum sering dikaitkan dengan moral atau kebiasaan yang dirasakan dan dilakukan dan secara khusus akhlak juga berkaitan dengan pengetahuan tentang moral. Seorang berakhlak baik sering disebut memi-

liki perilaku moral (*moral behaviour*) yang baik, begitu pula seorang berakhlak buruk sering dikatakan memiliki perilaku moral (*moral behaviour*) buruk. Untuk itu diperlukan sebuah acuan (referensi) moral.

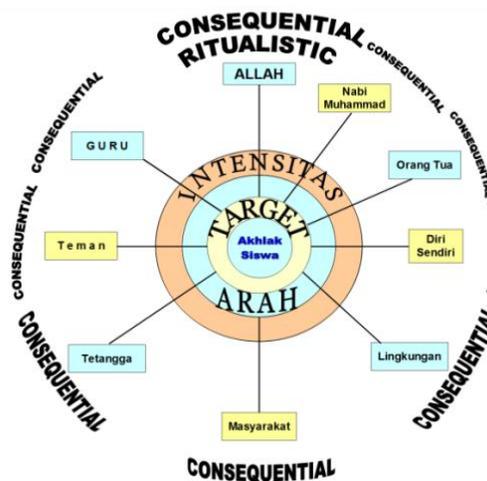
Menurut *Divine Command theories of morality* acuan moral diturunkan oleh Tuhan, seperti tercantum dalam Bible dan Qur'an (Waluchow, 2003, p.19). Teori *Divine Command* sejalan dengan pendapat George N. Braswell (1996) yang mengatakan bahwa pandangan Islam tentang *ethic* seperti pandangan Kristen yaitu *moral absolutes*. Dalam iman Kristen, Bible ("Injil") menjadi dasar dalam kaitan pemahaman terhadap karakter dasar Tuhan. Tuhan mempunyai hak untuk mengultimatum suatu perbuatan dikatakan baik atau tidak. Sehingga muslim mengetahui moral yang baik melalui perkataan Tuhan yang ada di dalam al-Qur'an (Braswell, 1996, p.11).

Dalam kaitan *Islamic ethic* sebagai konsep moral, Abdalati (1978:32) menyimpulkan bahwa konsep moral dalam Islam (*the concept of morality in Islam*) berkisar pada keyakinan dan prinsip-prinsip dasar yang sudah ditentukan. Konstruk akhlak peserta didik. Konstruk akhlak dikembangkan berdasarkan dimensi *ritualistic* dan *consequencial* dari Glock and Stark (Robinson & Shaver: 1969, pp.556-557). Dimensi *ritualistic* dan *consequencial* dikembangkan pada aspek/dimensi akhlak kepada Allah. Dimensi *consequencial* dikembangkan pada aspek/dimensi akhlak kepada ciptaan Allah.

Secara skematik penilaian akhlak peserta didik berdasarkan konsep *religiosity* dan karakteristik afektif disajikan pada Gambar 1.

Pengembangan instrumen penilaian akhlak peserta didik Madrasah Aliyah didasarkan pada karakteristik penilaian afektif yaitu *target, direction and intencity* (Fishbein, M, & Ajzen, I: 1975; Gable: 1986, p.56). Target atau sasaran penilaian akhlak diimplementasikan dengan akhlak kepada Allah dan akhlak kepada ciptaan Allah yang terdiri atas akhlak kepada Nabi Muhammad, diri sendiri, orang tua, guru, teman/tetangga/masyarakat, dan lingkungan. Arah (*direction*)

ditunjukkan oleh pernyataan akhlak baik (positif) dan merupakan perbuatan yang diperintahkan atau dianjurkan untuk dilakukan. Arah negatif ditunjukkan oleh pernyataan akhlak tidak baik atau tercela (negatif) yang merupakan perbuatan yang tidak boleh atau dilarang untuk dilakukan. Intensitas (*intencity*) dinyatakan dengan frekuensi perilaku akhlak dilakukan.



Gambar 1. Bagan Penilaian Akhlak Peserta Didik sebagai Konsep *Religiosity* dan Karakteristik Ranah Afektif

Berkaitan dengan intensitas, pengembangan instrumen pengukuran akhlak peserta didik Madrasah Aliyah memperhatikan daya ingat siswa tentang waktu atau kapan perbuatan tersebut telah dilakukan. Berdasarkan masukan dari Prof. Dough Willms (Supervisor di University of New Brunswick, Canada), teman sejawat, hasil *Focus Group Discussion (FGD)* dan saran para ahli, maka pada instrumen *self report* dijelaskan tentang rentang waktu penilaian. Variasi waktu diberikan tergantung pada intensitas akhlak yang dilakukan.

Dengan memperhatikan sistematika akhlak yang dikembangkan oleh Anshari (1986, p.31) serta pendapat Miskawaih (1968, p.30), Al-Qardhawi (1981, pp.106-109), Shihab (1996, p.261), tentang definisi akhlak dapat disimpulkan bahwa akhlak dapat dikategorikan ke dalam dua dimensi yaitu akhlak kepada Allah (sebagai *kehalig* atau pencipta) dan akhlak kepada makhluk

Allah (yang diciptakan Allah). Akhlak kepada Allah mencakup aspek kewajiban manusia sebagai makhluk ciptaan-Nya yaitu mengabdikan kepada-Nya, melakukan perbuatan yang diperintahkan-Nya serta menjauhi yang dilarang-Nya. Akhlak kepada Allah dibatasi pada keluasan aspek ibadah kepada Allah sebagaimana pendapat Majid (2003, p.23), Rahim (2001, p.39) dan Yani (2009, p.34) serta disesuaikan dengan kondisi usia peserta didik sehingga akhlak kepada Allah dibatasi pada aspek shalat, puasa, dzikir dan tidak syirik.

Untuk dimensi akhlak kepada makhluk Allah, terdiri dari aspek akhlak kepada Nabi Muhammad SAW., akhlak kepada diri sendiri, akhlak kepada orang tua, akhlak kepada guru, akhlak kepada teman sejawat, akhlak kepada kakak/adik, akhlak kepada masyarakat/tetangga dan akhlak kepada lingkungan. Pengembangan indikator untuk tiap aspek dari dimensi akhlak tersebut didasarkan pada referensi ayat al-Qur'an, hadis Nabi Muhammad SAW., pendapat ulama atau ahli agama, penulis buku akhlak dan *ijtihad*. *Ijtihad* berdasarkan hasil pengamatan dan pengalaman. Dalam menetapkan indikator yang digunakan dalam pengembangan model penilaian akhlak peserta didik didasarkan pada asumsi bahwa indikator tersebut potensial dapat diukur, (Robinson dan Shaver, 1966, p.552)

Proses penilaian akhlak peserta didik dilakukan dengan mempertimbangkan aspek-aspek penilaian afektif di antaranya adalah hasil dalam bentuk kualitatif (huruf). Konversi nilai kuantitatif ke kualitatif melalui proses penentuan Kriteria Ketuntasan Minimum yang diperoleh melalui *expert judgment*. Kriteria "baik" untuk akhlak peserta didik adalah syarat kelulusan peserta didik dari Madrasah Aliyah. Penentuan kriteria ketuntasan "baik" menggunakan pendekatan Lawshe (1975, p.5) pada proses *quantifying consensus*. Sebagai penilai (rater) terdiri atas peserta didik yang menilai diri sendirinya, guru akhlak, guru konseling, wali kelas, dan teman sekelasnya.

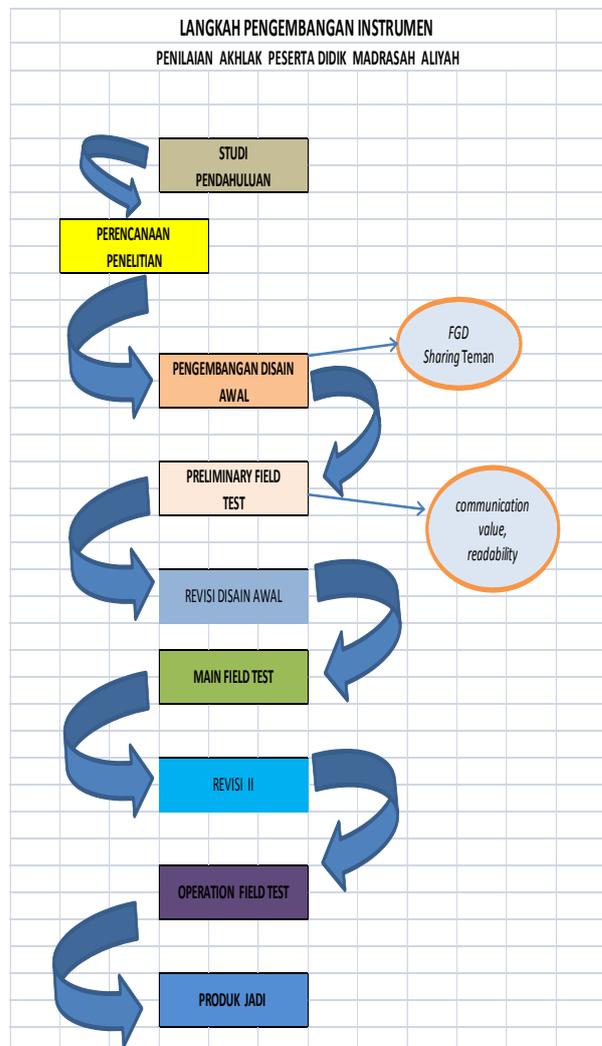
Penelitian ini juga menghasilkan sistem penilaian dan penentuan nilai akhir akhlak

peserta didik. Untuk kepentingan kelulusan peserta didik dari Madrasah Aliyah berdasarkan jumlah persentase penilaian "baik" dan "kurang baik" hasil respon terhadap instrumen dengan menggunakan formulasi Lawshe (1975).

## Metode Penelitian

### Jenis Penelitian

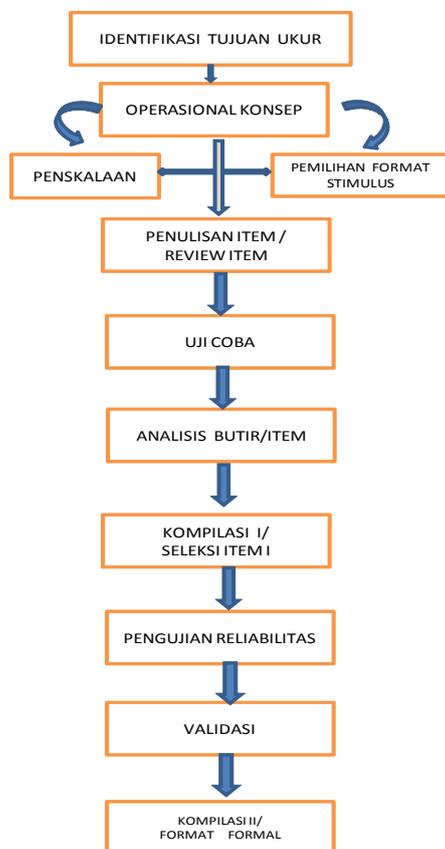
Penelitian ini termasuk kategori *Research and Development* (R&D) dengan memodifikasi model yang direkomendasikan oleh Borg dan Gall (2003: 274). Alur pengembangan model instrumen penilaian akhlak peserta didik secara skematik ditampilkan pada Gambar 2.



Gambar 2. Model Pengembangan Instrumen Penilaian Akhlak

### Langkah Pengembangan Instrumen

Pengembangan instrumen mengadaptasi langkah pengembangan instrumen dari Azwar (2000, p.11) yang nampak pada Gambar 3.



Gambar 3. Langkah Pengembangan Instrumen

### Metode Pengambilan Data

Pengukuran akhlak peserta didik Madrasah Aliyah ini menggunakan dua metode yang disarankan untuk digunakan dalam penilaian afektif yaitu *self report* (lapor diri) dan observasi (Anderson, 1981, p.90). Metode lapor diri digunakan dengan menggunakan instrumen jenis angket dengan teknik *paper and pencil test*. Metode observasi yang digunakan adalah observasi secara tidak langsung (*indirect observation*), yaitu menggunakan penilai yang lebih memahami responden yaitu guru akhlak, wali kelas serta guru bimbingan konseling dan penilaian oleh teman sekelas. Hasil penilaian guru akhlak, wali kelas dan guru konseling ter-

sebut digunakan sebagai bahan penilaian pembandingan dalam tahapan penilaian akhir. Responden penelitian ini adalah siswa kelas XII, sehingga penentuan nilai akhir yang digunakan adalah pendekatan Lawshe (1975).

### Pengujian Kualitas Instrumen

Penelitian ini menggunakan dua inventori (siswa dan guru) dengan empat instrumen. Kualitas instrumen diuji melalui proses validasi, reliabilitas dan stabilitas. Proses validasi meliputi validasi tampak dan isi (*face validity* dan *content validity*) melalui kegiatan *communication value*, *readability*, *expert judgment* dan *Focus Group Discussion (FGD)*. Validasi empirik dilakukan berdasarkan data yang diperoleh dengan menggunakan analisis butir dan Confirmatory Factor Analysis (CFA). Untuk mengetahui tingkat reliabilitas, dilakukan analisis internal reliabilitas (untuk *instrumen self report*) dan reliabilitas interater (untuk instrumen observasi tak langsung). Untuk mengetahui tingkat stabilitas instrumen digunakan pendekatan tes re-tes dengan waktu jeda 3 minggu.

### Hasil Penelitian dan Pembahasan

#### Pengembangan Disain Awal

Hasil analisis kajian teori menghasilkan disain instrumen yang diawali dengan menyusun tabel spesifikasi dan kisi-kisi. Dengan mengacu pada kisi-kisi, disusun *draft* instrumen pengukuran akhlak peserta didik Madrasah Aliyah. Disain awal terdiri atas 2 dimensi akhlak dan 8 sub dimensi yaitu (a) Akhlak kepada Allah dengan 16 indikator, (b) Akhlak kepada makhluk Allah dengan subdimensi (1) akhlak kepada nabi Muhammad SAW. dengan 3 indikator, (2) akhlak kepada diri sendiri dengan 11 indikator, (3) akhlak kepada orang tua dengan 14 indikator, (4) akhlak kepada guru dengan 6 indikator, (5) akhlak kepada teman dengan 11 indikator, (6) akhlak kepada teman/tetangga dengan 5 indikator, (7) akhlak kepada masyarakat dengan 5 indikator, (8) akhlak kepada lingkungan dengan 4

indikator. Disain awal ini akan digunakan sebagai bahan kajian *preliminary field study* dan *focus group discussion* serta penilaian teman sejawat.

*Preliminary Field Test*

*Preliminary field test* bertujuan untuk proses *communication value, readability* instrumen kepada calon responden yaitu peserta didik dan guru. Pada tahapan ini digunakan siswa dan guru MAS Nurul Falah sebagai responden. MAS Nurul Falah terletak di perbatasan kota dan kabupaten Tangerang. Secara geografis karakteristik siswa MAS tersebut dapat mewakili karakteristik siswa wilayah kota dan kabupaten.

*Penilaian Teman Sejawat*

Tujuan kegiatan penilaian teman sejawat adalah untuk memperoleh masukan tentang draft instrumen 1A, 1B, dan instrumen 2A/B sebagai bagian dari *face validity*. Teman sejawat terdiri atas 2 orang pengawas PAI, 2 orang dosen PAI, 4 orang guru PAI, 2 rekan mahasiswa S3 UNY, jadi seluruhnya berjumlah 10 orang. Hasil masukan teman sejawat tersebut adalah sebagai berikut: a. Format instrumen diperbaiki agar lebih efisien; b. Petunjuk pengisian dipersingkat; c. Penggunaan beberapa kata yang tidak tepat diganti, misalnya: kata “Anda” pada petunjuk instrumen diganti “Saudara”

Hasil penilaian teman sejawat dan penilaian para ahli terhadap instrumen IA, IB, IIA dan IIB ditampilkan pada Tabel 1.

Tabel 1. Hasil Penilaian Teman Sejawat dan Ahli tentang Instrumen

Hasil Penilaian Instrumen					
No	Instumen	Sangat baik	Baik	cukup	kurang
1	IA	-	60%	40%	-
2	IB	-	80%	10%	10%
3	IIA	-	80%	20%	-
4	IIB	-	70%	30%	-

Tabel 1 menunjukkan dari 10 orang, 6 orang (60%) menilai instrumen IA baik, 4 orang (40 %) menilai cukup; 8 orang (80%)

menilai instrumen I B baik dan 1 orang (10%) menilai instrumen I B cukup dan 1 orang (10%) menilai kurang. Untuk instrumen II A 8 orang (80%) menilai baik dan 2 orang (20%) menilai cukup; sedangkan untuk instrumen II B 7 orang (70 %) menilai instrumen baik dan sisa nya 3 orang (30%) menilai cukup.

*Focus Group Discussion (FGD)*

Tujuan kegiatan *focus group discussion* adalah memperoleh masukan terhadap *draft* instrumen yang telah disusun, kesinambungan kisi-kisi dengan butir pernyataan angket serta beberapa pertanyaan berkaitan dengan pengembangan instrumen yang memerlukan jawaban ahli. Ahli yang dilibatkan yakni ahli pengukuran, ahli pendidikan karakter, ahli bahasa serta pengguna instrumen dari MAN 1 Yogyakarta.

Hasil rangkuman dan analisis *preliminary field test*, Penilaian teman sejawat, *focus group discussion* menghasilkan revisi instrumen 1A, 1B, II A dan II B

Main Field Test

*Estimasi Koefisien Reliabilitas Instrumen I A*

Analisis internal konsistensi instrumen IA selain secara total juga dilakukan secara parsial berdasarkan kategori jurusan IPA, IPS, dan Agama, juga berdasarkan pada ketegori jenis kelamin (putra dan putri). Hasil rangkuman analisis reliabilitas internal instrumen I A hasil respon siswa di MAN Balaraja ditampilkan dalam Tabel 2.

Tabel 2. Hasil Analisis Reliabilitas Internal Data *Main Field Test* (MAN Balaraja)

No	Kategori	Jumlah Siswa	Koefisien ALPHA
1	Total (IPA/IPS/AGAMA)	50	0,984
2	Siswa putra	14	0,910
3	Siswa putri	36	0,887
4	Siswa IPA	20	0,915
5	Siswa IPS	20	0,865
6	Siswa Agama	10	0,906

Hasil Analisis Butir Instrumen IA

Dengan bantuan program SPSS versi 16 dan teknik “*item-total correlated*” diperoleh hasil analisis butir instrumen IA. Berdasarkan hasil tersebut disusun kembali kisi-kisi instrumen dengan urutan nomor disesuaikan kembali. Ada beberapa item yang menurut analisis butir mempunyai nilai “*item-total correlated*” kurang dari 0,3, namun dengan pertimbangan item tersebut merupakan bagian dari item yang potensial dapat diukur dan tidak bernilai negatif, maka peneliti masih memasukkannya ke dalam bagian dari instrumen revisi. Jumlah keseluruhan item adalah 84 termasuk 2 item berkarakteristik gender khusus putri yaitu pernyataan mengenai keterlaksanaan ‘*qadha*’ Ramadhan dan keterlaksanaan menjaga aurat (jilbab).

Operational Field Test

Hasil Analisis Diskriptif

Tabel 3 menunjukkan analisis diskriptif respon siswa terhadap instrumen 1A. Tampak bahwa rerata skor respon siswa dari empat Madrasah Aliyah Negeri dan Swasta terhadap butir instrumen berkisar dari 2.403 sampai 2.838. Jika rerata skor standar (minimum – maksimum) berkisar antara 0 – 3.952 maka skor rerata respon termasuk dalam kategori baik. Hasil rerata keseluruhan skala instrumen berkisar antara 146.600 sampai 174.710. Jika total skor berdasarkan skala berkisar antara 0 – 332 maka rerata skor total yang diperoleh termasuk kategori cukup baik. Varians butir berkisar antara 0,127 sampai 0,332 dan varians skala instrumen antara 295.983 sampai 978.067, sedangkan standard deviasi berkisar antara 17.204 – 31.274. Berdasarkan hasil analisis tersebut dapat disimpulkan bahwa instrumen IA direspon secara variatif oleh peserta didik dari empat Madrasah Aliyah Negeri dan Swasta tersebut.

Tabel 3. Analisis Diskriptif Respon Peserta Didik terhadap Instrumen IA (pada *operational field test*)

No	Nama Madrasah	Kategori	N Siswa	N butir	Rerata Butir	Rerata Skala Statistik	Varians Butir	Varians Skala	Standar Deviasi		
1	MAN Kota Tangerang	Jurusan IPA	38	84	2,218	174,71	0,231	316,644	17,794		
		Jurusan IPS	30		2,718	166,5	0,161	695,983	17,204		
		Jurusan Agama	15		2,472	153,267	0,309	833,924	28,878		
		Total (IPA-IPS-Agama)		83		2,719	168,690	0,177	450,245	21,218	
		Putra (IPA-IPS-Agama)		40	82	2,667	164,75	0,149	523,967	22,891	
		Putri (IPA-IPS-Agama)		43	84	2,777	172,162	0,262	365,378	19,114	
2	MAN Tiga Raksa	Jurusan IPA	32	82	2,526	156,613	0,134	416,445	20,407		
		Jurusan IPS	26		2,838	170,110	0,230	501,089	22,729		
		Total (IPA-IPS-Agama)			58		2,648	164,175	0,141	518,112	22,762
		Putra (IPA-IPS-Agama)			21	82	2,607	161,636	0,211	517,385	22,746
		Putri (IPA-IPS-Agama)			37	84	2,718	165,771	0,151	527,005	22,965
3	MAS Al Ijtihad Kota Tangerang	Total (IPS)	29	82	2,578	159,862	0,173	687,480	26,219		
		Putra (IPS)		10	82	2,403	146,600	0,127	455,156	21,334	
		Putri (IPS)		19	84	2,691	166,842	0,199	692,696	26,319	
4	MAS Al Mubaroq	Jurusan IPS	35	82	2,684	166,400	0,218	636,482	25,228		
		Putra (IPS)		16	82	2,734	166,750	0,338	978,067	31,274	
		Putri (IPS)		19	84	2,625	162,737	0,274	356,205	18,873	

Hasil Analisis Reliabilitas Internal Instrumen IA

Hasil analisis reliabilitas internal instrumen 1A ditampilkan pada tabel 4. Tabel

4 memperlihatkan hasil analisis internal reliabilitas instrumen IA yaitu berkisar antara 0,813 sampai 0,927. Standar mini-

num skala Prince tentang *internal consistency reliability for difference values inventory* untuk 80– 85 butir adalah minimal 0,75.

Tabel 4 memperlihatkan hasil analisis internal reliabilitas instrumen IA yaitu ber-

kisar antara 0,813 sampai 0,927. Standar minimum skala Prince tentang *internal consistency reliability for difference values inventory* untuk 80 – 85 butir adalah minimal 0,75.

Tabel 4. Relibilitas Internal Instrumen 1A berdasarkan Kategori Jurusan dan Jenis Kelamin.

No	Nama Madrasah	Kategori	N Siswa	N butir	Koefisien Alpha
1	MAN Kota Tangerang	Jurusan IPA	38	83	0,848
		Jurusan IPS	30		0,813
		Jurusan Agama	15		0,921
		Total (IPA-IPS-Agama)	83	0,874	
		Putra (IPA-IPS-Agama)	40	83	0,892
		Putri (IPA-IPS-Agama)	43	84	0,850
2	MAN Tiga Raksa	Jurusan IPA	32	83	0,850
		Jurusan IPS	26		0,900
		Total (IPA-IPS-Agama)	58		0,885
		Putra (IPA-IPS-Agama)	21	83	0,878
		Putri (IPA-IPS-Agama)	37	84	0,895
		Total (IPS)	29	83	0,911
3	MAS Al Ijtihad Kota Tangerang	Putra (IPS)	10	83	0,840
		Putri (IPS)	19	84	0,927
		Total IPS	35	83	0,887
4	MAS Al Mubaroq	Putra (IPS)	16	83	0,920
		Putri (IPS)	19	84	0,824
		Kategori	Rerata Koefisien Alpha		
		IPA	64	0,874	
		IPS	126	0,865	
		Agama	15	0,921	
		Putra	87	0,883	
		Putri	118	0,874	

Tabel 4 juga menunjukkan analisis reliabilitas internal instrumen IA berdasarkan kategori jurusan yaitu untuk responden siswa jurusan IPA menghasilkan koefisien alpha 0,874, untuk reponden jurusan IPS sebesar 0,865, reponden jurusan agama 0,921. Hasil analisis reliabilitas berdasarkan kategori jenis kelamin diperoleh koefisien alpha untuk responden siswa putra sebesar 0,883, dan responden siswa putri 0,874. Dengan demikian, hasil analisis internal reliabilitas untuk instrumen IA berdasarkan kategori jurusan IPA/IPS, kategori jenis kelamin keseluruhannya sangat baik.

Dengan demikian, dapat disimpulkan bahwa instrumen IA reliabel dan dapat di-

gunakan untuk mengukur akhlak peserta didik secara variatif berdasarkan kategori jurusan IPA, IPS dan Agama serta jenis kelamin putra dan putri di Madrasah Aliyah di wilayah kota dan kabupaten.

*Hasil analisis reliabilitas interater instrumen I B, II A dan II B*

Penilaian rater dianalisis dengan menggunakan teknik korelasi antara skor hasil penilaian antar rater per butir per siswa dengan menggunakan program Microsoft Excell 2007. Tabel 5 menunjukkan rangkuman hasil penilaian interrater tiap dimensi akhlak.

Tabel 5. Rangkuman Hasil Analisis Korelasi Interrater Per Dimensi Akhlak

No	Dimensi	Instrumen IB dan II A Teman (siswa) dan guru akhlak	Intrumen II B Wali kelas/guru non- akhlak dan guru konseling
1	Semua dimensi	0,837	-
2	Akhlak kepada Allah	0,774	-
3	Akhlak kepada Nabi Muhammad	0,885	-
4	Akhlak kepada orang tua	0,883	0,788
5	Akhlak kepada guru	0,922	0,950
6	a.Akhlak kepada diri sendiri gender	0,833	0,844
	b.Akhlak kepada diri sendiri nongender	0,891	0,820
7	Akhlak kepada masyarakat	0,871	0,880
8	Akhlak kepada lingkungan	0,734	0,892
	Rerata dengan butir gender	0,843	0,862
	Rerata dengan butir nongender	0,851	0,866

Penentuan reliabilitas interrater juga menggunakan formulasi koefisien Cohen's Kappa. Setelah data interval yang diperoleh dikategorikan menjadi data nominal, data dianalisis dengan bantuan software SPSS 16 dengan menggunakan prosedur *cross-tabs*. Interpretasi Koefisien Kappa menggunakan acuan sebagaimana tabel *Benchmarks for Interpreting Kappa* dari Landis dan Koch (1977) yaitu yang ditampilkan dalam tabel berikut,

Kappa Statistic	Strength of Agreement (Kekuatan konsesus)
<0,00	<i>Poor</i> (sangat rendah)
0,00 – 0,20	<i>Slight</i> (rendah)
0,21 – 0,40	<i>Fair</i> (cukup)
0,41 – 0,60	<i>Moderate</i> (baik)
0,61 – 0,80	<i>Substantial</i> (sangat baik)
0,81 – 1,00	<i>Almost Perfect</i> (hampir sempurna)

Hasil penghitungan Cohens' Kappa diperoleh koefisien Kappa antara rater guru akhlak dengan siswa (antarteman) diperoleh 0,770, sedangkan antara rater wali kelas dengan guru konseling diperoleh 0,713. Berdasarkan acuan interpretasi dari Landis dan Koch (2004) dapat disimpulkan bahwa tingkat konsensus antar rater sangat baik.

#### Hasil Analisis Stabilitas Instrumen.

Untuk menganalisis stabilitas instrumen IA, IB, IIA, dan IIB digunakan teknik korelasi Product Moment. Pendekatan yang dilakukan adalah mengkorelasikan skor *test – re test*. Responden yang digunakan adalah siswa MAN Kota Tangerang. Hasil respon pada penelitian I dikorelasikan dengan skor respon pada penelitian II. *Delay time* nya 3 minggu. Hasil analisis korelasi dengan menggunakan program Microsoft Excel 2007 diperoleh untuk instrumen IA tiap kategori siswa (IPA-IPS-Agama) 0,749; instrumen IB 0,827; instrumen II A 0,858.; instrumen IIB rater guru wali kelas 0,767; instrumen IIB rater guru konseling 0,715. Berdasarkan referensi *Values Inventory by Coopersmith* (1967) maka dapat disimpulkan stabilitas instrumen I A baik, stabilitas instrumen I B sangat baik, stabilitas instrumen IIA sangat baik, stabilitas instrumen II B baik.

#### Analisis Diskriptif Hasil Operational Field Test.

#### Distribusi Skor Hasil Respon terhadap Instrumen IA.

Hasil analisis distribusi kategori skor menunjukkan data cenderung condong ke arah skor cukup dan tinggi. Berarti *skewness*

positif. Hal ini menunjukkan data tidak berdistribusi normal.

Tabel 6. Distribusi Kategori Skor Perolehan Data Siswa (Instrumen IA)

Kategori	Rentang Skor	Frekuensi	Persentase
Missing	0	0	0
Sangat rendah	$0 \leq \text{skor} < 99,6$	0	0
Rendah	$99,6 \leq \text{skor} < 132,8$	18	8,53
Cukup	$132,8 \leq \text{skor} < 199,2$	114	54,03
Tinggi	$199,2 \leq \text{skor} < 265,6$	78	36,97
Sangat tinggi	$265,6 \leq \text{skor} < 332$	1	0,47
Total		211	100

*Distribusi Kategori Skor Perolehan Data Siswa (instrumen IB)*

Hasil analisis distribusi kategori skor menunjukkan bahwa data cenderung condong ke arah skor tinggi yang berarti *skewness* positif. Berdasarkan hal tersebut dapat disimpulkan data tidak terdistribusi normal

Tabel 7. Distribusi Kategori Skor Perolehan Data Siswa (Instrumen IB)

Kategori	Rentang Skor	Frekuensi	Persentase
Missing	-	4	1,9
Sangat rendah	$0 \leq \text{skor} < 44$	3	1,42
Rendah	$44 \leq \text{skor} < 80,67$	13	6,16
Cukup	$80,67 \leq \text{skor} < 132$	41	19,43
Tinggi	$132 \leq \text{skor} < 176$	135	63,98
Sangat tinggi	$176 \leq \text{skor} < 220$	15	7,11
Total		211	100

*Distribusi Kategori Skor Perolehan Data Guru Akhlak (instrumen IIA)*

Hasil analisis distribusi kategori skor menunjukkan bahwa data cenderung condong ke arah skor tinggi, berarti *skewness* positif. Hal ini menunjukkan bahwa data tidak terdistribusi normal.

Tabel 8. Distribusi Kategori Skor Perolehan Data Guru Akhlak (Instrumen IIA)

Kategori	Rentang Skor	Frekuensi	Persentase
Missing	-	7	3,32
Sangat rendah	$0 \leq \text{skor} < 44$	0	0
Rendah	$44 \leq \text{skor} < 80,67$	1	0,47
Cukup	$80,67 \leq \text{skor} < 132$	26	12,32
Tinggi	$132 \leq \text{skor} < 176$	170	80,57
Sangat tinggi	$176 \leq \text{skor} < 220$	7	3,32
Total		211	100

*Distribusi Kategori Skor Perolehan Data Wali kelas (instrumen II B)*

Hasil analisis distribusi kategori skor menunjukkan data cenderung condong ke arah skor tinggi, berarti *skewness* positif. Hal ini menunjukkan bahwa data tidak terdistribusi normal.

Tabel 9. Distribusi Kategori Skor Perolehan Data Wali Kelas/Guru Lain (Instrumen IIB)

Kategori	Rentang Skor	Frekuensi	Persentase
Missing	-	7	3,32
Sangat rendah	$0 \leq \text{skor} < 28$	1	0,48
Rendah	$28 \leq \text{skor} < 56$	15	7,14
Cukup	$56 \leq \text{skor} < 84$	34	16,19
Tinggi	$84 \leq \text{skor} < 112$	125	59,52
Sangat tinggi	$112 \leq \text{skor} < 140$	29	13,81
Total		211	100

*Distribusi Kategori Skor Perolehan Data Guru Konseling (instrumen II B)*

Hasil analisis distribusi kategori skor menunjukkan data cenderung condong ke arah skor tinggi, berarti *skewness* positif. Hal ini menunjukkan secara umum data tidak terdistribusi normal.

Tabel 10. Distribusi Kategori Skor Perolehan Data Konseling(Instrumen IIB)

Kategori	Rentang Skor	Frekuensi	Persentase
Missing	-	3	1,42
Sangat rendah	$35 \leq \text{skor} < 56$	0	0
Rendah	$56 \leq \text{skor} < 77$	2	0,95
Cukup	$77 \leq \text{skor} < 98$	29	13,74
Tinggi	$98 \leq \text{skor} < 119$	149	70,62
Sangat tinggi	$119 \leq \text{skor} < 140$	28	13,27
Total		211	100

Pengujian Model Pengukuran

Hasil Pengujian Model Konstruksi Akhlak Peserta Didik.

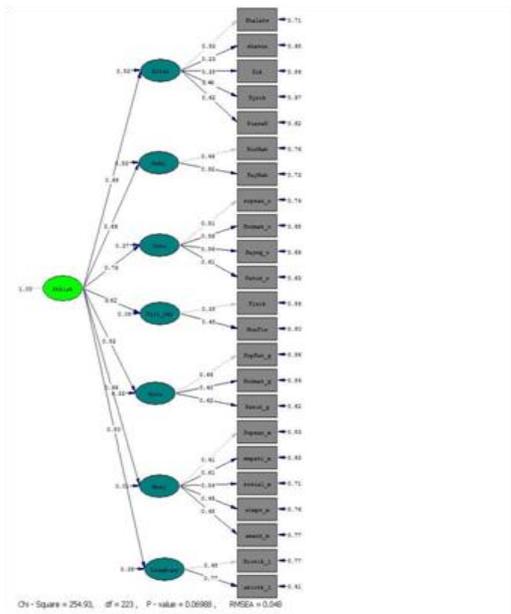
Persiapan analisis menemukan bahwa data secara multivariat tidak normal (Lampiran *Out put Lisrell*), karena itu metode estimasi *Robust Maximum Likelihood (RML)* digunakan dalam penelitian ini. Hasil pengujian model konstruk akhlak peserta didik Madrasah Aliyah (data siswa dalam skor baku) dengan butir gender disertakan dan tidak disertakan dapat dilihat pada Tabel 11.

Tabel 11. Rangkuman Hasil Pengujian Model Pengukuran

No Konstruk	$\chi^2$	df	p-value	RMSEA	Kesimpulan	Muatan Faktor Tertinggi	Muatan Faktor Terendah
1 Akhlak Peserta didik							
a.Nongender Butir	254,9	223	0,0699	0,048	Model fit	Akhlak kepada masyarakat (0,99)	Akhlak kepada diri sendiri (0,62)
b.Gender Butir	245,1	223	0,1477	0,0356	Model fit	Akhlak kepada masyarakat (0,99)	Akhlak kepada diri sendiri (0,62)
2 Akhlak kepada Allah							
a.Nongender Butir	8,86	6	0,1815	0,048	Model fit	Shalat wajib (0,996)	Syirik (0,052)
b.Gender Butir	3,96	5	0,5555	0,000	Model fit	Shalat wajib (0,996)	Syirik (0,052)
3 Akhlak kepada Nabi Muhammad	0	0	1,0000	0,000	Model fit	Hormat kepada nabi (0,628)	Sayang kepada nabi (0,399)
4 Akhlak kepada orang tua	0,47	2	0,0000	0,000	Model fit	Patuh kepada orang tua (0,811)	Sayang kepada orang tua (0,390)
5 Akhlak kepada diri sendiri							
a.Nongender Butir	12,78	4	0,1240	0,002	Model fit	Menjaga kehalalan makan/minum (0,831)	Menjaga kesehatan ruhani (0,603)
b.Gender Butir	15,04	7	0,0855	0,074	Model fit	Menjaga kehalalan makan/minum (0,831)	Menjaga kesehatan ruhani (0,605)
6 Akhlak kepada guru	91,99	17	0,0800	0,045	Model fit	Aspek hormat (memberi salam) (0,990/0,791)	Aspek sopan (membungkuk hormat) (0,621)
7 Akhlak kepada teman/tetangga	7,32	5	0,1977	0,000	Model fit	Empati (0,606)	Simpati (0,325)
8 Akhlak kepada lingkungan	17,00	5	0,0645	0,007	Model fit	Aspek abiotik (udara) (0,640)	Aspek biotik (cinta hewan) (0,517)

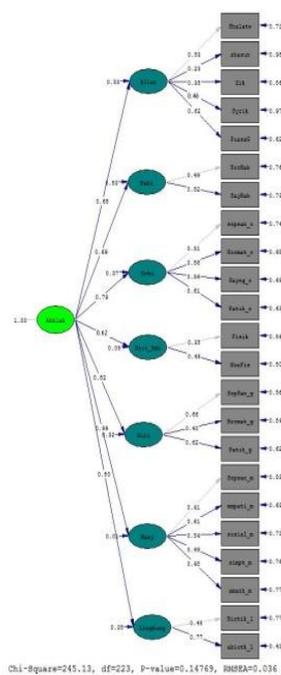
Hasil pengujian model konstruk akhlak peserta didik dan analisis individual

model konstruk tiap dimensi akhlak ditampilkan pada Gambar 4 sampai Gambar 14.



Gambar 4. Hasil Pengujian Model Konstruk Akhlak Peserta Didik

Madrasah Aliyah (dengan butir nongender)



Gambar 5. Hasil Pengujian Model Konstruk Akhlak Peserta Didik Madrasah Aliyah (dengan butir gender)

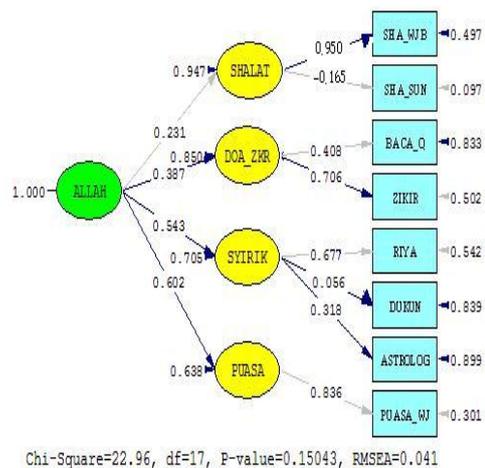
Hasil Pengujian Konstruk Akhlak kepada Allah (dimensi 1)

Gambar 6 dan 7 menunjukkan hasil analisis uji konstruk akhlak kepada Allah.

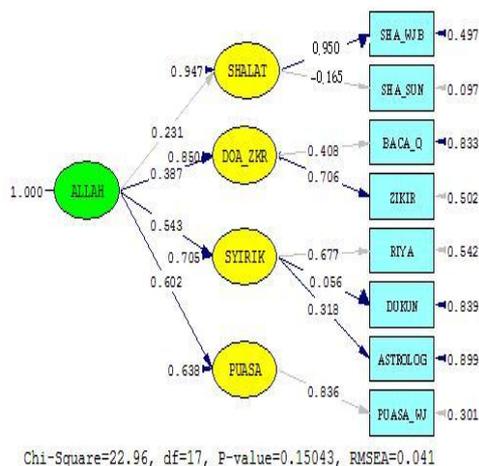
Pada level aspek dimensi, muatan faktor tertinggi dimiliki oleh sholat wajib (0,930) sedangkan muatan faktor terendah pada aspek doa dan zikir (0,442). Pada level indikator nilai muatan faktor hampir keseluruhannya positif kecuali faktor shalat sunnah dan pergi ke dukun (-0,019 dan -0,005). Muatan faktor baku terbesar dimiliki oleh faktor zikir (0,989) dan terkecil pada faktor pergi ke dukun (0,144). Hasil ini menunjukkan bahwa faktor shalat wajib merupakan faktor dominan dalam mengukur akhlak kepada Allah.

Hasil analisis juga menunjukkan bahwa akhlak peserta didik kepada Allah dapat dijelaskan oleh empat faktor yaitu keterlaksanaan shalat wajib, keterlaksanaan puasa wajib maupun sunnah, doa dan zikir, serta perbuatan syirik. Untuk indikator dengan muatan faktor negatif dilakukan kajian ulang. Gambar 10 dan 11 hasil modifikasi dua model akhlak kepada Allah.

Berdasarkan kriteria model fit yang telah ditentukan, dari dua model tersebut, keduanya fit (cocok) dengan model yang dihipotesiskan. Hal tersebut dibuktikan dengan hasil  $\chi^2=43,93$ ; p-value=0,1036; RMSEA=0,069, untuk model akhlak kepada Allah dengan tidak menyertakan butir gender disertakan dan  $\chi^2=22,96$ ; p-value=0,1504; RMSEA=0,041 untuk model akhlak kepada Allah dengan menyertakan butir gender.



Gambar 6. Model Akhlak kepada Allah Nongender

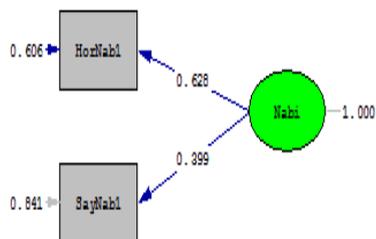


Gambar 7. Model Akhlak kepada Allah dengan Gender

Hasil Pengujian Konstruk Akhlak kepada Nabi Muhammad (dimensi 2)

Uji konstruk secara keseluruhan terhadap dimensi 2 dilakukan satu kali tahapan uji. Nilai muatan faktor untuk dimensi akhlak kepada Nabi Muhammad SAW. keseluruhannya positif dan menunjukkan nilai  $\geq 0,3$ . Nilai muatan faktor hormat kepada Nabi Muhammad SAW. 0,628, sedangkan nilai muatan faktor sayang kepada Nabi Muhammad SAW. 0,399. Hasil ini menunjukkan bahwa faktor hormat kepada nabi lebih dominan dalam mengukur akhlak kepada Nabi Muhammad SAW.

Hasil analisis menunjukkan bahwa akhlak peserta didik kepada Nabi Muhammad SAW. dapat dijelaskan oleh dua faktor yaitu hormat kepada Nabi dan sayang kepada nabi Muhammad SAW. Gambar 8 menunjukkan hasil pengujian konstruk akhlak kepada nabi Muhammad.



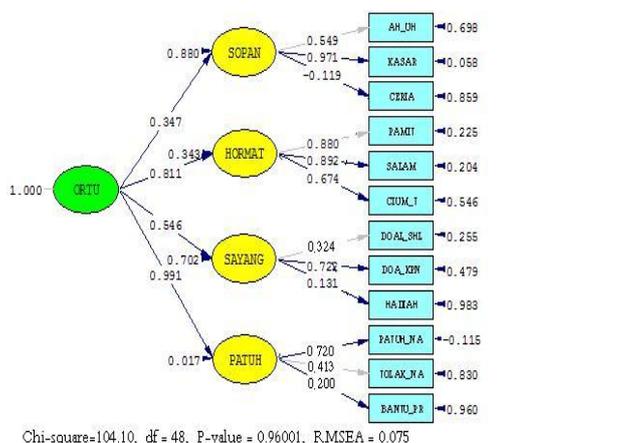
Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000

Gambar 8. Model akhlak kepada Nabi Muhammad SAW.

Hasil Pengujian Konstruk Akhlak kepada Orang tua (dimensi 3)

Uji konstruk secara keseluruhan terhadap dimensi 3 dilakukan satu tahapan uji. Hasil pengujian pada level aspek dimensi menunjukkan nilai muatan faktor untuk dimensi akhlak kepada orang tua keseluruhannya positif dan menunjukkan nilai  $\geq 0,3$ . Muatan faktor paling tinggi dimiliki oleh faktor patuh kepada orang tua yaitu 0,991. Nilai muatan faktor paling rendah dimiliki faktor sopan kepada orang tua yaitu 0,347. Pada level indikator semua muatan faktor bernilai positif kecuali indikator “berwajah ceria” (-0,119). Hasil ini menunjukkan bahwa faktor patuh kepada orang tua merupakan faktor dominan dalam mengukur akhlak kepada orang tua. Hasil analisis menunjukkan bahwa akhlak peserta didik kepada orang tua dapat dijelaskan oleh empat faktor yaitu patuh, sopan santun, hormat, dan sayang. Gambar 13 menunjukkan model akhlak kepada orang tua.

Berdasarkan kriteria model *fit* yang telah ditentukan, model tersebut *fit* (cocok) dengan yang dihipotesiskan, hal tersebut berdasarkan hasil  $\chi^2=104,10$ ; p-value= 0,96001; RMSEA=0,075. Dapat disimpulkan bahwa model akhlak kepada orang tua fit dengan model yang dihipotesiskan.

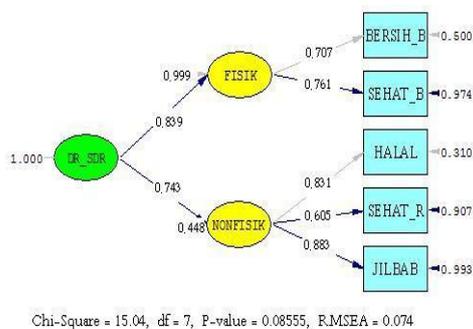


Gambar 9. Model Akhlak Kepada Orang tua

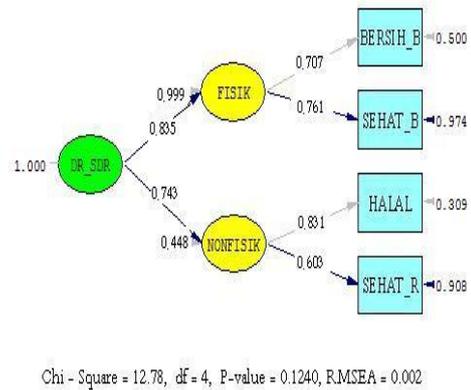
Hasil Pengujian Konstruk Akhlak kepada Diri Sendiri (dimensi 4)

Uji konstruk secara keseluruhan terhadap dimensi 4 dilakukan dua tahapan uji yaitu uji model dimensi akhlak kepada diri sendiri dengan butir nongender, dan dengan butir gender. Pada level aspek dimensi untuk kedua model (dengan butir gender maupun nonbutir gender), muatan faktor fisik mempunyai nilai lebih tinggi dari aspek nonfisik (0,839 dan 0,835). Pada level indikator semua muatan faktor bernilai positif. Muatan faktor tertinggi jika butir gender disertakan diperoleh indikator “menggunakan jilbab” (0,833), terendah pada indikator “menjaga kesehatan rohani” (0,605). Untuk model dengan non gender butir muatan faktor tertinggi diperoleh indikator “menjaga kehalalan makanan/minuman” (0,831), muatan faktor terendah pada indikator “menjaga kesehatan rohani” (0,603).

Hasil ini menunjukkan bahwa faktor fisik merupakan faktor dominan dalam mengukur akhlak kepada diri sendiri. Hasil analisis menunjukkan bahwa akhlak peserta didik kepada diri sendiri dapat dijelaskan oleh faktor fisik dan nonfisik. Gambar 10 dan 11 hasil modifikasi dua model akhlak kepada diri sendiri. Berdasarkan kriteria model *fit* yang telah ditentukan, maka dari dua model yang hipotesiskan fit atau cocok ( $\chi^2 = 15,04$  dan  $12,78$ ;  $p\text{-value} = 0,0855$  dan  $0,1240$ ;  $RMSEA = 0,074$  dan  $0,002$ ).



Gambar 10. Model Akhlak kepada Diri Sendiri dengan Butir Gender

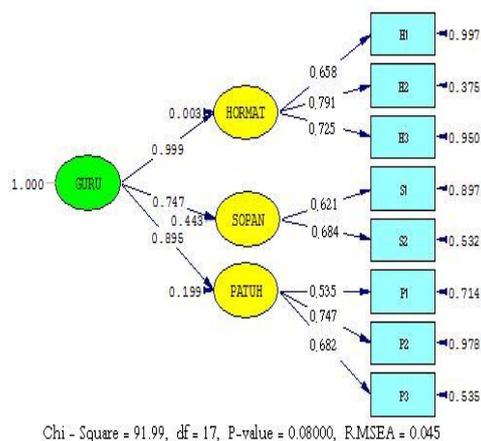


Gambar 11. Model Akhlak kepada Diri Sendiri dengan Nonbutir Gender

Hasil Pengujian Konstruk Akhlak kepada Guru (dimensi 5)

Uji konstruk secara keseluruhan terhadap dimensi 5 dilakukan satu tahapan uji. Nilai muatan faktor untuk dimensi akhlak kepada guru keseluruhannya positif dan menunjukkan nilai  $\geq 0,3$ . Pada level aspek dimensi, muatan faktor paling tinggi dimiliki oleh faktor hormat kepada guru (0,999), muatan faktor paling rendah dimiliki faktor sopan kepada guru yaitu 0,747. Untuk level indikator muatan faktor tertinggi terdapat pada indikator “memberi salam”, muatan faktor terendah pada indikator “membolos satu pelajaran” (0,535). Hasil ini menunjukkan bahwa faktor hormat kepada guru merupakan faktor dominan dalam mengukur akhlak kepada guru. Hasil analisis menunjukkan bahwa akhlak peserta didik kepada guru dapat dijelaskan oleh tiga faktor yaitu hormat, sopan santun, dan patuh.

Gambar 12 menunjukkan hasil modifikasi model akhlak kepada guru. Berdasarkan kriteria model fit yang telah ditentukan, maka model tersebut fit (cocok) dengan model yang dihipotesiskan. Hal tersebut dibuktikan dengan hasil  $\chi^2=91,99$ ;  $p\text{-value} = 0,0800$ ; dan  $RMSEA=0,045$ . Dapat disimpulkan bahwa model akhlak kepada guru fit (cocok) dengan model yang dihipotesiskan.

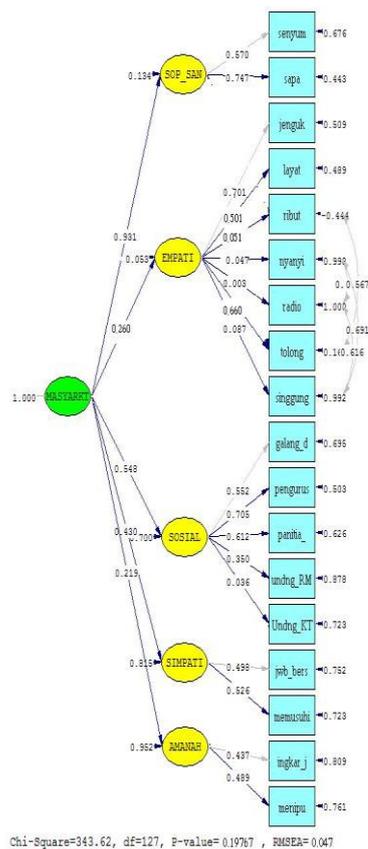


Gambar 12. Model Akhlak kepada Guru

*Hasil Pengujian Konstruk Akhlak kepada Teman/ Tetangga/ Masyarakat (dimensi 6)*

Uji konstruk secara keseluruhan terhadap dimensi akhlak kepada teman/tetangga/masyarakat dilakukan satu tahapan uji. Hasil pengujian pada level aspek maupun pada level indikator menghasilkan nilai muatan faktor keseluruhannya positif dan  $\geq 0,3$ . Pada level aspek dimensi muatan faktor paling tinggi dimiliki oleh aspek sopan santun kepada masyarakat yaitu 0,931, sedangkan nilai muatan faktor paling rendah dimiliki faktor simpati kepada masyarakat yaitu 0,219. Pada level indikator, muatan faktor tertinggi pada indikator “menjadi pengurus organisasi sosial”, muatan faktor terendah pada indikator “menyalakan radio keras-keras”. Hasil analisis menunjukkan bahwa akhlak peserta didik kepada masyarakat dapat dijelaskan oleh empat faktor yaitu empati, sopan santun, sosial, amanah dan sosial. Gambar 13 memperlihatkan hasil modifikasi model akhlak kepada masyarakat.

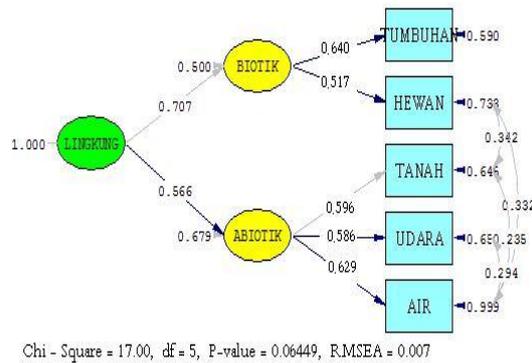
Berdasarkan kriteria model fit yang telah ditentukan, maka model tersebut fit (cocok). Hal tersebut dibuktikan dengan hasil  $\chi^2 = 343,62$ ;  $p\text{-value} = 0,198$ ; dan  $RMSEA = 0,047$ . Dapat disimpulkan bahwa model akhlak kepada masyarakat *fit* (cocok).



Gambar 13. Model Akhlak kepada Teman/Tetangga/ Masyarakat

*Hasil Pengujian Konstruk Akhlak kepada Lingkungan (dimensi 7)*

Uji konstruk secara keseluruhan terhadap dimensi 7 dilakukan satu tahapan uji. Uji konstruk tersebut menghasilkan nilai muatan faktor untuk dimensi akhlak kepada lingkungan keseluruhannya positif dan menunjukkan nilai  $\geq 0,3$ . Nilai muatan faktor lingkungan abiotik (0,968) lebih tinggi dari nilai muatan faktor lingkungan biotik (0,383). Hasil ini menunjukkan bahwa faktor lingkungan abiotik merupakan faktor dominan dalam mengukur akhlak kepada lingkungan. Hasil analisis menunjukkan bahwa akhlak peserta didik kepada lingkungan dapat dijelaskan oleh dua faktor yaitu perilaku terhadap lingkungan abiotik dan perilaku terhadap lingkungan biotik. Gambar 14 hasil modifikasi model akhlak kepada lingkungan.



Gambar 14. Model Akhlak kepada Lingkungan

Model Penilaian Akhlak Peserta Didik Madrasah Aliyah (PAPDMA)

Penentuan skor minimal tiap dimensi akhlak peserta didik.

Hasil kesepakatan para ahli (*expert judgment*) yang terdiri atas dosen PAI, guru PAI atau guru akhlak di Madrasah Aliyah Negeri dan Swasta, Guru PAI atau guru akhlak di SMA, Pengawas PAI di Madrasah Aliyah dan SMA, disimpulkan sebagaimana ditampilkan pada tabel 12.

Tabel 12. Skor Minimal Tiap Dimensi Akhlak

No Dimensi/Aspek	Skor minimum butir
1 Akhlak kepada Allah	
a. Sholat wajib	3
b. Sholat sunnah	2
c. Zikir dan doa	3
d. Puasa	
1) Wajib	4
2) Sunnah	2
e. Syirik	
1) Riya	2
2) Percaya astrologi	3
3) Pergi ke dukun/paranormal	4
2 Akhlak kepada Nabi Muhammad SAW	2
3 Akhlak kepada diri sendiri	
a. Kebersihan badan	3
b. Kesehatan badan	1
c. Menjaga kehalalan makan/minum	4
d. Menutup aurat	3
4 Akhlak kepada orang tua	3
5 Akhlak kepada guru	3
6 Akhlak kepada masyarakat	2
7 Akhlak kepada lingkungan	2

Keterangan skor minimum

1. **Jarang** untuk pernyataan arah positif atau **sering** untuk pernyataan arah negatif
2. **kadang-kadang** untuk pernyataan arah positif atau arah negatif
3. **sering** untuk pernyataan arah positif atau **jarang** untuk pernyataan arah negatif
4. **selalu** untuk pernyataan arah positif atau **tidak pernah** untuk pernyataan arah negatif.

Sebagai contoh, untuk akhlak kepada Allah skor minimum untuk butir shalat wajib skor minimum 3, hal ini berarti bahwa untuk pernyataan positif (misalnya: melaksanakan shalat subuh di awal waktu), minimum dilakukan **sering**. Untuk pernyataan negatif (misalnya: meninggalkan shalat subuh) skor minimumnya adalah **jarang**.

Untuk akhlak kepada nabi Muhammad SAW. skor minimum ditetapkan 2, hal ini berarti untuk tiap indikator butir aspek akhlak kepada nabi Muhammad SAW. minimum dilakukan **kadang-kadang** untuk pernyataan positif maupun pernyataan negatif.

Untuk dimensi akhlak kepada diri sendiri ditetapkan skor minimum 4 untuk aspek menjaga kehalalan makanan dan minuman, hal ini berarti untuk pernyataan positif minimum **selalu** dilakukan, sedangkan untuk pernyataan negatif minimum **tidak pernah** dilakukan. Untuk aspek menjaga kebersihan badan dan menutup aurat, skor minimum ditetapkan 3, hal ini berarti untuk tiap indikator butir aspek menutup aurat minimum dilakukan **sering** untuk pernyataan positif, untuk pernyataan negatif minimum dilakukan **jarang**. Untuk aspek menjaga kesehatan ruhani ditetapkan skor minimum 2, hal ini berarti untuk tiap indikator butir aspek menjaga kesehatan ruhani minimum dilakukan **kadang-kadang** untuk pernyataan positif dan pernyataan negatif. Untuk aspek menjaga kesehatan badan skor minimum ditetapkan 1, hal ini berarti untuk tiap indikator menjaga kesehatan minimum **jarang di-**

**lakukan** untuk pernyataan positif atau **sering dilakukan** untuk pernyataan negatif

Untuk dimensi akhlak kepada orang tua dan guru ditetapkan skor minimum tiap indikator adalah 3 hal ini berarti untuk semua butir indikator aspek akhlak kepada orang tua dan guru minimum **sering dilakukan** untuk pernyataan positif, untuk pernyataan negatif minimum dilakukan **jarang**. Untuk dimensi akhlak kepada masyarakat dan kepada lingkungan ditetapkan skor minimum 2, hal ini berarti untuk setiap indikator akhlak kepada masyarakat dan lingkungan minimum kadangkadang dilakukan baik untuk pernyataan positif maupun negatif.

*Kriteria Ketuntasan Minimum (KKM)*

Setelah skor minimum tiap dimensi ditentukan, selanjutnya ditentukan Kriteria

Ketuntasan Minimum (KKM) untuk setiap instrumen. Tabel 13, 14 dan 15 menunjukkan KKM untuk instrumen IA, IB, IIA dan IIB

Berdasarkan penentuan KKM dari table 13, 14 dan 15 tersebut maka KKM untuk instrumen IA (*self report*) adalah sebesar 219 untuk kategori gender (bila butir bermuatan gender disertakan), untuk kategori non gender (bila butir gender tidak disertakan) KKM berjumlah 212. Selanjutnya KKM untuk penilaian antar teman (instrumen IB) dan penilaian oleh guru akhlak (instrumen IIA) adalah sebesar 159 untuk kategori gender (bila butir bermuatan gender disertakan) KKM berjumlah 156. Besar KKM untuk instrumen II B adalah sebesar 105 untuk kategori gender. Untuk kategori nongender, KKM sebesar 102.

Tabel 13. Penentuan KKM Instrumen IA

No	Dimensi/Aspek	Skor minimum butir	Jumlah butir	Jumlah skor minimum
1	Akhak kepada Allah			
	a. Sholat wajib	3	10	30
	b. Sholat sunnah			
	1) Rawatib	2	1	2
	2) Dhuha	2	1	2
	3) Tahajud	1	1	1
	c. Zikir dan doa	2	2	4
	d. Puasa			
	1) Wajib 1	4	1	4
	2) Wajib 2 (gender)	4	1	4
	3) Sunnah	2	1	2
	e. Syirik			0
	1) Riya	3	1	3
	2) Percaya astrologi	3	1	3
	3) Pergi ke dukun/paranormal	4	1	4
2	Akhilak kepada Nabi Muhammad SAW.	2	3	6
3	Akhilak kepada diri sendiri			0
	a. Kebersihan badan	4	2	8
	b. Kesehatan badan	2	1	2
	c. Menjaga kehalalan makan/minum	4	5	20
	d. Menjada kesehatan ruhani	2	2	4
	e. Menutup aurat	3	1	3
4	Akhilak kepada orang tua	3	12	36
5	Akhilak kepada guru	3	9	27
6	Akhilak kepada masyarakat	2	18	36
7	Akhilak kepada lingkungan	2	9	18
			Gender	219
			Nongender	212

Tabel 14. Penentuan KKM untuk Instrumen I B dan II A (rater: antarteman dan guru akhlak)

No Dimensi/Aspek	Skor minimum butir	Jumlah butir	Jumlah skor minimum
1 Akhlak kepada Allah	3	14	42
2 Akhlak kepada Nabi Muhammad SAW	3	2	6
3 Akhlak kepada diri sendiri	3	8	18
4 Akhlak kepada Orang tua	3	4	12
5 Akhlak kepada guru	3	8	24
6 Akhlak kepada Masyarakat	3	15	45
7 Akhlak kepada lingkungan	3	4	12
Jumlah	Gender Nongender	55 54	159 156

Tabel 15. Rancangan Sistem Penilaian Instrumen II (rater: guru bidang studi lain dan konseling)

No Dimensi/Aspek	Skor minimum butir	Jumlah butir	Jumlah skor minimum
1 Akhlak kepada orang tua	3	4	12
2 Akhlak kepada diri sendiri	3	8	24
3 Akhlak kepada guru	3	6	18
4 Akhlak kepada masyarakat	3	15	45
5 Akhlak kepada lingkungan	3	2	6
Jumlah	Gender Nongender	35 34	105 102

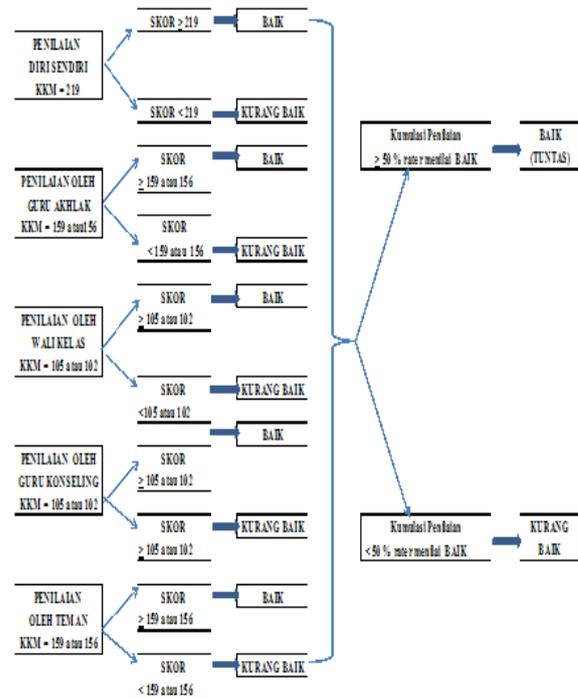
Implementasi Sistem Penilaian Akhlak Peserta Didik

Berdasarkan sistem penilaian yang telah ditentukan, serta mengacu pada tabel

Tabel 16. Nilai Akhir Akhlak Peserta Didik

No	Nama Madrasah	Jumlah siswa dinilai	Siswa Kategori berakhlak "Baik"		Siswa kategori berakhlak "kurang baik"	
			Jumlah (siswa)	Persentase (%)	Jumlah (siswa)	Persentase (%)
1	MAN Kota Tangerang	83	83	100	0	0
2	MAS Al Ijtihad	29	21	72,41	8	27,59
3	MAN Tiga Raksa	57	48	84,21	9	15,79
4	MAS Mubaroq	35	28	80,00	24	11,76
Jumlah Total		204	180	88,24	24	11,76

13, 14, dan 15, sistem penilaian akhlak peserta didik dimodifikasi sebagaimana tampak pada gambar 15. Dalam penelitian ini, model penilaian diterapkan untuk siswa kelas XII sehingga penilaian akhir menggunakan dua kategori yaitu kategori "baik" dan "kurang baik".



Gambar 15. Sistem Penilaian Akhlak Peserta Didik

Dengan mengacu pada pendapat Lawshe (1975) pada proses *quantifying consensus*, diperoleh hasil penilaian akhir akhlak peserta didik oleh setiap rater dalam bentuk persentase ditampilkan pada tabel 16.

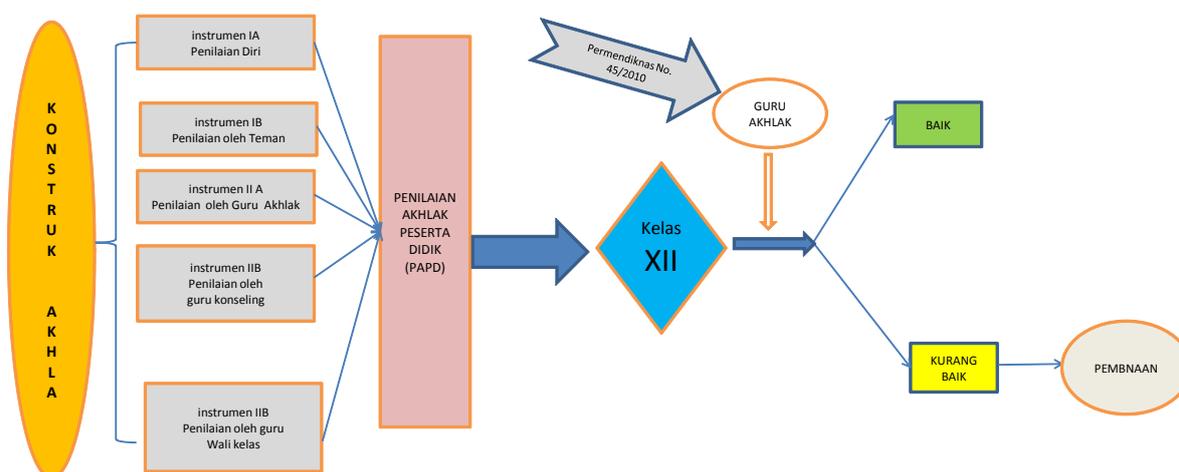
Berdasarkan hasil tersebut, pada MAN Kota Tangerang jumlah peserta didik yang memperoleh penilaian akhlak kategori “baik” berjumlah 100 % atau sejumlah 83 siswa responden; pada MAS al-Ijtihad Kota Tangerang jumlah peserta didik yang memperoleh penilaian akhlak kategori “baik” berjumlah 21 siswa dari 29 responden atau 72,41%; pada MAN Tigaraksa Kabupaten Tangerang jumlah peserta didik yang memperoleh penilaian akhlak kategori “baik” berjumlah 48 siswa atau 84,21 dari 57 siswa responden; pada MAS al-Mubaroq Kabupaten Tangerang jumlah peserta didik yang memperoleh penilaian akhlak kategori “baik” berjumlah 28 siswa atau 80 % dari 35 siswa responden.

Secara keseluruhan dapat disimpulkan 180 siswa atau 88,24 % dari 204 siswa yang dinilai memperoleh nilai akhlak “baik” dan sebanyak 24 siswa atau 11,76 % dari 204 siswa memperoleh penilaian akhlak kategori baik “kurang baik”. Hasil konfirmasi penilaian akhlak terhadap 85 siswa MAN kota Tangerang yang telah diperoleh hasil penilaian “baik” dan “kurang baik” diperoleh kecocokan mendekati 100 persen (hanya 2

siswa dari 85 siswa yang tidak cocok). Hasil konfirmasi penilaian akhlak peserta didik di MAN Tigaraksa 1 Kabupaten Tangerang diperoleh kecocokan 98% yaitu hanya 3 orang siswa yang tidak cocok.

#### Skema Model Penilaian Akhlak Peserta Didik Madrasah Aliyah (PAPDMA)

Gambar 16 menjelaskan alur penilaian akhlak peserta didik Madrasah Aliyah dimulai dengan menggunakan instrumen pengukuran akhlak yang terdiri dari 4 instrumen dengan 5 penilai. Hasil penilaian 5 penilai terhadap siswa kelas XII diolah guru akhlak dengan menggunakan *software* penilaian akhlak. Pengolahan nilai menghasilkan dua kategori nilai akhir akhlak peserta didik yaitu peserta didik berakhlak baik dan berakhlak kurang baik. Peserta didik berakhlak baik mendapat sertifikat atau sejenis surat keterangan berakhlak **baik** dari Madrasah Aliyah yang ditandatangani Kepala MA dan guru akhlak. Bagi peserta didik yang mendapat nilai akhir **kurang baik** dilakukan pembinaan pada aspek yang masih perlu dibina (tergantung hasil penilaian dari lima rater tersebut).



Gambar 16. Model Penilaian Akhlak Peserta Didik Madrasah Aliyah (PAPDMA)

#### Simpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan hal-hal sebagai berikut.

Konstruk akhlak peserta didik didasarkan pada definisi akhlak yaitu keadaan jiwa peserta didik yang secara komperhensif menghasilkan perbuatan yang secara rutin

dilakukan dan berulang. Dimensi akhlak mencakup dimensi akhlak kepada Allah, akhlak kepada Nabi Muhammad, akhlak kepada orang tua, akhlak kepada diri sendiri, akhlak kepada guru, akhlak kepada masyarakat/tetangga/teman dan akhlak kepada lingkungan. Pengukuran akhlak peserta didik menggunakan instrumen yang disusun berdasarkan dimensi akhlak tersebut. Model Pengukuran Akhlak Peserta Didik yang disusun berdasarkan konstruk akhlak peserta didik *fit* (cocok) dengan model yang ditekankan ( $\chi^2=254,93$ ;  $p=0,06988$  dan  $RMSEA=0,048$ ).

Konstruk akhlak kepada Allah didefinisikan dengan keterlaksanaan shalat wajib, puasa wajib, puasa sunnah, berzikir dan berdoa, dan syirik. Konstruk akhlak kepada Nabi Muhammad SAW., terdiri atas aspek hormat dan sayang. Konstruk akhlak kepada orang tua terdiri atas aspek sopan santun, hormat, patuh dan sayang. Konstruk akhlak kepada diri sendiri terdiri atas aspek fisik dan non fisik. Konstruk akhlak kepada guru terdiri atas aspek sikap sopan santun, hormat, dan patuh. Akhlak kepada masyarakat/tetangga/teman terdiri atas aspek sikap sopan, simpati, empati, sikap sosial dan amanah.

Hasil pengujian model konstruk akhlak kepada Allah, akhlak kepada Nabi Muhammad, akhlak kepada orangtua, akhlak kepada diri sendiri, akhlak kepada guru, akhlak kepada teman/tetangga/masyarakat, dan akhlak kepada lingkungan, seluruhnya menghasilkan model *fit* dengan yang ditekankan terbukti dengan hasil berturut-turut:

$\chi^2=43,93$ ;  $p\text{-value}=0,10362$ ; dan  $RMSEA=0,069$ ;  $\chi^2=0,000$ ;  $p\text{-value}=1,000$ ; dan  $RMSEA=0,000$ ;  $\chi^2=104,10$ ;  $p\text{-value}=0,96001$ ; dan  $RMSEA=0,075$ ;  $\chi^2=15,04$ ;  $p\text{-value}=0,0855$ ; dan  $RMSEA=0,074$ .  $\chi^2=91,99$ ;  $p\text{-value}=0,08000$ ; dan  $RMSEA=0,045$ ;  $\chi^2=343,62$ ;  $p\text{-value}=0,19767$ ; dan  $RMSEA=0,047$ ;  $\chi^2=17,000$ ;  $p\text{-value}=0,06449$ ; dan  $RMSEA=0,007$ . Hasil pengujian model akhlak kepada Nabi Muham-

mad, diperoleh  $df=0$ , hal ini menunjukkan bahwa model *saturated* dan *over confident* atau model hanya dapat digunakan pada sampel yang digunakan dan data input lebih besar dari jumlah parameter yang dianalisis.

Instrumen pengukuran akhlak peserta didik terdiri dari 4 instrumen yaitu instrumen IA berisi penilaian diri peserta didik; instrumen IB berisi penilaian oleh teman, instrumen IIA berupa penilaian oleh guru akhlak; dan instrumen IIB berisi penilaian oleh guru konseling/walikelas.

Reliabilitas konsistensi internal instrumen IA berdasarkan kategori jurusan diperoleh untuk jurusan IPA koefisien alpha 0,874; IPS koefisien alpha 0,865; Agama koefisien alpha 0,921. Untuk kategori gender diperoleh untuk kategori putra koefisien alpha 0,883 dan kategori putri koefisien alpha 0,874.

Hasil analisis konsensus interrater antara guru akhlak dengan teman dan antara guru konseling dan wali kelas menghasilkan konsensus yang dapat diterima karena di atas 70 % yaitu 0,843 dan 0,866. Begitu pula hasil analisis dengan formulasi Cohens' Kappa diperoleh koefisien Kappa antara guru akhlak dengan teman sebesar 0,770; antara guru konseling dengan walikelas sebesar 0,713. Menurut Landis dan Koch (2004) hal tersebut menunjukkan konsensus antar rater sangat baik.

Tingkat stabilitas instrumen IA **baik** ( $r_{\text{test re-test}}=0,749$ ); instrumen IB sangat baik ( $r_{\text{test re-test}}=0,827$ ); instrumen IIA sangat baik ( $r_{\text{test re-test}}=0,858$ ), rerata instrumen II B rater walikelas dan konseling baik ( $r_{\text{test re-test}}=0,767$  dan 0,715).

Hasil pengembangan model penilaian akhlak peserta didik Madrasah Aliyah terdiri atas: seperangkat instrumen penilaian akhlak peserta didik beserta pedomannya, sistem penilaian dan skema model penilaian dan *software* penilaian yang berisi manual cara penilaian dan sistem aplikasi penilaian akhlak peserta didik. Model ini disebut : Model Penilaian Akhlak Peserta Didik Madrasah Aliyah (PAPDMA).

## Daftar Pustaka

- Abdalati, Hammuda. (1978). *Islam in focus*. Indianapolis, USA: Amanna Publications.
- Al-Qardhawi, Yusuf. (1981). *Kumpulan hikmah al-Qur'an*. Jakarta: Bumi Aksara Press.
- Anderson, Lorin W. (1981). *Assessing affective characteristics in the schools*. Boston Massachusetts: Allyn and Bacon, Inc.
- Anshari, Endang Saefuddin. (1986). *Wawasan Islam*. Jakarta: Rajawali.
- Azwar, Saefuddin. (2012). *Reliabilitas dan Validitas*. (edisi 4). Pustaka Pelajar, Yogyakarta.
- Gall, J., Borg. W., & Gall, M. (2003). *Educational research: An introduction* (7th ed.). Boston: Pearson Education
- Braswell, George W. (1996). *Islam; It's prophet, people, politics and power*. Nashville, TN: Broadman and Holman Publisher.
- BSNP, (2009). *8 Standar Nasional Pendidikan*. Jakarta: Depdiknas.
- \_\_\_\_\_. (2010). *POS Ujian Nasional 2011*. Jakarta: Kemendiknas.
- Depdiknas. (2006). *Undang-Undang Republik Indonesia Nomor 20 tahun 2005 tentang Sistem Pendidikan Nasional*. Jakarta: DepDikNas.
- Depdiknas. (2005). *Peraturan Pemerintah Republik Indonesia Nomor 19 tahun 2005 Tentang Standar Nasional Pendidikan*. Jakarta: DepDikNas.
- Depdiknas. (2007). *Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 16 tahun 2007 Tentang Standar Kualifikasi Akademik dan Kompetensi Guru*. Jakarta: DepDikNas.
- Goodlad, John I. (1984). *A place called school, prospect for the future*. USA: McGraw-Hill Book Company.
- Griffin, Patrick E. (1991). *Educational assessment and reporting*. Harcourt Brace Jovanovich Group. Australia.
- Lawshe. (1975). *A quantitative approach to content validity*. Journal Personal Psychology. India: Purdue University.
- Levin, Henry M. (2011). *The utility and need for incorporating non-cognitive skill into large scale educational assessments*. Presented at ETS Invitational on International Large – Scale Assessment and to be published by Educational Testing Service. USA: Teachers College, Columbia University.
- Miskawaih. (1968). *Tahdzibul akhlak*. Edisi bahasa Indonesia diterjemahkan oleh Mustofa. Pustaka Setia. Bandung.
- Shihab, Quraish. (1996). *Wawasan al-Qur'an*. Bandung: Mizan.
- Rahim, Husni. (2001). *Arab baru pendidikan Islam di Indonesia*. Jakarta: Logos Wacana Islam.
- Robinson, John. P and Shaver, Philip R. (1969). *Measures of social psychological attitudes. dimension of religiosity*. Survey Research Center (Institute for Social Research) The University of Michigan, Michigan, USA.
- Siddiqui, Ataulah. (1997). *Ethics in Islam: key concepts and contemporary challenges*. Journal of Moral Education, V.26. No:4. Liechester, UK: Islamic Foundation.
- Steven M. Cahn, Peter J. Markie. (2006). *Ethics: history, theory, and contemporary issues*. Oxford University Press.
- Swarup, Ram. (1983). *Understanding Islam through hadis*. New Delhi, India. Voice of India.
- Waluchow, Wilfrid.J. (2003). *The Dimension of ethics: an introduction to ethical theory*. Broadviewpress. Canada
- Yani, Ahmad. (2009). *Be excellent, menjadi pribadi yang terpuji*. Jakarta: Khairu Ummah.