

Pengembangan dan Validasi Instrumen “Persepsi Siswa terhadap Karakter Moral Guru” di Indonesia dengan Model Rasch¹

Ifa H Misbach² dan Bambang Sumintono³

ABSTRAK

Tujuan dari penelitian ini adalah melakukan studi pendahuluan (*pre-liminary study*) untuk mengembangkan sistem pendidikan moral di Indonesia yang lebih akuntabel dan terukur (*accountable* dan *measureable*). *Pre-liminary study* ini dilakukan dengan cara mengembangkan instrumen baru untuk mengukur apakah karakter moral guru di Indonesia masih dapat dijadikan *role model* berdasarkan persepsi siswa. Instrumen ini dinamakan “Persepsi Siswa terhadap Karakter Moral Guru”. Sampel penelitian terdiri dari siswa SMA kelas 12 sebanyak 228 orang dari enam buah sekolah di wilayah kota Bandung. Format instrumen diisi secara manual (*paper-and-pencil*). Metodologi yang digunakan adalah kuantitatif, menggunakan kuesioner dengan peringkat Likert dengan 7 pilihan dalam bentuk *semantic differential*. Pegujian analisis data menggunakan model Rasch melalui perangkat lunak Winsteps. Temuan menarik yang didapatkan bahwa siswa sekolah negeri dan swasta menunjukkan pola respon yang berbeda; sedangkan analisis aitem menunjukkan adanya dua aitem yang tidak sesuai (*misfit*) maupun aitem dalam satu konstruk mengukur logit yang sama; instrumen memenuhi syarat unidimensionalitas dan beberapa aitem terdeteksi mengandung bias; serta validitas peringkat menunjukkan penyederhaan peringkat akan mendapatkan hasil pengukuran yang lebih tepat.

PENDAHULUAN

Model pendidikan di Indonesia memiliki sistem keyakinan bahwa secara tradisi guru memiliki status istimewa sebagai *role model* yang dipercaya memegang peranan penting sebagai panutan siswa (Thomas, 1962). Namun demikian, sejak sistem ujian nasional (UN) diberlakukan dari tahun 2003 sampai 2014, sekolah tidak lagi menjadi tempat pertumbuhan pendidikan dalam membina karakter moral siswa moral (D. Joesoef, Juni 2011). Sekolah telah berubah fungsi dari pusat pendidikan nilai-nilai karakter menjadi pabrik bimbingan belajar dalam bentuk latihan (*drill*) pemadatan materi, menghafalkan segala jenis soal UN yang sama dengan tahun sebelumnya. Dari tahun 2003 sampai 2012, eskalasi jumlah kecurangan yang dilakukan oknum guru telah meningkat lebih dari 750 kasus sejak pertama UN dilaksanakan (Jibi, 2012). Data tersebut hanya diduga sebagian kecil dari permukaan gunung es masalah kecurangan yang sebetulnya lebih banyak lagi yang tidak terdeteksi. Peningkatan kecurangan yang dilakukan oleh banyak oknum guru untuk meluluskan siswa membuat keresahan di kalangan pemerhati pendidikan bahwa banyak siswa di Indonesia menghadapi krisis *role model* untuk bisa membedakan mana yang baik dan buruk, khususnya di lingkungan sekolah.

¹ Paper dipresentasikan dalam Seminar Nasional Psikometri dengan tema “Pengembangan Instrumen penilaian Karakter yang Valid” di Hotel Lorin Solo, Fakultas Psikologi, Universitas Muhammadiyah Surakarta.

² dosen Jurusan Psikologi, Fakultas Ilmu Pendidikan, Universitas Pendidikan Indonesia, Bandung.
email: ifahmisbach@gmail.com

³ dosen pada Institute of Educational Leadership, Universiti Malaya, Kuala Lumpur, Malaysia
Email: bambang@um.edu.my dan deceng@gmail.com

Federasi Serikat Guru Indonesia (FSGI) melaporkan bahwa guru berada dalam posisi yang sangat rentan untuk disalahkan oleh kepala sekolah atau kepala dinas ketika siswa gagal UN (Pratiwi & Djumena, 2011). Banyak ahli pendidikan di Indonesia berpendapat bahwa pelaksanaan UN telah mengacaukan paradigma moral dalam sistem pendidikan di Indonesia, yang telah memberikan tekanan psikologis dalam bentuk dilema moral yang hebat pada guru untuk melakukan perilaku demoralisasi, dalam bentuk memanipulasi nilai siswa, mendistribusikan jawaban, dan bahkan lebih buruk, mengintimidasi siswa untuk berbagi jawaban mereka dengan seluruh kelas. Keberhasilan guru dalam melakukan pendidikan hanya diukur dari angka, lulus atau tidak lulus, sementara nilai-nilai luhur kemanusiaan yang mendasar tentang kejujuran telah diabaikan.

Kasus UN 2011 di SDN Gadel Surabaya merupakan luka moral dimana sistem pendidikan moral di Indonesia sering mengabaikan sudut pandang siswa dengan selalu menempatkan siswa dalam posisi yang inferior yang diliputi ketakutan. Seorang siswa dan orang tua yang melaporkan oknum guru yang memaksa anaknya memberikan jawaban ke seluruh kelas, harus dipaksa meminta maaf secara terbuka karena warga satu desa yang takut semua anaknya tidak lulus karena terbongkarnya kasus contek massal ke media. Ironisnya, orang tua dan anak tersebut harus mengungsi dari rumahnya sendiri demi menghindari tekanan agresif yang intensif dari para warga sekitar. Berdasarkan perspektif *social learning*, anak lain yang melihat kejadian ini akan belajar melihat konsekuensi bahwa kejujuran diasosiasikan dengan hukuman keras. Anak akan berpikir dalam *moral reasoning*-nya bahwa lebih baik saya menjadi seorang pembohong tetapi selamat daripada saya jujur tetapi saya dihukum keras oleh masyarakat yang tidak suka jika saya jujur. Inilah keruntuhan moral yang akan dipanen saat anak dewasa berganti memimpin negeri ini sebagai ancaman munculnya *the lost generation*.

Dalam konteks budaya feodal seperti di Indonesia, sebagian besar guru-guru tidak sepenuhnya menyadari bahwa peran mereka sebagai *role model* dalam posisi superior, memiliki dampak langsung untuk ditiru siswa dalam posisi yang lebih inferior. Perilaku demoralisasi pada guru sebagai role model akan memiliki implikasi langsung pada persepsi siswa tentang karakter moral gurunya. Salah satu dampaknya adalah siswa dapat mengalami kesulitan di dalam memahami relevansi nilai-nilai moral yang dipelajari di ruang kelas bertentangan dengan perilaku pelanggaran moral yang mereka amati di luar kelas. Lumpkin (2008) menunjukkan bahwa siswa yang secara langsung mengalami kecurangan, ketidakjujuran, atau korupsi yang ditunjukkan oleh guru, akan memaknakan bahwa semua bentuk perilaku curang dianggap wajar karena siswa diperbolehkan untuk meniru hal serupa.

Sayangnya, di Indonesia, tidak tersedia cukup kajian penelitian yang luas untuk mengevaluasi sejauhmana dampak karakter moral guru pada karakter moral siswa di Indonesia (Misbach, 2013). Meskipun tantangan yang signifikan karena tidak ada literatur yang komprehensif mengenai desain pendidikan moral maupun instrumen moral untuk menilai karakter moral *role model* dalam sistem pendidikan di Indonesia, peneliti merasa perlu mengembangkan instrument karakter moral untuk mengukur konstruk moral dengan cara yang lebih dapat dipertanggungjawabkan. Bagian berikutnya akan membahas kajian pustaka yang

berhubungan, pendekatan metodologi penelitian ini menggunakan model Rasch dimana analisis dan pembahasan temuan riset ditampilkan serta ditutup dengan kesimpulan.

KAJIAN PUSTAKA

Peran Model dalam Teori Belajar Sosial

Menurut teori belajar sosial, peran seorang model adalah salah satu alat yang paling kuat untuk menularkan nilai-nilai, sikap, dan pola pikir dan perilaku pada orang lain (Bandura, 1986). Asumsi teoritis dari penelitian ini adalah guru dapat berperan sebagai model yang mengajarkan nilai-nilai moral dan karakter moral (Kohlberg, 1981; Lickona, 1991; Noddings, 1992) di sekolah-sekolah Indonesia. Meninjau guru sebagai panutan, mereka sering diakui sebagai komponen penting dari pendidikan moral yang mempengaruhi perilaku moral siswa (Bandura, 2002). Teori pembelajaran sosial berfokus pada bagaimana siswa belajar dengan mengamati dan melakukan proses imitasi dari model yang menjadi panutannya (Bandura, 1963, 2004).

Bandura (1971) mengusulkan beberapa karakteristik yang perlu dimiliki agar seseorang dapat menjadi model yang efektif. Pertama, model harus dinilai kompeten dalam perilaku yang diamati. Jika guru dipandang kompeten, perilaku mereka lebih mungkin untuk ditiru siswa. Kedua, model memiliki prestise dan kekuasaan. Karena budaya dan masyarakat Indonesia mengakui guru sebagai representasi dari orang tua siswa saat mereka berada di sekolah, guru menduduki posisi penting sebagai panutan yang luar biasa dan dihargai. Guru memegang posisi istimewa sebagai figur otoritas yang memiliki status sosial, rasa hormat, dan kekuasaan. Ketiga, model berperilaku dalam cara yang tidak bias sesuai stereotip gendernya. Ada harapan budaya bahwa guru perempuan harus melakukan perilaku dalam sosok keibuan. Sebaliknya, guru laki-laki harus mengekspresikan perilaku sosok kebabakan. Baik guru perempuan dan laki-laki dihargai sebagai pengganti orang tua di sekolah. Keempat, perilaku model relevan dengan situasi pengamat. Agar menjadi model yang efektif, perilaku guru harus sesuai dengan apa yang mereka ucapkan. Sebagai pengamat, siswa akan mengamati apakah yang guru lakukan bertentangan dengan apa yang guru katakan. Proses imitasi muncul ketika siswa sebagai individu merasakan perilaku yang sama atau relevan antara diri mereka sendiri dengan guru yang telah dianggap sebagai model (Bandura, 2003; Holyoak & Thagard, 1989).

Teori belajar sosial mengakui bahwa individu belajar melalui konsekuensi, baik oleh penguatan atau hukuman. Hal ini dapat terjadi dalam beberapa cara yang mungkin. Pertama, tindakan pengamat diperkuat melalui tindakan yang meniru model. Kedua, pengamat diperkuat oleh orang ketiga, peserta lain dalam lingkungan, seperti siswa lain atau guru lain. Ketiga, perilaku yang ditiru dapat memperkuat konsekuensi positif. Keempat, konsekuensi perilaku yang dialami sendiri oleh model akan mempengaruhi perilaku pengamat. Hal ini dikenal sebagai *vicarious reinforcement* (Bandura, Ross & Ross, 1963). Artinya, respon yang dihasilkan model akan membuat peningkatan respon serupa yang dilakukan pengamat. Misalnya, ketika siswa mengamati bahwa guru menunjukkan perilaku tidak bermoral, mereka memiliki kecenderungan meningkat untuk meniru perilaku dengan cara yang sama. Sebaliknya, ketika siswa mengamati bahwa korban yang berjuang untuk kejujuran dihukum oleh masyarakat, mereka belajar bahwa

perilaku yang baik belum tentu didukung di masyarakat. Akibatnya, siswa akan belajar untuk menjadi sadar dari kedua perilaku dan mengkalkulasi konsekuensi yang juga akan mereka dapatkan. Misalnya siswa akan berpikir bahwa mungkin lebih baik untuk berbohong daripada mengatakan yang sebenarnya, untuk menghindari hukuman dari masyarakat setempat.

Karakter Moral

Moralitas berada di persimpangan kedua isu, yaitu etika normatif dan psikologi empiris (Timpe, 2007). Perspektif dan asumsi yang berbeda dalam mengkaji literatur moral, akan memiliki teori pendekatan moral yang berbeda. Dalam konteks penelitian ini, moralitas pada akhirnya merupakan bentuk karakteristik tindakan (Blasi, 1980).

Berkowitz (2002) mendefinisikan karakter sebagai satu set karakteristik psikologis individu yang mempengaruhi kemampuan dan kecenderungan seseorang untuk berfungsi secara moral. Menurut Berkowitz, menjadi orang yang bermoral dalam hal apa yang disebutnya anatomi moral, terbentuk dari "perilaku moral, nilai-nilai moral, kepribadian moral, emosi moral, penalaran moral, identitas moral, dan karakteristik dasar" (Berkowitz, 2002, hal.48).

Sejalan dengan posisi Berkowitz, Lickona (1991) mendefinisikan karakter yaitu bagaimana individu melakukan hal yang benar tanpa tekanan untuk menghadapi sesuatu yang ditentang. Dia mengusulkan sebuah model untuk menilai karakter yang terdiri dari tiga komponen psikologis, yaitu: pengetahuan moral, perasaan moral, dan tindakan moral. Karakter adalah sebuah fenomena universal yang deskriptif pada orang yang memiliki keberanian dan keyakinan untuk hidup dengan kebajikan moral (Lumpkin, 2008). Ukuran keutamaan karakter terletak pada tindakan (Lickona dan Davidson (2005).

Dimensi-dimensi Karakter

Konstruk karakter moral secara konseptual lebih luas daripada konstruk penalaran moral. Vessel (1998) meneliti konstruk karakter moral yang menggabungkan domain kognisi atau penalaran moral, yang juga berkaitan dengan domain afektif dan perilaku. Vessel (1998) memberikan tinjauan ekstensif dari 12 dimensi karakter yang paling sering dikutip dalam banyak literatur. Sejumlah dimensi karakter yang digunakan dalam penelitian ini adalah: integritas, kejujuran, kepercayaan, keadilan, rasa hormat, kesetiaan, tidak mementingkan diri sendiri, kasih sayang, penghargaan spiritual, kerjasama, peduli, dan tanggung jawab (Barlow, 2002; Lickona, 1997; Nodding, 1984, 2005).

Model Pendidikan dalam Budaya Indonesia

Di Indonesia, setiap warga negara wajib mengikuti wajib belajar sembilan tahun, yang dibagi menjadi enam tahun SD dan tiga tahun SMP. Sistem pendidikan nasional berfungsi untuk mengembangkan dan membentuk karakter moral dan martabat dalam rangka untuk mendidik masyarakat Indonesia, yang bertujuan untuk mengembangkan potensi peserta didik siswa potensi untuk menjadi setia dan adil, mulia, sehat, berilmu, cakap, kreatif, mandiri, sehingga menjadi

warga negara yang demokratis dan bertanggung jawab (UU Sistem Pendidikan Indonesia No 20/2003).

Sebuah studi longitudinal oleh Hofstede (1983) menggambarkan model budaya masyarakat Indonesia dipengaruhi oleh hubungan jarak kekuasaan yang tinggi dalam nilai-nilai kolektivisme. Biasanya, hubungan jarak kekuasaan yang tinggi terjadi dalam budaya yang memiliki nilai-nilai feodal. Adanya hubungan hirarkis antara individu yang berstatus tinggi dan rendah dianggap sebagai hubungan alami. Oleh karena itu, tingkat hirarki dalam derajat agama, akademik, kekuasaan, dan status sosial sangat menentukan apakah individu tersebut memiliki status yang tinggi atau rendah dalam masyarakat Indonesia (Lubis, 2001). Secara alami, di bawah kekuatan individu dengan status yang tinggi, orang lain dengan status yang lebih rendah akan mematuhi dan mendedikasikan hidup mereka untuk mereka yang statusnya lebih tinggi. Umumnya, dilakukan dengan rasa hormat, menghargai juga takut.

Idealnya, adanya saling ketergantungan dari hubungan antara superior dan inferior mencerminkan harapan budaya untuk menjaga dan melindungi satu sama lain (Koentjaraningrat, 1985). Mereka yang memiliki kekuatan superior memiliki tanggung jawab besar untuk berperan sebagai model yang baik untuk memberikan bimbingan moral dan kebijaksanaan, sehingga mereka yang memiliki kekuatan lebih rendah dapat mengikuti apa yang baik dan meninggalkan apa yang buruk (Koentjaraningrat & Schwartz, 2002).

Dalam konteks penelitian ini, contoh hubungan jarak kekuasaan antara posisi superior dan inferior berasal dari orang tua dan anak di rumah, guru dan siswa di sekolah, atau atasan dan bawahan di tempat kerja. Di Indonesia, di mana lingkungan sekolah dipengaruhi oleh hubungan jarak kekuasaan yang tinggi, para siswa yang berada di posisi lebih inferior akan memiliki kecenderungan untuk mengikuti tuntutan guru dan perilaku mereka karena guru berada dalam posisi superior. Budaya Indonesia mengakui bahwa guru memiliki peran tradisional untuk mencerminkan peran orang tua dalam lingkungan sekolah. Untuk memahami budaya kolektif di Indonesia, seorang anak akan belajar untuk berpikir identitas sebagai "kita" bukan identitas sebagai "aku". Orang-orang pada umumnya tidak terbiasa untuk memiliki pendapat yang berbeda dari komunitas mereka sendiri demi menjaga harmonisasi. Kompromi dan penyesuaian aspirasi lebih penting daripada berdebat dengan orang lain atas pendapat pribadi (Koentjaraningrat, 2004).

Mengungkap kesalahan yang dibuat oleh seseorang meskipun tujuannya dianggap baik, tetapi juga dianggap sebagai serangan pribadi bila dilakukan di depan umum. Oleh karena itu, kebanyakan orang Indonesia telah belajar bahwa menjaga perasaan orang lain lebih penting daripada mengatakan hal yang sebenarnya jika itu menyakiti orang lain. Ada nilai yang sangat kuat tidak menyakiti orang lain, karena biasanya menyebabkan reaksi negatif, khususnya dalam budaya Jawa, yang dianggap sebagai kelompok etnis yang paling kuat di Indonesia. Mereka tidak suka berbicara lugas; bahkan ada kecenderungan bahwa mereka suka berbohong untuk melindungi perasaan orang lain. Pastor Van Lith, seorang misionaris Katolik yang juga dikenal sebagai pakar bahasa Jawa dan filsafat, membuat interpretasi persepsi ini dan menyatakan, *"Orang-orang Barat tidak bisa memahami sikap Jawa dalam hubungan sosial. Dalam*

masyarakat Barat, anak-anak dididik 'jangan berbohong'. Sebaliknya, anak-anak Jawa dikondisikan untuk menumbuhkan sikap 'jangan menyakiti perasaan orang lain', " (Sumantri & Suharnomo, 2001, hal. 21).

METODOLOGI

Penelitian ini menggunakan pendekatan kuantitatif, data didapatkan dengan menggunakan kuesioner yang dikembangkan oleh peneliti dengan responden adalah siswa-siswa SMA pada kelas 12 di Kota Bandung. Analisis data pada untuk pengujian instrument menggunakan perangkat lunak (*software*) analisis model Rasch yaitu Winsteps versi 3.73.

Instrumen Riset

Instrumen yang digunakan adalah kuesioner dengan dua belas konstruk dengan masing-masing konstruk terdiri dari dua buah aitem. Dua belas konstruk yang digunakan merupakan karakter moral guru yang dinilai muridnya yang meliputi integritas, kejujuran, kepercayaan, keadilan, rasa hormat, kesetiaan, tidak mementingkan diri sendiri, kasih sayang, penghargaan spiritual, kerjasama, peduli, dan tanggung jawab (Barlow, 2002; Lickona, 1997; Nodding, 1984, 2005). Responden memberikan jawaban pada setiap aitem dalam bentuk *semantic differential* dengan diberikan tujuh pilihan jawaban. Jawaban responden dilihat dari kecenderungan apakah jawaban responden cenderung bergerak ke kolom yang paling kiri atau ke kolom yang paling kanan, yang mempertentangkan kualitas karakter moral yang baik atau buruk di setiap aitem.

Responden Penelitian

Sampel untuk penelitian ini dipilih secara acak, yaitu siswa perempuan (39%) dan lelaki (61%) kelas 12 SMA sebanyak 228 orang di Kota Bandung, Jawa Barat. Siswa yang terpilih berasal dari enam buah sekolah, yaitu tiga sekolah negeri dan tiga sekolah swasta. Semua siswa berpartisipasi secara sukarela dalam pengisian kuesioner dan sebelumnya mendapat penjelasan tentang tujuan riset yang diberikan oleh pihak sekolah.

Analisis model Rasch

Georg Rasch mengembangkan satu model analisis dari teori respon butir (atau *Item Response Theory*, IRT) pada tahun 1960-an, yang kemudian dipopulerkan oleh Ben Wright (Linacre, 2011). Dengan data mentah berupa data dikotomi (berupa benar dan salah) yang mengindikasikan kemampuan siswa, Rasch memformulasikan hal ini menjadi satu model yang menghubungkan antara siswa dan aitem (Sumintono & Widhiarso, 2013).

Sebagai ilustrasi, seorang siswa yang mampu mengerjakan 80% soal dengan benar tentu mempunyai abilitas yang lebih baik dari siswa lain yang hanya bisa mengerjakan 65% soal. Hal ini menunjukkan bahwa data mentah yang diperoleh adalah jenis data ordinal yang menunjukkan peringkat dan tidak linier. Oleh karena data ordinal tidak mempunyai interval yang sama, maka data tersebut perlu diubah menjadi data rasio untuk keperluan analisis statistik. Sehingga bila seseorang mendapat skor 80%, maka kemungkinan suksesnya adalah 80:20, yang tidak lain adalah data rasio yang lebih tepat untuk tujuan pengukuran. Melalui data rasio ini Rasch

mengembangkan model pengukuran yang menentukan hubungan antara tingkat kemampuan siswa (*person ability*) dan tingkat kesulitan aitem (*item difficulty*), dimana bisa disimpulkan bahwa tingkat kesuksesan siswa dengan tingkat abilitas yang tinggi akan mampu mengerjakan soal dengan tingkat kesulitan yang lebih rendah (Bond & Fox, 2007).

Selain data dikotomi, model Rasch juga bisa melakukan analisis untuk data politomi seperti yang dikembangkan oleh Andrich, yang tetap berlandaskan pada dua teorema dasar: tingkat kemampuan seseorang dan tingkat kesulitan aitem. Model Rasch berasumsi bahwa kesulitan aitem adalah sifat yang dipengaruhi oleh jawaban responden, dan kemampuan seseorang adalah sifat yang dipengaruhi oleh estimasi kesulitan aitem (Linacre, 1999).

Analisis dengan model Rasch menghasilkan analisis statistik kesesuaian (*fit statistics*) yang memberikan informasi pada peneliti apakah data yang didapatkan memang secara ideal menggambarkan bahwa orang yang mempunyai abilitas tinggi memberikan pola jawaban terhadap aitem sesuai dengan tingkat kesulitannya. Parameter yang digunakan adalah *infit* dan *outfit* dari kuadrat tengah (*mean square*) dan nilai terstandarkan (*standardized values*). Menurut Sumintono dan Widhiarso (2013), *infit* (*inlier sensitive* atau *information weighted fit*) adalah kesensitifan pola respon terhadap aitem sasaran pada responden (*person*) atau sebaliknya; sedangkan *outfit* (*outlier sensitive fit*) mengukur kesensitifan pola respon terhadap aitem dengan tingkat kesulitan tertentu pada responden atau sebaliknya.

Dalam analisis pada tingkat instrument dengan perangkat lunak Winsteps (Linacre, 2011), bila data sesuai dengan model Rasch, maka nilai jumlah kuadrat tengah (*mean square*) adalah 1,0 sedangkan nilai terstandarkannya (*Z-standardized values*) adalah 0,0. Sedangkan pada tingkat masing-masing aitem atau responden, maka parameter yang menunjukkan kesesuaian atau tidak sesuainya, ada tiga kriteria yang harus dipenuhi, yaitu :

Point Measure Correlation (x) : $0.32 < x < 0.8$

Outfit Mean Square (y) : $0.5 < y < 1.5$

Outfit Z standard (z) : $-2.0 < z < +2.0$

Dalam konteks pengujian aitem ini, maka suatu aitem yang tidak sesuai (*misfit*) adalah aitem yang terlalu mudah (nilai logit terlalu negatif) ataupun sangat sulit (nilai logit positif yang besar) dari pola jawaban responden yang diberikan; ataupun nilai dari tiga kriteria yang dihasilkan dari analisis dengan perangkat lunak menunjukkan bahwa aitem tidak memenuhi syarat, yang mengindikasikan aitem tersebut tidak mengukur ciri-sifat yang diinginkan (Sumintono & Widhiarso, 2013).

HASIL DAN PEMBAHASAN

Analisis yang dilakukan pada data yang dikumpulkan terbagi dalam ringkasan statistik yang memberikan informasi mengenai reliabilitas instrument, kemudian validitas person dan aitem.

Reliabilitas di tingkat instrumen: responden dan aitem

Pada Tabel 1 menampilkan ringkasan statistik dari analisis model Rasch, data yang diperoleh dari 228 responden yang menjawab 24 aitem pada instrument yang diberikan yang

merupakan hasil pengolahan dari perangkat lunak Winsteps. Dengan total jumlah data yang diberikan sebanyak 5472 menghasilkan nilai Chi-kuadrat 16474 dengan derajat bebas (df) 5216 ($p = 0,000$) yang menunjukkan keseluruhan pengukuran sangat bagus dan signifikan hasilnya. Untuk tabel yang mengukur pola jawab responden didapati bahwa nilai Infit MNSQ dan Outfit MNSQ yang ekspektasinya adalah 1.0 ; terlihat untuk person nilai Infit MNSQ adalah 1.11 dan Outfit MNSQ 1.08 hal ini menunjukkan bahwa secara keseluruhan pola jawaban respon pada instrument adalah bagus. Sedangkan untuk nilai Infit ZSTD dan Outfit ZSTD ekspektasinya adalah 0.0 ; untuk person nilai Infit ZSTD adalah -0.1 dan Outfit ZSTD -0.1 hal ini juga menunjukkan bahwa secara keseluruhan pola jawaban responden punya kesesuaian dengan model. Reliabilitas responden secara keseluruhan juga bagus, yaitu 0,86.

Table 1. Ringkasan Statistik Instrumen: Responden dan Aitem

SUMMARY OF 228 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	121.1	24.0	.55	.17	1.11	-.1	1.08	-.1
S.D.	20.3	.0	.55	.04	.77	2.0	.80	2.0
MAX.	162.0	24.0	2.45	.39	5.52	5.5	8.21	6.9
MIN.	64.0	24.0	-.72	.14	.17	-4.7	.18	-4.6
REAL RMSE	.21	TRUE SD	.51	SEPARATION	2.43	Person	RELIABILITY	.86
MODEL RMSE	.17	TRUE SD	.52	SEPARATION	3.03	Person	RELIABILITY	.90
S.E. OF Person MEAN = .04								
Person RAW SCORE-TO-MEASURE CORRELATION = .98								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .90								
SUMMARY OF 24 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	1150.9	228.0	.00	.05	1.04	-.1	1.08	.0
S.D.	155.8	.0	.40	.00	.41	3.6	.56	4.0
MAX.	1373.0	228.0	1.17	.07	2.48	9.9	3.13	9.9
MIN.	661.0	228.0	-.69	.05	.65	-4.5	.66	-4.4
REAL RMSE	.06	TRUE SD	.40	SEPARATION	6.90	Item	RELIABILITY	.98
MODEL RMSE	.05	TRUE SD	.40	SEPARATION	7.47	Item	RELIABILITY	.98
S.E. OF Item MEAN = .08								
UMEAN=.0000 USCALE=1.0000								
Item RAW SCORE-TO-MEASURE CORRELATION = -.99								
5472 DATA POINTS. LOG-LIKELIHOOD CHI-SQUARE: 16474.05 with 5216 d.f. p=.0000								
Global Root-Mean-Square Residual (excluding extreme scores): 1.2843								

Untuk pengujian instrument pada aitemnya, terlihat bawah untuk item nilai Infit MNSQ adalah 1.04 dan Outfit MNSQ 1.08 (dari nilai ekspektasinya 1,0); sedangkan untuk ZSTD nilai Infit ZSTD adalah -0.1 dan Outfit ZSTD 0.0 (dari nilai ekspektasinya 0,0). Kedua hal ini menunjukkan bahwa keseluruhan instrument adalah bagus, apalagi diperkuat dengan nilai reliabilitas instrument yang 0,98.

Nilai alpha cronbach (KR-20) yang mengukur interaksi antara responden dan aitem pun menunjukkan nilai reliabilitas yang bagus yaitu 0,90. Secara keseluruhannya hal ini menunjukkan bahwa data aktual yang diperoleh dalam riset ini sesuai dengan baik pada persyaratan model Rasch, sehingga analisis lebih lanjut layak untuk dilakukan.

Validitas

Konsep validitas sangat penting dalam pengukuran, validitas instrument adalah seberapa jauh pengukuran oleh instrument dapat mengukur atribut apa yang seharusnya diukur. Dalam konteks analisis model Rasch interpretasi pengukuran khususnya validitas konstruk dan isi dapat diinvestigasi secara lebih tepat. Selain itu model Rasch juga dapat mengukur validitas responden, dimana bila didapati pola jawaban responden yang tidak konsisten dapat terdeteksi yang menunjukkan tingkat kesahannya.

Validitas responden

Gambar 1 menunjukkan sebaran pola jawaban responden (sebelah kiri) dan tingkat persetujuan aitem (sisi kanan). Bagian sebelah kanan menunjukkan pengelompokan responden berdasarkan asal sekolahnya, dimana kode S menunjukkan sekolah negeri (dilambangkan dengan huruf A, C dan D) dan sisanya adalah sekolah swasta yaitu sekolah DT, MB dan SP (dilambangkan dengan huruf B, E dan F). Pola respon siswa berdasar sekolah menunjukkan kecenderungan menarik, dimana siswa dari tiga sekolah negeri yang menjadi responden pada penelitian ini menunjukkan pola yang beragam dari yang mudah setuju (nilai logit positif besar, bagian atas peta di bagian kanan) ke siswa yang sulit menyetujui (nilai logit paling negatif, berada di bagian bawah peta di sebelah kanan); sedangkan pada tiga sekolah swasta menunjukkan pola yang berbeda beda, siswa dari sekolah DT menunjukkan pola jawaban yang cenderung tidak setuju; sekolah MB mayoritas memilih posisi tengah; sedangkan siswa pada sekolah SP cenderung untuk menjawab mudah setuju terhadap aitem-aitem yang mengukur karakter moral guru. Hal ini mengindikasikan bahwa siswa pada tiga sekolah negeri yang menjadi responden studi ini mempunyai sebaran yang beragam, sedangkan siswa sekolah swasta cenderung mempunyai pola respon yang tidak begitu beragam.

Seperti terdapat pada Tabel 1, rata-rata logit responden adalah +0,55 logit yang menunjukkan keseluruhan responden punya kecenderungan untuk lebih menyetujui pernyataan tentang karakter moral gurunya. Dengan nilai separation 2,43 (Tabel 1) maka strata responden pada penelitian ini dengan menggunakan formula person strata yaitu H, sehingga nilai $H = [(4 * \text{separation}) + 1] / 3$ (Nazlinda & Beh, 2013), maka $H = [(4 * 2.43) + 1] / 3$, $H = 3,57$. Nilai person strata (H), sebesar 3 menunjukkan bahwa kelompok responden terbagi dalam tiga kelompok besar yaitu kelompok siswa yang menyetujui bahwa guru sebagai model karakter moral, kelompok kedua berada yang berada di tengah (dekat dengan rata-rata logit) yang cenderung melihat guru ada di posisi menengah; dan kelompok ketiga yang menilai guru bukanlah model karakter moral mereka.

Validitas Aitem

Seperti yang terlihat pada Tabel 1, nilai logit rata-rata aitem adalah 0,0 logit yang menunjukkan bahwa instrument secara keseluruhan bisa mengukur. Nilai rata-rata aitem 0,0 logit adalah nilai acak yang ditetapkan untuk menyatakan kemungkinan 50:50 yang tidak lain adalah ukuran sama antara tingkat abilitas responden dan tingkat kesulitan soal (Bond & Fox, 2007). Bila didapati bahwa rata-rata logit aitem tidak 0,0 maka secara keseluruhan instrument tidak bagus.

Pada Gambar 1 terlihat bahwa semua aitem yang terdapat di sebelah kanan, mulai dari yang mempunyai nilai logit negatif paling kecil (Sprt) sampai ke yang logit positif terbesar (Hon2) berada dalam nilai logit responden di bagian sebelah kiri peta. Hal ini menunjukkan bahwa keseluruhan aitem memang mengukur opini siswa pada model moral karakter guru.

Table 2. Kesesuaian Aitem (*item misfit*)

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR.	EXP.	OBS%	EXP%	Item
4	661	228	1.17	.05	2.48	9.9	3.13	9.9	A-.06	.60	17.1	25.4	Hon2
2	1119	228	.11	.05	1.76	6.9	2.17	9.5	B .15	.56	24.6	28.0	Int2
19	1371	228	-.68	.07	1.52	4.1	1.27	2.2	C .53	.45	31.6	37.3	Sprt1
5	1114	228	.13	.05	1.32	3.3	1.42	4.1	D .40	.56	27.2	27.8	Loy1
3	1319	228	-.48	.06	1.08	.8	1.41	3.3	E .43	.48	37.3	33.2	Hon1
1	1194	228	-.08	.05	1.06	.6	1.37	3.4	F .45	.54	33.8	30.1	Int1
9	1059	228	.26	.05	1.24	2.5	1.26	2.7	G .52	.58	28.9	27.4	Comp1
18	1189	228	-.07	.05	1.26	2.6	1.18	1.7	H .58	.54	27.6	29.8	Resp2
20	1373	228	-.69	.07	1.19	1.6	1.11	.9	I .55	.45	35.1	37.8	Sprt2
14	801	228	.84	.05	1.15	1.8	1.15	1.6	J .48	.61	24.6	25.6	Res2
6	1057	228	.27	.05	.99	-.1	1.03	.3	K .53	.58	32.5	27.4	Loy2
12	1269	228	-.31	.06	.93	-.7	.84	-1.6	L .62	.51	40.4	32.3	Care2
17	1227	228	-.18	.05	.84	-1.7	.76	-2.5	l .68	.53	38.6	30.5	Resp1
10	1082	228	.21	.05	.83	-1.9	.80	-2.3	k .67	.57	33.8	27.8	Comp2
22	1197	228	-.09	.05	.82	-2.0	.82	-2.0	j .62	.54	35.5	29.9	Coop2
13	1237	228	-.21	.05	.77	-2.5	.72	-3.0	i .64	.52	37.3	30.5	Res1
15	1161	228	.01	.05	.77	-2.7	.77	-2.6	h .69	.55	34.2	29.0	Fair1
16	1205	228	-.12	.05	.72	-3.3	.70	-3.4	g .68	.53	41.2	30.4	Fair2
7	1139	228	.06	.05	.71	-3.5	.70	-3.6	f .67	.56	32.9	28.2	Self1
11	1163	228	.00	.05	.71	-3.5	.71	-3.4	e .67	.55	40.8	29.3	Care1
24	1193	228	-.08	.05	.70	-3.5	.68	-3.7	d .71	.54	40.4	30.1	Trust2
23	1244	228	-.23	.06	.70	-3.4	.68	-3.6	c .67	.52	42.5	30.9	Trust1
21	1229	228	-.18	.05	.69	-3.5	.66	-3.8	b .61	.52	40.4	30.7	Coop1
8	1019	228	.35	.05	.65	-4.5	.66	-4.4	a .65	.59	36.4	26.6	Self2
MEAN	1150.9	228.0	.00	.05	1.04	-.1	1.08	.0			33.9	29.8	
S.D.	155.8	.0	.40	.00	.41	3.6	.56	4.0			6.2	3.0	

Pada Tabel 2 menunjukkan urutan kesesuaian setiap aitem (*item misfit order*), yang memberikan informasi menarik. Terdapat satu aitem yang mempunyai nilai *Point Measure Correlation* negative yaitu Hon2 (aitem kejujuran-2), yang lainnya adalah positif; dengan nilai rata-rata kesalahan pengukuran yang sangat kecil +0,05 logit. Untuk mengetahui mana aitem yang tidak fit dapat diketahui dengan menjumlahkan nilai rata-rata infit kuadrat tengah (*mean Infit MNSQ*) yaitu 1.04 dengan deviasi standarnya (*Infit MNSQ S.D.*) yaitu 0.41, sehingga nilainya adalah $1.04 + 0.41 = 1.45$; sehingga nilai Infit MNSQ lebih besar dari 1.45 adalah indikasi aitem tidak

sesuai. Patokan lainnya adalah bila nilai $z_{std} > +/- 2$ (Mohd Saidfudin, tanta tahun). Oleh karena itu aitem 4 (Hon2), 2 (Int2) dan 19 (Sprt1), termasuk aitem yang tidak fit dengan model karena nilai Infit MNSQ-nya lebih besar dari jumlah 1.45 serta nilai Z-std yang lebih besar dari 2.

Kriteria lain yang bisa digunakan untuk menilai kesesuaian aitem adalah dengan parameter nilai *outfit MNSQ*, *outfit Z-std* serta *PT-Measure correlation*, yang mendapati hal yang sama untuk dua aitem yang *misfit*, yaitu aitem 4 dan 2 (ketiga syarat untuk ketiga parameter tidak memenuhi syarat); namun tidak untuk aitem 19 (hanya nilai z-std yang tidak memenuhi syarat).

Bila dilihat lebih lanjut terdapat dua buah aitem yang mempunyai nilai logit yang hampir sama yang keduanya berasal dari konstruk yang sama, yaitu aitem 19 (logit -0.68) dan aitem 20 (logit -0.69). Hal ini memberikan informasi bahwa kedua aitem menurut responden mengukur hal yang sama sehingga tidak perlu memunculkan sampai dua aitem.

Dari keseluruhan aitem pada instrument ‘persepsi siswa terhadap karakter moral guru’ ini tiga aitem yang berasal dari tiga konstruk harus diperbaiki ulang. Hal ini karena aitem yang ada tidak sesuai dengan model (*misfit*) yaitu pada konstruk kejujuran (aitem ke-4) dan satu aitem pada konstruk integritas (aitem ke-2); sedangkan untuk aitem ke-19 (kejujuran-1) memberikan pola respon yang sama dengan aitem yang lain pada konstruk kejujuran (aitem ke-20), sehingga juga perlu disusun ulang.

Unidimensionalitas

Unidimensionalitas adalah hal yang penting untuk mengetahui apakah ia mengukur apa yang seharusnya diukur, yang dalam hal ini adalah mengukur persepsi siswa terhadap karakter moral guru. Dalam hal ini analisis Rasch model menggunakan Analisis Komponen Utama (Principal Component Analysis, PCA) dari residual, yaitu mengukur sejauh mana keragaman dari instrument mengukur apa yang seharusnya diukur.

Table 3. Keragaman residu tersrandarkan
(*standardized residual variance* dalam unit Eigenvalue)

		-- Empirical --		Modeled
Total raw variance in observations	=	40.9	100.0%	100.0%
Raw variance explained by measures	=	16.9	41.3%	43.4%
Raw variance explained by persons	=	5.1	12.4%	13.1%
Raw Variance explained by items	=	11.8	28.8%	30.3%
Raw unexplained variance (total)	=	24.0	58.7%	56.6%
Unexplned variance in 1st contrast	=	2.6	6.4%	10.9%
Unexplned variance in 2nd contrast	=	2.3	5.6%	9.5%
Unexplned variance in 3rd contrast	=	2.0	4.8%	8.2%
Unexplned variance in 4th contrast	=	1.7	4.2%	7.1%
Unexplned variance in 5th contrast	=	1.5	3.7%	6.3%

Pada tabel di atas terlihat hasil pengukuran keragaman (*raw variance*) data adalah 41.3% tidak jauh beda dengan nilai ekpektasinya yaitu 43,4%. Hal ini menunjukkan persyaratan minimum unidimensionalitas 20% terpenuhi, namun juga secara bersamaan batas unidimensi Rasch pun

terpenuhi yaitu di atas 40% (Reckase, 1979; Linacre, 2011). Hal lain yang mendukung, keragaman yang tidak dapat dijelaskan (unexplained variance) semua di bawah 7% yang menunjukkan tingkat independensi aitem dalam instrument yang baik.

Keberfungsian Aitem Differensial (DIF)

Aitem maupun instrument pengukuran dapat bersifat bias karena adanya perbedaan dalam responden (misalnya karena perbedaan gender, etnisitas, latar belakang keluarga dll) dimana aitem tertentu akan lebih memihak pada satu jenis tertentu. Analisis model Rasch menampilkan hal ini dalam keberfungsian aitem differensial (*differential item functioning*, DIF). Pada Table 4 ditampilkan hasil analisis DIF yang menunjukkan beberapa aitem terjangkit bias, dimana bisa diketahui dengan nilai probabilitas di bawah 5% (0,05) (Sumintono & Widhiarso, 2013).

Table 4. Keberfungsian Aitem Differensial (DIF)

Person CLASSES	SUMMARY DIF			BETWEEN-CLASS		Item	
	CHI-SQUARE	D.F.	PROB.	MEAN-SQUARE	t=ZSTD	Number	Name
6	25.5086	5	.0001	.9549	.1385	1	Int1
6	57.4479	5	.0000	2.9579	2.2765	2	Int2
6	17.9371	5	.0030	.8333	-.0689	3	Hon1
6	45.6466	5	.0000	2.3615	1.7841	4	Hon2
6	30.4182	5	.0000	1.2938	.6361	5	Loy1
6	8.2386	5	.1433	.3284	-1.2601	6	Loy2
6	5.1861	5	.3933	.1650	-1.9309	7	Self1
6	5.5687	5	.3501	.1601	-1.9569	8	Self2
6	13.7160	5	.0175	.5545	-.6357	9	Comp1
6	26.4525	5	.0001	1.1732	.4703	10	Comp2
6	2.3886	5	.7931	.0738	-2.5430	11	Care1
6	4.6224	5	.4634	.1775	-1.8668	12	Care2
6	13.9077	5	.0162	.6008	-.5302	13	Res1
6	10.7811	5	.0558	.3633	-1.1478	14	Res2
6	13.8594	5	.0165	.5754	-.5872	15	Fair1
6	4.1474	5	.5281	.1018	-2.3176	16	Fair2
6	10.9691	5	.0519	.3793	-1.0989	17	Resp1
6	28.2860	5	.0000	1.2483	.5748	18	Resp2
6	55.3136	5	.0000	3.9711	2.9789	19	Sprt1
6	16.0959	5	.0066	.6964	-.3282	20	Sprt2
6	4.2431	5	.5147	.1339	-2.1056	21	Coop1
6	7.4144	5	.1913	.2652	-1.4852	22	Coop2
6	12.1349	5	.0329	.3708	-1.1247	23	Trust1
6	13.2307	5	.0212	.5185	-.7220	24	Trust2

Pada table di atas terlihat bahwa aitem-aitem yang tidak terjangkit bias adalah aitem 6, 7, 8, 11, 12, 14, 16, 17, 21 dan 22. Banyaknya aitem yang terjangkit bias, hal ini menunjukkan bahwa penilaian persepsi siswa terhadap karakter moral guru dipengaruhi oleh bergagai hal yang dalam tahapan ini belum dipastikan, diantaranya bisa berasal dari jenis sekolah (negeri dan swasta), jenis kelamin, etnisitas dan latar belakang pendidikan orang tua.

Validitas Skala Peringkat

Validitas skala peringkat adalah hal yang sangat penting dalam berbagai system pengukuran. Sehingga validitas skala sangat menentukan secara keseluruhan pengukuran yang dilakukan. Analisis pemodelan Rasch memberikan proses verifikasi yang unik kepada asumsi peringkat

yang diberikan dalam instrument. Dalam studi ini, tujuh pilihan jawaban dalam bentuk *semantic differential* untuk setiap aitem.

Responden memberikan jawaban pada setiap aitem dengan diberikan. Jawaban responden dilihat dari kecenderungan apakah jawaban responden cenderung bergerak ke kolom yang paling kiri atau ke kolom yang paling kanan, yang mempertentangkan kualitas karakter moral guru yang baik atau buruk di setiap aitem. Pada Tabel 5 terlihat bahwa rata-rata observasi dimulai dari logit -0.22 untuk pilihan 1 (paling tidak setuju), meningkat sampai ke logit +1.17 (pilihan 7, yaitu sangat setuju). Namun, untuk pilihan peringkat 3, didapati nilainya adalah logit -0,03 yang menunjukkan kenaikan yang tidak konsisten; hal ini menunjukkan pilihan jawaban 1-7, bisa lebih disederhanakan misalnya dengan digabungkan menjadi 1-6 atau 1-5.

Tabel 5. Validitas Skala Peringkat

CATEGORY LABEL	SCORE	OBSERVED COUNT	OBSERVED %	OBSVD AVRGE	SAMPLE EXPECT	INFIIT MNSQ	OUTFIT MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE
1	1	244	4	-.22	-.50	1.48	1.96	NONE	(-2.09) 1 (tidak setuju)
2	2	285	5	-.23*	-.22	1.02	1.03	-.52	-1.05 2
3	3	341	6	-.03	.03	.87	.85	-.28	-.51 3
4	4	1082	20	.25	.28	.92	.91	-1.00	-.07 4
5	5	968	18	.47	.52	.92	.84	.51	.41 5
6	6	1247	23	.74	.78	1.09	1.13	.40	1.08 6
7	7	1305	24	1.17	1.08	.94	.97	.88	(2.34) 7 (setuju)

Ukuran lain untuk mengetahui ukuran validitas peringkat adalah apa yang disebut *Rasch-Andrich threshold*, yang menunjukkan transisi yang terjadi pada pengambilan keputusan oleh responden dari satu peringkat ke peringkat berikutnya (Linacre, 2011). Seperti halnya pada rata-rata pengamatan, *Rasch-Andrich threshold* tidak menunjukkan peningkatan yang konsisten dari peringkat 3 ke 4 (dalam skala tidak setuju [1] ke setuju [7]).

KESIMPULAN

Makalah ini menyajikan riset tentang pengembangan dan validasi instrument persepsi siswa terhadap karakter moral guru. Dua belas konstruk berdasar kajian pustaka tentang kualitas moral guru diidentifikasi, dimana tiap konstruk terdiri dari dua aitem disusun dalam instrumen riset ini dengan memberikan tujuh pilihan jawaban dalam bentuk *semantic differential*. Sebanyak 228 siswa kelas 12 SMA dari enam buah sekolah di Kota Bandung (masing-masing tiga sekolah negeri dan swasta) dipilih secara acak mengisi instrument riset secara sukarela.

Hasil analisis dengan pemodelan Rasch menunjukkan pada tingkat instrument, keseluruhan aitem dan responden menunjukkan tingkat reliabilitas yang bagus. Pada tingkat responden, terlihat bahwa pola respon untuk sekolah negeri cenderung sama dengan varian dari sangat setuju menuju ke tidak setuju terhadap karakter moral gurunya; namun pada tiga sekolah swasta terdapat pola yang berbeda-beda, satu sekolah cenderung tidak setuju, satu sekolah lain berada di opsi tengah dan sekolah terakhir cenderung setuju dengan karakter moral gurunya.

Hasil pengujian aitem menunjukkan terdapat dua buah aitem (kejujuran dan integritas) yang *misfit* dan dua buah aitem dari kostruk sama (penghargaan spiritual) yang mempunyai nilai logit yang sama; sehingga dua aitem terdahulu perlu disusun ulang, sedangkan aitem dengan logit yang sama perlu membuat aitem yang baru.

Instrumen riset pada penelitian ini menunjukkan bahwa ia mengukur apa yang seharusnya diukur, yaitu memenuhi syarat unidimensionalitas. Beberapa aitem juga terjangkiti bias, hal ini menunjukkan hal yang wajar dilihat dari keragaman responden dari segi asal sekolah (negeri dan swasta), jenis kelamin, etnisitas dan latar belakang pendidikan orang tua. Dari segi validitas skala peringkat didapati bahwa memberikan pilihan 1 sampai 7 menyebabkan responden tidak bisa secara tepat opini tentang karakter moral guru mereka, sehingga penyederhanan diperlukan untuk menjadi pilihan 1-5 atau 1-6.

DAFTAR PUSTAKA

- Bandura, A. (1963). The role of imitation in personality. *The Journal of Nursery Education*, 18 (3).
- Bandura, A. (1971). *Social learning theory*. New York: General Learning Press.
- Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1995). *Social learning*. In A. S. R. Manstead & M. Hewstone (Eds.), *Blackwell encyclopedia of social psychology* (pp. 600-606). Oxford: Blackwell.
- Bandura, A. (2002). Social cognitive theory in cultural context. *Applied Psychology: An International Review*, 151, 269-290.
- Bandura, A. (2003). Observational learning. In J. H. Byrne (Eds.), *Encyclopedia of learning and memory*. (2nd ed., pp. 482-484). New York: Macmillan.
- Bandura, A. (2004). Modeling. In E. W. Craighead & C. B. Nemeroff (Eds.), *Encyclopedia of Psychology and Behavioral Sciences* (pp. 575-577). New York: Wiley.
- Bandura, A., Ross, D., & Ross, S. A. (1963). Vicarious reinforcement and imitative learning. *Journal of Abnormal and Social Psychology*, 67, 601-607.
- Barlow, C. B. (2002). *Character assessment across Air Force professional military education: A descriptive investigation*. (Res. Rep. No. AU/ACSC/015/2002-4). Air University, Maxwell AFB, Alabama.
- Berkowitz, M. W. (2002). The science of character. In W. Damon (Eds.), *Bringing in a new era in character education* (pp. 43-63). Stanford, CA: Hoover Institute Press.
- Blasi, A. (1980). Bridging moral cognition and moral action: A critical review of the literature. *Psychological Bulletin*, 88, 1-45. Boston: Kluwer.

- Bond, T.G., & Fox, C. (2007). *Applying the Rasch Model. Fundamental measurement in the Human Sciences*. Lawrence Erlbaum Associates, Publishers. Mahwah. New Jersey
- Hofstede, G. (1983). National culture in four dimensions: A research-based theory of cultural differences among nations. *Journal of International Studies of Management & Organization*, XIII (1-2), 46-74.
- Holyoak, K.J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13, 295-355.
- Jibi. (2012, April, 19). The complained number of the National Exam reached 752 cases. Retrieved from <http://www.regionaltimur.com/index.php/jumlah-pengaduan-un-mencapai-752-kasus>.
- Koentjaraningrat, R. (2004). *People and culture in Indonesia*. Jakarta, Indonesia: Djambatan Publisher.
- Koentjaraningrat, R. (1985). *Javanese culture*. London: Oxford University Press.
- Kohlberg, L. (1981). *The philosophy of moral development*. New York, NY: Harper and Row.
- Kohlberg, L. (1981). *The meaning and measurement of moral development*. Worcester, MS: Clark University Press.
- Kohlberg, L. (1984). *The psychology of moral development: The nature and validity of moral stages*. New York, NY: Harper and Row.
- Lickona, T. (1991). *Educating for character: How our schools can teach respect and responsibility*. New York: Bantam.
- Lickona, T. (1997). *Educating for character: A comprehensive approach*. In A. Molinar, (Ed.), *The construction of children's character* (pp.45-62). Chicago, IL: The National Society for the Study of Education.
- Lickona, T. & Davidson, M. (2005). *Smart and good high school: Integrating excellence and ethics for success in school, work, and beyond*. Cortland, NY: Center for the 4th and 5th RS: Respect and Responsibility. Washington DC: Character Education Partnership.
- Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement*. 3(2), 103-122.
- Linacre, J.M. (2011). *A User's guide to WINSTEPS Ministeps; Rasch-model Computer Program. Program Manual 3.73*.

- Lubis, M. (2001). *Indonesian People*. (5th ed.). Jakarta, Indonesia: Yayasan Obor Indonesia.
- Lumpkin, A. (2008). Teachers as role models, teaching character and moral virtues. *Journal of Jouperd*, 79 (2).
- Misbach, I. H. (2013). The Pilot Study of Students' Perception on Teachers' Moral Character Scale in Indonesia. Tesis MA (S2) tidak dipublikasikan pada University of Connecticut.
- Mohd Saidfudin Masodi. (tanpa tahun). Instrument Construct Validation of VFMA_35 for the conduct of Value for Money Audit using Rasch Model. Makalah tidak diterbitkan, Accounting Research Institute, UiTM, Syah Alam, Malasia.
- Nazlinda Abdullah dan Beh Kian Lim. (2013). Parallel Circuit Conceptual Understanding Test (PCCUT). *Procedia - Social and Behavioral Sciences* 90 (2013) 431 – 440
- Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education*. Berkeley: University of California Press.
- Noddings, N. (2005). Identifying and responding to needs in teacher education. *Journal of Education*, 35 (2).
- Pratiwi, L., & Djumena, E. (2011, January 23). This is the United Federation of Teachers Indonesia. Retrieved from <http://edukasi.kompas.com/read /2011/01 /23/17094173/>
- Reckase, M.D. (1979). *Unifactor Latent Trait Models Applied to Multifactor Tests: Results and Implications*, *Journal of Educational Statistics*, Vol. 4, No. 3, pp. 207-230
- Schwartz, A. (2000). *A nation in waiting: Indonesia's search for stability*. USA: Westview Press.
- Sumantri, S., & Suharnomo (2007). Kajian proposisi hubungan antara dimensi budaya nasional dengan motivasi dalam suatu organisasi usaha. Retrieved on December, 11, 2011 from <http://www.pustaka.unpad.ac.id>
- Sumintono, B dan Widhiarso, W. (2013). *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial*. Cimahi: Trim Komunikata Publishing House.
- Thomas, R. M. (1962). The prestige of teachers in Indonesia. *Journal of Education*, 40(3), 150-156. Santa Barbara, CA: Taylor & Francis, Ltd.
- Timpe, K. (2007). Moral character. *Online Research of Philosophy*. Retrieved 20 February 2012, from <http://philpapers.org/rec/TIMMC>.
- Vessels, G. (1998). *Character education*. In J. Guthrie (Ed.), *The encyclopedia of education* (2nd ed.), New York: Macmillan.