



## UJI RELIABILITAS DAN VALIDITAS SKALA KARAKTER BAIK (SKB) UNTUK SISWA SMA/SEDERAJAT DI INDONESIA

Aviva Lutfiana<sup>1</sup>, Hasna Hafizhah Salma<sup>2</sup>, Gary Collins Brata Winardy<sup>3</sup>, dan Ali Nina Liche Seniati<sup>4</sup>

Fakultas Psikologi, Universitas Indonesia, Depok<sup>1,2,3,4</sup>

*E-mail:* [avivalutfiana21@gmail.com](mailto:avivalutfiana21@gmail.com)

### Abstract

Good character was a congruence of thoughts, feelings, and behaviors by principles of morality. Since 2016, the Indonesian education system attempted to develop good character in students using principles for encouraging character education (PPK, Penguatan Pendidikan Karakter) based on Pancasila, Indonesia's national ideology. These five principles are religiosity, nationalism, independence, cooperation, and integrity. Currently, there are no instruments for measuring good character by the principle of PPK. This study aimed to develop a good character measurement tool based on a good character construct (Lickona, 1991) consisting of three dimensions: moral knowing, moral feeling, and moral action to describe good character for high school students in Indonesia. This study tests the reliability with item discrimination analysis using SPSS v.21. Construct validity was measured by confirmatory factor analysis with JASP emulation Mplus. Group norm developed with an average of 10 and a standard deviation of 3. The subjects of this study are 728 Indonesian high school students. The result of this study showed that the instrument is reliable with Cronbach alpha of 0.892. This instrument also showed a model fit p-value of 0.000, SRMRS of 0.078, and GFI of 0.957, which can be interpreted as reliable in measurement.

*Keyword:* character education PPK, good character, high school/equivalent

### Abstrak

Karakter baik (*good character*) merupakan bentuk keselarasan dari pemikiran, sikap, dan tingkah laku yang sesuai moral. Sejak tahun 2016, Indonesia berupaya untuk membentuk karakter baik bagi setiap anak bangsa dengan menerapkan nilai-nilai Penguatan Pendidikan Karakter (PPK) yang bersumber dari Pancasila di dalam dunia pendidikan. Lima nilai prioritas PPK adalah religius, nasionalis, mandiri, gotong royong, dan integritas. Namun, sampai saat ini belum ditemukan pengukuran karakter yang baik dengan kontekstualisasi kelima nilai PPK tersebut. Oleh karena itu, peneliti mengembangkan alat ukur karakter baik berdasarkan konstruk *good character* (Lickona, 1991) yang terdiri dari tiga dimensi, yaitu: pengetahuan moral, sikap moral, dan tingkah laku moral sehingga dapat digunakan untuk menggambarkan karakter yang baik bagi siswa/i SMA/ sederajat di Indonesia. Pengembangan alat ukur ini dilakukan dengan menguji reliabilitas dan analisis diskriminasi item menggunakan SPSS Versi 21, menguji validitas konstruk melalui *confirmatory factor analysis* menggunakan JASP emulation MPlus, dan menyusun norma kelompok menggunakan skor standar dengan skor rata-rata = 10 dan deviasi standar = 3. Populasi dari pengujian alat ukur ini adalah siswa/i SMA sederajat di Indonesia dengan sebanyak 728 sampel. Hasilnya menunjukkan bahwa alat ukur Skala Karakter Baik adalah alat ukur yang reliabel (Cronbach alpha = 0,892) dan valid (Model Fit RMSEA p-value=0,000, SRMS=0,078, GFI=0,957).

Kata kunci: Karakter Baik, Penguatan Pendidikan Karakter (PPK), SMA/Sederajat

## 1. Pendahuluan

Pendidikan karakter merupakan salah satu konsep pendidikan yang hangat dibicarakan di seluruh penjuru dunia. Keberhasilan penerapan pendidikan karakter suatu negara dapat menciptakan individu yang memiliki pemahaman, keterlibatan, rasa peduli, dan bertingkah laku sesuai dengan nilai-nilai etika seperti rasa hormat, keadilan, kewarganegaraan, dan tanggung jawab untuk diri sendiri dan orang lain baik di dalam institusi pendidikan maupun komunitas yang lebih besar, bahkan dalam konteks bernegara (Pala, 2011; Dempster, 2020). Sejak tahun 2016, di Indonesia, Penguatan Pendidikan Karakter (PPK) menjadi salah satu butir Nawacita Presiden Joko Widodo melalui Gerakan Nasional Revolusi Mental. Pendidikan Nasional berfokus pada penguatan karakter, bukan hanya meningkatkan kompetensi. Kombinasi karakter yang kuat dan kompetensi yang tinggi melahirkan jati diri bangsa yang kokoh, kolaborasi, dan daya saing bangsa meningkat, hingga menjawab tantangan abad-21 (Kemendikbud, 2017). Kementerian Pendidikan dan Kebudayaan (Kemendikbud) menerjemahkan PPK menjadi

program pendidikan yang diimplementasikan secara masif di satuan pendidikan (sekolah/lembaga pendidikan) khususnya jenjang menengah (SMP dan SMA) di seluruh Indonesia.

Pendidikan karakter di Indonesia membutuhkan pembuktian keberhasilannya. Keberhasilan pendidikan karakter dilihat dari dua aspek, yaitu perubahan positif perilaku peserta didik secara individual dan perubahan positif pada budaya sekolah satuan pendidikan. Evaluasi pendidikan karakter melalui observasi (pengamatan langsung), analisis dokumen, wawancara maupun diskusi untuk mengumpulkan data, baik data-data administratif maupun catatan pendukung. Metode ini disesuaikan dengan jenis data yang diperoleh dan target dari evaluasi. Namun, sejauh ini evaluasi dilakukan secara kualitatif (Kemendikbud, 2017).

Konsep pendidikan karakter dalam konstruk psikologi pertama dikemukakan oleh Psikolog Pendidikan Lickona awal tahun 1900-an. Lickona (1991) menyebutkan pendidikan karakter diinternalisasi pada tiap individu sebagai karakter baik, yang secara konsisten ditampilkan melalui pengetahuan tentang kebaikan (pengetahuan moral) yang menjadi bentuk pemikiran, keinginan melakukan kebaikan (sikap moral) yang terinternalisasi menjadi sikap atau emosi, dan melakukan kebaikan (perilaku moral) yang dapat dilihat dalam kebiasaan sehari-hari. Lebih mendetail lagi Lickona menjelaskan ketiga dimensi tersebut dalam beberapa indikator, yakni (1) pengetahuan moral terdiri dari indikator kesadaran moral, pengetahuan tentang nilai moral, pengambilan sudut pandang (*perspective taking*), alasan moral, pengambilan keputusan, dan pengetahuan diri; (2) sikap moral terdiri dari indikator kesadaran (*conscience*), harga diri, empati, cinta kebaikan, kontrol diri, dan kerendahan hati (*humility*); serta (3) tingkah laku moral terdiri dari indikator kompetensi atau kemampuan, keinginan, dan kebiasaan. Adanya ketiga dimensi ini sangat mempengaruhi proses pengembangan alat ukur, yaitu peneliti memberikan kontekstualisasi atau situasi yang sama selaras dengan nilai-nilai prioritas Penguatan Pendidikan Karakter (PPK) di Indonesia pada setiap indikator yang saling berkaitan satu sama lain. Hal ini bertujuan untuk melihat keselarasan antara aspek pemikiran, sikap, dan tingkah laku dalam menunjukkan karakter baik.

Terdapat beberapa penelitian pendidikan karakter yang didasari oleh konsep Lickona dengan penerapan yang disesuaikan dengan kondisi institusi atau negara dimana penelitian dilakukan, seperti (1) Penelitian Akademi Militer di Taiwan berkenaan dengan perbandingan tiga model intervensi pendidikan karakter pada taruna baru (Yu, 2018); (2) Penelitian pengembangan program pendidikan karakter memberikan usulan kebijakan pada pendidikan sekretaris di Amerika Serikat (Elias, Nayman, & Duffell, 2018); (3) Penelitian metode studi literatur mengenai implikasi karakter pada pendidikan karakter menjelaskan lanjut tentang moral dan negarawan (Baehr, 2017); selain itu, (4) Berkowitz & Bier (2014) mengumpulkan penjelasan “Penelitian Berbasis Pendidikan Karakter” dengan pendekatan studi literatur.

Di Indonesia, penelitian pendidikan karakter telah dilakukan dengan berbagai metode, dari studi literatur, pengambilan data, hingga proses intervensi, diantaranya: (1) Penelitian yang dilakukan oleh Taqwiem (2017) menemukan bahwa pendidikan karakter mempunyai peran penting di Indonesia dengan budaya yang beragam; (2) Fatoni (2017) yang menemukan model pendidikan karakter di Indonesia bertujuan meningkatkan karakter perilaku peserta didik, dimana pemerintah dapat menekankan pada pendidikan karakter yang sistematis; (3) Penelitian Ismail et al. (2016) yang menemukan model pendidikan karakter efektif meningkatkan kemampuan akademik siswa dan kemandirian di Pondok Pesantren Sinjai. Peneliti menyimpulkan satu pola kelemahan dari berbagai penelitian tersebut, yakni tidak menampilkan kejelasan alat ukur yang digunakan. Kemudian, peneliti menemukan bukti uji validitas pengembangan alat ukur karakter yang menggunakan konsep Lickona oleh Layyinah & Roebianto (2020), tetapi dalam penelitian tersebut tidak dikaitkan dengan kontekstualisasi nilai PPK yang berlaku di Indonesia. Sementara, jika mengacu pada konsep yang digunakan Lickona (1991) bahwa pengukuran karakter baik harus disesuaikan dengan nilai-nilai yang dianut dalam kelompok/institusi/negara karena idealnya nilai-nilai baik yang diajarkan adalah nilai yang dibagikan dan dibentuk dari masyarakat di mana pendidikan karakter diterapkan (Lickona, 1991).

Oleh karena itu, peneliti hendak mengembangkan lebih lanjut alat ukur Skala Karakter Baik didasarkan pada konstruk *good character* (karakter baik) dari Lickona (1991) yang terdiri dari tiga dimensi: pengetahuan moral, sikap moral, dan tingkah laku moral, dengan mengkontekstualisasikan pada lima nilai prioritas penguatan pendidikan karakter (PPK) di Indonesia. Tujuan pembuatan alat ukur ini adalah membuat alat ukur yang reliabel dan valid sehingga dapat menggambarkan karakter baik pada siswa/i SMA/ sederajat.

## 2. Metode Penelitian

### Populasi dan Sampel

Populasi dalam pengembangan alat ukur ini adalah siswa/i SMA/ sederajat yang tinggal di Indonesia dengan teknik *accidental sampling*. Hal ini didasari dari pengambilan sampel yang dilakukan dengan prinsip kemudahan dalam mengakses populasi pengambilan sampel (Kumar, 2011). Responden yang diperoleh sejumlah 728 siswa. Hasil uji KMO (*Kaiser-Meyer-Olkin of Sampling Adequacy*) adalah sebesar 0,875 dan signifikansi Bartlett Test sebesar 0,00 ( $p < 0,05$ ). Menurut Field (2018), nilai yang mendekati 1,00 untuk KMO dan signifikan untuk Bartlett Test menunjukkan bahwa sampel memenuhi persyaratan signifikansi dan dapat diproses dalam analisis faktor. Oleh karena itu, analisis faktor dapat dilakukan.

## Uji Keterbacaan

Uji keterbacaan dilakukan dalam rentang waktu 25-26 Oktober 2020. Subjek dalam uji keterbacaan ini berjumlah 12 siswa/i SMA/ sederajat, yang berasal dari Malang, Surakarta, Depok, Kab. Bogor, Bandung, dan Jakarta. Pada uji keterbacaan, penguji melakukan wawancara dengan subjek secara daring melalui aplikasi Zoom untuk membahas umpan balik pada setiap itemnya. Wawancara dilakukan secara berkelompok dengan maksimal setiap sesinya berjumlah 3 orang siswa. Hal yang dilihat dari uji keterbacaan adalah kemudahan item untuk dipahami, baik dari segi bahasa maupun maknanya. Pada uji keterbacaan dieliminasi 3 item karena sulit dipahami dan ambigu. Akhirnya didapatkan 63 item dari awalnya 66 item untuk dilakukan analisis faktor.

## Proses Pengambilan Data

Proses pengambilan data berlangsung sejak tanggal 18 November 2020 hingga 04 Desember 2020 yang diadministrasikan secara daring melalui *google form*. Tanggapan yang terekam dalam *google form* berjumlah 769 data. Dari data tersebut, peneliti melakukan pembersihan data dengan membuang tanggapan “tidak bersedia” dan data ganda. Total data yang dibuang adalah 36 data, sehingga total data bersih sementara adalah 733 data. Setelah melakukan pembersihan data, peneliti melakukan uji normalitas data Kolmogorov-Smirnov dan mendapatkan hasil signifikansi  $p > 0.05$  (0.128). Perolehan hasil tersebut menunjukkan bahwa data berdistribusi normal. Kemudian langkah berikutnya adalah membuang 5 data yang berada di luar kurva distribusi normal sehingga jumlah data bersih yang digunakan dalam pengolahan dan perhitungan adalah 728 data.

## Prosedur Pengujian

Pada pengujian reliabilitas dalam pengembangan alat ukur Skala Karakter Baik, peneliti memilih metode *Cronbach alpha* dengan kriteria yang digunakan untuk menyatakan item reliabel adalah di atas 0,7 (Chadha, 2009). Pengukuran reliabilitas tersebut menggunakan SPSS Versi 25. Adapun untuk pengujian validitas, Skala Karakter Baik merupakan alat ukur multidimensi yang terdiri atas tiga dimensi, yaitu pengetahuan moral, sikap moral, dan tingkah laku moral. Dengan begitu, pengujian validitasnya dilakukan dengan metode CFA *second-order*, dengan langsung melihat keterkaitan antara konstruk *Good Character* kepada ketiga dimensi dan dari setiap dimensi ke dalam item-item. Kemudian, dalam untuk menentukan apakah model empirik fit dengan model teoritis, peneliti menggunakan indeks model-fit Chi square ( $p > 0.05$ ), RMSEA p-value ( $< 0.08$ ), CFI ( $> 0.90$ ), GFI ( $> 0.90$ ), dan SRMR ( $< 0.08$ ). Kriteria ini mengacu pada kriteria Hair, dkk. (2019), dimana model dianggap fit ketika minimal tiga nilai indeks model fit memenuhi kriteria. Pengukuran validitas menggunakan JASP versi 0.14.1 dengan *emulation MPlus*.

## Kisi-Kisi Skala Karakter Baik (SKB)

Peneliti membuat Skala Karakter Baik berlandaskan konsep teori Lickona (1991) yang terdiri dari tiga dimensi, yakni pengetahuan moral, sikap moral, dan tingkah laku moral. Selain itu, peneliti juga menyesuaikannya dengan penerapan pendidikan karakter di Indonesia melalui lima nilai prioritas Penguatan Pendidikan Karakter (PPK), yakni religius, nasionalis, mandiri, gotong royong, dan integritas.

Pada saat awal pembuatan Skala Karakter Baik, peneliti melakukan studi awal di lapangan dengan mewawancarai 8 orang siswa/i SMA peneliti sederajat. Wawancara ini bertujuan untuk mendapatkan gambaran tingkah laku subjek berkaitan dengan tiga dimensi karakter baik dan nilai-nilai PPK. Hasil wawancara tersebut menunjukkan bahwa terdapat indikator karakter baik yang peneliti tiadakan, yaitu indikator kesadaran moral pada dimensi pengetahuan moral. Indikator tersebut telah diwakili dalam dimensi sikap moral sehingga sulit memunculkan item spesifik mengarah ke indikator tersebut. Hal ini terjadi karena semua karakter baik pasti membutuhkan kesadaran akan moral dan juga penghayatan emosi dari kesadaran proses berpikirnya. Pada tabel 1 berikut ini dijelaskan *blueprint* alat ukur Skala Karakter Baik (SKB):

Tabel 1. *Blueprint* Skala Karakter Baik

Dimensi	Indikator	Contoh Item
Pengetahuan Moral ( <i>Moral Knowing</i> )	<i>Knowing Moral Values</i> (Subjek mengetahui nilai moral yang penting selama menjalani pendidikan)	Saya memilih untuk melakukan kegiatan ibadah ketika saya berada dalam kondisi bingung
	<i>Moral Reasoning</i> (Subjek memahami alasan harus bertingkah laku sesuai moral)	Membantu teman yang kesulitan dalam belajar adalah bagian dari tanggung jawab saya
	<i>Decision Making</i> (Subjek dapat mengetahui dampak dari keputusan yang dipilihnya)	Saya lebih memilih untuk remedial daripada mencontek saat ujian

	<i>Perspective Taking</i> (Subjek dapat memahami pandangan orang lain yang berbeda darinya)	Saya memahami pandangan hidup teman yang berbeda suku/budaya dengan saya
	<i>Self Knowledge</i> (Subjek dapat mengetahui kekuatan/kelemahan dirinya)	Saya dapat menuliskan apa saja kelebihan saya yang dapat membantu saya menggapai cita-cita
<b>Sikap Moral (Moral Feeling)</b>	<i>Self Esteem</i> (Subjek mampu menghargai dirinya)	Saya merasa dihargai orang lain saat saya berusaha taat beribadah
	<i>Empathy</i> (Subjek mampu merasakan sikap yang dirasakan oleh orang lain)	Saya ikut sedih saat teman saya dihina karena sukunya
	<i>Loving the Good</i> (Subjek mampu merasakan kenikmatan dalam melakukan hal yang baik)	Saya puas saat bisa berkontribusi menyelesaikan tugas saya sebagai anggota kelompok
	<i>Self-Control</i> (Subjek mampu mengesampingkan keinginannya sendiri)	Saya ingin mengingatkan guru saat beliau lupa tugas pada hari itu
	<i>Humility</i> (Subjek membuka diri pada kebenaran yang baru ia temui)	Saya membuka diri terhadap ide baru untuk menunjang proses mencapai tujuan
<b>Tingkah Laku Moral (Moral Action)</b>	<i>Competence</i> (Subjek mampu melakukan perilaku moral dengan mempertimbangkan pengetahuan yang dimiliki)	Saya mengajak teman saya untuk saling menghargai perbedaan suku/budaya
	<i>Will</i> (Subjek mampu menahan tekanan kelompok untuk melakukan perilaku moral)	Saya tetap memilih menjalankan ibadah tepat waktu meskipun teman saya membujuk untuk menundanya
	<i>Habit</i> (Subjek mampu mengembangkan kebiasaan dalam melakukan perilaku moral)	Saya membiasakan diri untuk melakukan apapun yang sudah saya katakan

### 3. Hasil

#### Uji Reliabilitas Tahap Awal

Sebelum dilakukan uji validitas, peneliti mencoba melihat uji reliabilitas pada 63 item awal. Didapatkan hasilnya sebagai berikut:

**Tabel 2. Hasil Uji Reliabilitas Tahap Awal**

Uji Reliabilitas	Item Awal (63 item)	
	Jumlah Item	Nilai <i>Cronbach Alpha</i>
Skala Karakter Baik	63	0,885
Pengetahuan Moral	23	0,703
Sikap Moral	17	0,708
Perilaku Moral	23	0,795

Berdasarkan nilai *Cronbach alpha* pada tabel 2, dapat diketahui bahwa dalam tahap awal ini baik alat ukur SKB secara keseluruhan dan juga setiap dimensinya (pengetahuan moral, sikap moral, dan tingkah laku moral) memenuhi kriteria reliabel.

#### Uji Validitas Tahap Awal

Uji validitas SKB dilakukan dengan *Confirmatory Factor Analysis (CFA) second order*, karena SKB merupakan alat ukur multidimensional yang terdiri dari tiga dimensi, yaitu pengetahuan moral, sikap moral, dan tingkah laku moral. Untuk alat ukur yang bersifat multidimensional, peneliti dapat langsung melihat interaksi *second order* dari konstruk ke dimensi dan dimensi ke item-item, sehingga tidak diperlukan lagi untuk menguji *first order* secara terpisah (Yudiana, 2020). Hasil pengujian indeks *model-fit* yang diperoleh sebagai berikut:

**Tabel 3. Hasil Model Index Fit (Second-Order) Tahap Awal**

<i>Model Index Fit</i>	<b>Kriteria (Hair et al., 2019)</b>	<b>Hasil</b>	<b>Keterangan</b>
<b>Chi-square <i>p-value</i></b>	>0,05	<0,01	Model Tidak Fit
<b>RMSEA <i>p-value</i></b>	<0,07	0,000	Model Fit
<b>SRMR</b>	<0,08	0,078	Model Fit
<b>GFI</b>	>0,90	0,956	Model Fit
<b>CFI</b>	>0,90	0,565	Model Tidak Fit

Hasil tabel 3 menyatakan bahwa model pengukuran Skala Karakter Baik (SKB) dengan *second-order* dinyatakan fit berdasarkan nilai indeks RMSEA *p-value*, SRMR, dan GFI. Hal ini menunjukkan bahwa Skala Karakter Baik dalam kontekstualisasi nilai-nilai PPK valid dalam mengukur konstruk karakter baik. Hanya saja jika dilihat dari nilai indeks *chi-square* dan CFI menunjukkan bahwa model tidak valid. Hal ini masih dapat ditoleransi dan model *second-order* tetap dianggap valid karena *chi-square* dan CFI bias terhadap jumlah sampel besar (Yudiana, 2020). Kemudian, untuk melihat lebih lanjut bagaimana signifikansi item dalam merepresentasikan dimensi dan kualitas dimensi dalam merepresentasikan konstruk/variabel laten, maka dilihat nilai muatan faktor (*factor loading*) pada dimensi dan itemnya (Yudiana, 2020). Berikut ini merupakan tabel 4 yang menjelaskan hasil *second-order factor loading*:

**Tabel 4. Factor Loading Tiga Dimensi Skala Karakter Baik (Second-Order Factor Loading)**

	<b>Dimensi</b>	<i>Std.</i> (semakin kecil semakin baik)	<i>Error p-value</i> (<0,05)	<i>Factor Loading</i> (≥0,3)
<b>Skala Karakter Baik</b>	<b>Pengetahuan Moral</b>	0,509	<0,01	0,944
	<b>Sikap Moral</b>	0,313	<0,01	0,921
	<b>Tingkah Laku Moral</b>	1,411	0,002	0,975

Berdasarkan tabel 4, dapat diketahui bahwa nilai *factor loading* tiap dimensinya lebih dari 0,9, artinya ketiga dimensi benar merepresentasikan variabel laten/konstruk *good character* (Hair et.al, 2019). Setelah melihat nilai *factor loading* dimensi ke variabel laten, langkah berikutnya adalah melihat *factor loading* item-item di dalam setiap dimensinya, sebagai berikut:

**Tabel 5. Factor Loading Item-item dalam Dimensi**

<b>Dimensi</b>	<i>p-value</i> (<0,05)	<i>Factor Loading</i> (≥0,3)	<b>Jumlah Item</b>	<b>Nomor Item</b>
<b>Pengetahuan Moral</b>	<0,05	≥0,3	13 item	1, 3, 10, 11, 13, 15, 16, 17,18, 19, 20, 21, 23
	<0,05	<0,3	10 item*	2, 4, 5, 6, 7, 8, 9, 12, 14, 22
<b>Sikap Moral</b>	<0,05	≥0,3	12 item	1, 4, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16
	<0,05	<0,3	4 item*	2, 3, 5, 12
	>0,05	<0,3	1 item	17

<b>Tingkah Moral</b>	<b>Laku</b>	<0,05	≥0,3	15 item	3, 4, 5, 8, 9, 10, 11, 12, 13, 14, 15, 19, 20, 22, 23
		<0,05	<0,3	8 item*	1, 2, 6, 7, 16, 17, 18, 21

\*Item perlu dianalisis lebih lanjut

Berdasarkan tabel 5, dapat diketahui bahwa berdasarkan nilai *factor loading*-nya, terdapat 22 item yang perlu dianalisis lebih lanjut karena nilai *factor loading* yang kurang dari 0,3 menandakan bahwa item-item tersebut tidak mengukur dimensi sehingga butuh dianalisis lebih lanjut (Hair et.al., 2019).

### Uji Reliabilitas Tahap Lanjut

Peneliti melakukan analisis lanjutan secara integratif dengan (1) melihat nilai *item discrimination* melalui *corrected item-total correlation* yang terlihat pada *output* reliabilitas SPSS; (2) nilai-nilai *factor loading* yang rendah; dan (3) analisis kualitatif terhadap keterwakilan indikator dan kondisi kontekstual nilai PPK pada setiap itemnya. Setelah menelaah lebih lanjut item mana saja yang dapat dieliminasi dan mana yang dapat dipertahankan, maka didapatkan 42 item terpilih untuk dilakukan pengujian ulang mulai dari reliabilitas, analisis item dengan *item discrimination*, validitas, dan melihat *factor loading*. Kemudian, peneliti melakukan pengujian CFA kembali dengan jumlah item yang sudah dimodifikasi. Hasilnya ditunjukkan oleh tabel berikut ini:

**Tabel 6. Perbandingan Hasil Uji Reliabilitas Item Awal dan Item Terpilih**

Uji Reliabilitas	Item Awal		Item Terpilih	
	Jumlah Item	Nilai Cronbach Alpha	Jumlah Item	Nilai Cronbach Alpha
<b>Skala Karakter Baik</b>	<b>42</b>	<b>0,885</b>	<b>42</b>	<b>0,892</b>
Pengetahuan Moral	23	0,703	14	0,710
Sikap Moral	17	0,708	12	0,736
Perilaku Moral	23	0,795	16	0,795

Berdasarkan nilai *Cronbach alpha* pada tabel 6, dapat diketahui bahwa baik alat ukur secara keseluruhan maupun pada setiap dimensi setelah mengeliminasi 21 item, tetap memiliki konsistensi internal yang tinggi, bahkan semakin tinggi. Maka dari itu, alat ukur Skala Karakter Baik tetap dapat dipastikan reliabel sehingga menghasilkan konsistensi hasil yang sama jika dilakukan pengukuran pada sampel berbeda (dengan karakteristik sama) dan dalam waktu yang berbeda. Selain melihat nilai *Cronbach alpha*, peneliti juga melihat kembali nilai *corrected item-total correlation* dari dimensi dan total keseluruhan pada pengujian kedua. Hasilnya 39 item menampilkan nilai yang lebih baik, diatas 0,3 (Nunnally & Beirnsstein, 1994). Hal ini menunjukkan bahwa penyusun telah mengeliminasi item yang tepat. Meskipun demikian, terdapat 4 item yang masih memiliki nilai *corrected item-total correlation* 0,3 yang masih dipertahankan, karena pertimbangan item yang tidak dapat dieliminasi karena keterwakilannya terhadap indikator (Lickona, 1991) dan nilai PPK.

### Uji Validitas Lanjutan

Berikutnya, peneliti melakukan pengukuran validitas kembali dengan melihat hasil *index model-fit chi-square*, RMSEA p-value, SRMR, CFI, dan GFI-nya. Hasilnya didapatkan sebagai berikut:

**Tabel 7. Hasil Indeks Model-Fit dalam Pengujian Ulang**

Indeks Model-Fit	Kriteria (Hair et.al., 2019)	Sebelum Modifikasi		Setelah Modifikasi	
		Hasil	Keterangan	Hasil	Keterangan
<b>Chi-square</b>	>0,05	<0,01	Model Tidak Fit	<0,01	Model Tidak Fit
<b>RMSEA p-value</b>	<0,07	0,000	Model Fit	0,000	Model Fit
<b>SRMR</b>	<0,08	0,077	Model Fit	0,078	Model Fit

<b>GFI</b>	>0,90	0,942	Model Fit	0,957	Model Fit
<b>CFI</b>	>0,90	0,469	Model Tidak Fit	0,568	Model Tidak Fit

Meskipun hasilnya tetap tidak memenuhi kriteria untuk *chi-square* dan CFI, model skala karakter baik masih dianggap fit jika melihat indeks tabel RMSEA p-value, SRMS, dan GFI (Hair et.al., 2019). Berikutnya adalah melihat kembali perbedaan pengukuran *factor loading*-nya dari tiap dimensinya:

**Tabel 8. Factor Loading Tiga Dimensi Skala Karakter Baik (Second-Order Factor Loading) Setelah Pengujian Ulang**

Dimensi	Sebelum Modifikasi			Setelah Modifikasi			
	<i>Std.</i> (semakin semakin baik)	<i>Error</i> kecil	<i>p-value</i> (<0,05)	<i>Factor Loading</i> (>0,3)	<i>Std.</i> (semakin semakin baik)	<i>Error p-value</i> kecil (<0,05)	<i>Factor Loading</i> (>0,3)
<b>Pengetahuan Moral</b>	0,509		<0,01	0,944	1,248	0,002	0,969
<b>Sikap Feeling</b>	0,313		<0,01	0,921	0,210	<0,001	0,886
<b>Tingkah Laku Moral</b>	1,411		0,002	0,975	1,315	0,002	0,972

Berdasarkan tabel 8, dapat diketahui bahwa *factor loading* tiga dimensi setelah beberapa item dihapus masih menunjukkan angka yang tinggi untuk sampel yang besar (>0,3). Hasil ini menunjukkan bahwa ketiga dimensi valid dalam merepresentasikan variabel laten yang diukur (Hair et.al., 2019). Selain itu, peneliti juga melihat perubahan *factor loading* dari item-item tiap dimensinya. Hasilnya mayoritas item menunjukkan nilai *factor loading* yang baik, di atas  $\geq 0,3$  (Hair et.al., 2019). Adapun *model plot* yang berisi *factor loading second-order* terlampir pada *appendix*.

### Menentukan Norma

Agar skor SKB dapat dimaknai, peneliti menyusun norma *within group norm*. Norma ini dipilih karena alat ukur yang dirancang bertujuan mengevaluasi karakteristik kepribadian seseorang yaitu *good character* pada relasinya ke individu lain dalam kelompok yang sama yaitu siswa SMA. Distribusi skor akan dipengaruhi oleh representasi dari populasi yang diambil berdasarkan karakteristik yang dimiliki subjek (Chadha, 2009; Urbina, 2004; Crocker & Algina, 2008) pada kasus ini karakteristik tersebut adalah siswa SMA. Perhitungan norma akan dilakukan dengan menggunakan *standard score (T-Score)*, dengan alasan kemudahan pada saat membaca dan mengartikan norma, karena *T-score* tidak memiliki nilai negatif pada norma (Urbina, 2004). Pembagian kategori skala dilakukan berdasar prinsip pertimbangan probabilitas, agar peluang memperoleh subjek yang proporsional dalam setiap kategori (Azwar, 2012). Maka, berdasarkan perhitungan nilai *standard score* menggunakan *T-score* yang dilakukan terhadap sampel pengambilan data yang telah dibersihkan (n=728) dengan 42 item, dibuat kategorisasi norma skor sebagai berikut:

**Tabel 9. Kategorisasi Norma Standar Score**

Kategori	T-Score
Sangat Tinggi	60-80
Tinggi	40-59
Rendah	20-39
Sangat Rendah	<9-19

Berdasarkan kategori tersebut, didapatkan hasil norma dan pengelompokannya sebagai berikut:

**Tabel 10. Hasil Norma Skala Karakter Baik**

<i>Standard Score</i>	<i>SKB Raw Score (RS)</i>	<i>Pengetahuan Moral RS</i>	<i>Sikap Moral RS</i>	<i>Tingkah Moral RS</i>	<i>Laku</i>	<i>Kategorisasi</i>
<b>60-80</b>	149-168	49-56	44-48	57-64		<b>Sangat Tinggi</b>
<b>40-59</b>	122-148	40-48	36-43	46-56		<b>Tinggi</b>
<b>20-39</b>	108-121	34-39	30-35	34-45		<b>Rendah</b>
<b>&lt;9-19</b>	82-95	24-29	22-26	27-33		<b>Sangat Rendah</b>

Selanjutnya penyusun mencoba menginterpretasi skor responden, diambil 3 responden secara random dan ditemukan hasil interpretasinya sebagai berikut:

**Tabel 11. Contoh Interpretasi Skor Responden**

<i>No.</i>	<i>Raw Score</i>			<i>Standardized Score</i>					<i>Interpretasi</i>			
	<i>Peng.M</i>	<i>Per.M</i>	<i>Ting.M</i>	<i>SKB</i>	<i>MK</i>	<i>MF</i>	<i>MA</i>	<i>SKB</i>	<i>Peng.M</i>	<i>Per.M</i>	<i>Ting.M</i>	<i>SKB</i>
<b>243</b>	23	26	27	76	<9	10-19	<9	<9	Sangat Rendah	Sangat Rendah	Sangat Rendah	Sangat Rendah
<b>412</b>	45	35	47	127	20-29	30-39	40-49	40-49	Rendah	Rendah	Tinggi	Tinggi
<b>84</b>	54	48	63	165	70-80	70-80	70-80	60-69	Sangat Tinggi	Sangat Tinggi	Sangat Tinggi	Sangat Tinggi

**SKB=Skala Karakter Baik; Peng.M=Pengetahuan Moral; Per.M=Sikap Moral; Ting.M=Tingkah Laku Moral**

Tabel 11 menjelaskan dinamika interpretasi dari tiga contoh responden yang diambil secara random, dimana dapat dilihat pada responden 243 yang menampilkan karakter baik yang sangat rendah, dimana keseluruhan dimensi baik pengetahuan moral, sikap moral, maupun tingkah laku moral dalam kondisi yang sangat rendah. Hal ini sangat berkebalikan dari responden 84 dan menunjukkan bahwa ia menampilkan karakter baik yang sangat tinggi, dimana keseluruhan dimensi baik pengetahuan moral, sikap moral, maupun tingkah laku moral yang sangat baik. Hal ini menunjukkan ada keterpaduan pada responden 84 dalam melakukan kebajikannya, ia mampu berpikir dan merasakan esensi dari berperilaku baik, tidak hanya sekedar perilaku yang ditampilkannya saja. Sementara, pada responden 412 secara keseluruhan menampilkan karakter baik yang tinggi, dengan menampilkan pengetahuan moral dan tingkah laku moral yang baik, tetapi ia memiliki sikap moral yang rendah. Hal ini menunjukkan bahwa meskipun 412 memiliki perilaku yang baik, belum tentu ia memahami alasan dibalik tingkah laku yang ia tampilkan dan merasakan emosi positif ketika melakukan tingkah laku moral. Ia hanya sekedar berperilaku baik tanpa memahami esensi dari berperilaku baiknya tersebut.

#### **4. Diskusi**

Skala Karakter Baik (SKB) terbukti reliabel dan valid. Model SKB dapat mengukur konstruk karakter baik sesuai yang dikemukakan oleh Lickona (1991) dan sesuai dengan nilai prioritas pendidikan karakter di Indonesia. Hal ini sesuai dengan pernyataan Jerome (2019) yang mendalami pendidikan karakter, bahwa “*nilai yang berbeda dengan program yang berbeda di setiap daerah akan memberikan penekanan yang berbeda juga di setiap indikator pendidikan karakter*”. Meskipun demikian, ditemukan bahwa model tidak fit jika dilihat dari indeks model fit Chi-square dan CFI. Hal ini dapat terjadi karena *Chi-square* sendiri memiliki beberapa isu yang terkait dengan jumlah sampel dan variabel yang diukur dimana pengujian *Chi-square* menjadi bias dan tidak lagi bermakna seiring meningkatnya jumlah sampel dan variabel dalam kompleksitas model (Hair dkk., 2019). Selain itu tidak validnya indeks model CFI juga dapat terjadi karena CFI merupakan bentuk pengembangan Normed Fit Index (NFI) dimana ditetapkan norma pada *Chi-square* untuk mengurangi bias kompleksitas jumlah variabel,



meskipun demikian CFI tidak menjawab bias di dalam meningkatnya jumlah sampel. CFI yang dikembangkan dari chi-square memiliki kelemahan dan bias yang sama mengenai jumlah sampel. Sedangkan, GFI merupakan pengukuran statistik model fit yang dikembangkan untuk sensitif terhadap bias jumlah sampel besar (Hair dkk., 2019). Mengingat tujuan mendasar kedua pengukuran fit, sangatlah wajar bahwa CFI menjadi tidak fit saat GFI menunjukkan angka yang baik (diatas 0.9).

Pengujian alat ukur Skala Karakter Baik ini memiliki keterbatasan dalam pengujian validitas, yakni tidak dilakukan pengujian *content validity* melalui *expert judgement*. Padahal pengujian *expert judgement* dapat bermanfaat untuk melihat setiap kualitas bunyi item dalam tiap dimensi (Urbina, 2004; Chadha, 2009). Hal ini disebabkan oleh keterbatasan waktu sehingga tidak menemukan jadwal/waktu yang tepat untuk menghubungi seorang ahli di bidang Pendidikan Karakter. Oleh sebab itu, peneliti yang akan menggunakan alat ukur ini sebaiknya dapat melakukan proses *expert judgement* terlebih dahulu dan melakukan test *try-out* untuk menguji reliabilitas dan validitas kembali pada sampel yang akan dituju dalam penelitian. Penelitian pengembangan alat ukur menggunakan konstruk karakter baik (Lickona, 1991) juga pernah dilakukan oleh Layyinah dan Roebianto (2020) dengan alat ukur yang dinamakan “Karakter Moral Remaja”. Perbedaan mendasar dari Skala Karakter Baik (SKB) dengan Skala Pembentukan Karakter Moral Remaja adalah pada SKB terdapat kontekstualisasi nilai Prioritas PPK berdasarkan arahan Kemendikbud pada item-item alat uku, yakni: (1) religius, (2) nasionalis, (3) mandiri, (4) gotong royong, dan (5) integritas. Sementara, pada Skala Pembentukan Karakter Moral Remaja lebih didasarkan pada nilai-nilai moral yang universal. Selain itu, perbedaan berikutnya dilihat dari jumlah sampel yang dilibatkan. Sampel dalam pengujian alat ukur SKB lebih banyak ( $n=728$ ) dengan latar belakang sekolah dan daerah yang lebih beragam, mulai dari Bogor, Jakarta, Bandung, Surakarta, dan Malang. Meskipun demikian, persebaran sampel ini belum cukup merepresentasikan kondisi pelajar SMA/Sederajat di Indonesia sehingga alat ukur SKB ini setidaknya dapat mewakili kondisi pelajari di pulau Jawa. Sementara pada Skala Pembentukan Karakter Moral Remaja, jumlah sampel lebih sedikit ( $n=153$ ) yang melibatkan 2 sekolah menengah atas di Jakarta sehingga cakupannya lebih sempit. Kemudian, perbedaan berikutnya terletak pada cara pengujian CFA. Pada SKB, pengujian hanya dilakukan dengan CFA *second-order* dengan alasan SKB memang merupakan alat ukur yang dirancang untuk dapat dimaknai secara multidimensional sehingga peneliti langsung melihat dinamika keterkaitan konstruk ke dimensi dan dimensi ke item secara langsung dengan CFA *second-order*. Sementara pada pada alat ukur Karakter Moral Remaja oleh Layyinah dan Roebianto (2020) pengujian CFA dilakukan dengan *first-order* dan *second-order* yang mempertimbangkan alat ukur dapat dimaknai secara *unidimensional* maupun *multidimensional*.

## 5. Kesimpulan

Karakter baik pada pelajar dilihat berdasarkan keselarasan antara pengetahuan moral, sikap moral, dan perilaku moral. Di Indonesia sendiri penerapan karakter baik diwujudkan dalam pendidikan karakter yang disusun oleh Kemendikbud melalui program unggulannya, yakni Penguatan Pendidikan Karakter dengan mengenalkan lima nilai utama. Hal ini menjadi penting untuk diinternalisasi pada setiap pelajar, khususnya sekolah menengah, agar dapat menghadapi dunia dengan baik dan mampu membawa kebermanfaatannya dengan optimal.

Berdasarkan hasil pengujian alat ukur Skala Karakter Baik, setelah dilakukan uji reliabilitas dan validitas dari yang semula 63 item awal yang diujikan menjadi 42 item terpilih. Pada 42 item terpilih dapat disimpulkan bahwa Skala Karakter Baik reliabel dalam arti memiliki konsistensi internal yang tinggi. Selain itu, Skala Karakter Baik juga terbukti valid pada uji validitas dengan metode CFA *second-order* yang dilihat dari indeks model fit RMSEA p-value, SRMR, dan GFI dalam mengukur dimensi terhadap konstruk. Apabila hendak dilakukan pengembangan terhadap alat ukur Skala Karakter Baik lebih lanjut, peneliti menyarankan untuk melakukan uji validitas lainnya selain *construct validity*. Misalnya dengan *content validity* melalui *expert judgement* oleh orang yang ahli di bidang Pendidikan Karakter, serta mengkorelasikannya dengan alat ukur lain yang serupa seperti alat ukur Pembentukan Karakter Moral Remaja oleh Layyinah dan Roebianto (2020).

## 6. Referensi

- Baehr, J. (2017). The Varieties of Character and Some Implications for Character Education. *Journal of Youth and Adolescence*, 46(6), 1153-1161. doi:10.1007/s10964-017-0654-z
- Berkowitz, M. W., & Bier, M. C. (2014). Research-Based Fundamentals of the Effective Promotion of Character Development in Schools. *Handbook of Moral and Character Education*. doi:10.4324/9780203114896.ch14
- Chadha, N. K. (2009). *Applied Psychometry*. New Delhi: Sage Publications.
- Crocker, L. M., & Algina, J. (2008). *Introduction to classical and modern test theory*. Mason, OH: Cengage Learning.
- Dempster, M. 2020. An exploration of character education as a tool of ‘moral repair’ in the developing world. *Journal of Religious Education*. <https://doi.org/10.1007/s40839-020-00107-5> diakses pada 21 September 2020.

- Elias, M. J., Nayman, S. J., & Duffell, J. C. (2018). Scaling Up High-Quality Social-Emotional and Character Development in All Schools: A Set of Policy Recommendations to the US Secretary of Education. *The Springer Series on Human Exceptionality Emotional Intelligence in Education*, 321-350. doi:10.1007/978-3-319-90633-1\_12
- Fatoni, A. (2017). The strategy of character education in the globalization era. *International Journal of Scientific & Technology Research Volume 6, ISSUE 04, 112-114*
- Field, A. (2018). *Discovering Statistics using SPSS 5th ed.* London: SAGE Publication Inc.
- Hair, et al. (2019). *Multivariate data analysis.* Andover: Cengage Learning.
- Ismail, I., Thalib, S., Samad, S., & Mahmud, R. (2016). The Development of Character Education Model to Improve Students' Academic Independence in Islamic Boarding School in Sinjai District, Indonesia. *The New Educational Review*, 46(4), 29-39. doi:10.15804/tner.2016.46.4.02
- Jerome, L., & Kisby, B. (2019). *The Rise of Character Education in Britain.* doi:10.1007/978-3-030-27761-1
- Kemendikbud. (2017). *Konsep dan Pedoman Penguatan Pendidikan Karakter.* [https://repositori.kemdikbud.go.id/10075/1/Konsep\\_dan\\_Pedoman\\_PPK.pdf](https://repositori.kemdikbud.go.id/10075/1/Konsep_dan_Pedoman_PPK.pdf)
- Kumar, R. (2011). *Research Methodology Third Edition.* India: SAGE Publication.
- Layyindah., Roebianto, A. (2020). Uji Validitas Konstruk Pembentukan Karakter Moral Remaja. *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)* 9(2), 2020, 99-110. DOI: <http://dx.doi.org/10.15408/jp3i.v9i2.XXXXXX>. <http://journal.uinjkt.ac.id/index.php/jp3i>
- Lickona, T. (1991). *Educating For Character: How Our Schools Can Teach Respect and Responsibility.* New York: Bantam.
- Nunally, J.C. and Bernstein, I.H. (1994). The Assessment of Reliability. *Psychometric Theory*, 3, 248-292.
- Pala, A. (2011). The need for character education. *International journal of social sciences and humanity studies Vol 3*, 23-32.
- Taqwim, A. (2017). Urgency of Character Education Based on Multiculturalism. Proceedings of the 1st International Conference on Social Sciences Education - "Multicultural Transformation in Education, Social Sciences and Wetland Environment" (*ICSSE 2017*). doi:10.2991/icsse-17.2018.3
- Urbina, S. (2004). *Essentials of Psychological Testing.* Hoboken, NJ: Wiley.
- Yudiana, W. (2020, October 16). *Confirmatory Factor Analysis (CFA) Mata Kuliah : Statistik Lanjut PSPP801305.* Lecture. Universitas Indonesia
- Yu, Y. (2018). Comparing Moral Education Models at a Military Academy in Taiwan. *Journal of Academic Ethics*, 16(2), 173-193. doi:10.1007/s10805-018-9301-x

# Appendix

## Model Plot Second Order

