

Metode Alternatif untuk Mendeteksi Bias Respons *Social Desirability* pada Item-item Tes Kepribadian

Bina Ciptadi

Universitas Indonesia

Jahja Umar

Universitas Islam Negeri Jakarta

Abstrak

Penelitian ini membahas tentang metode untuk mendeteksi kerentanan item self-report inventory terhadap bias respons social desirability. Metode analisis faktor dengan pendekatan Confirmatory Factor Analysis (CFA) diajukan sebagai salah satu metode alternatif yang dapat digunakan untuk mengatasi social desirability pada level item. Dalam pendekatan CFA, social desirability dimodelkan sebagai suatu faktor. Item-item yang dinilai rentan dari hasil rating dijadikan kriteria sebagai item yang bias dan apakah analisis model CFA mampu membuktikan bahwa item-item tersebut tergolong ke dalam satu faktor yang sama yaitu faktor social desirability. Dari contoh analisis yang dilakukan terhadap hasil pengukuran salah satu inventori kepribadian, metode CFA berhasil mendeteksi item-item yang rentan terhadap bias respons social desirability.

Kata kunci:

social desirability, confirmatory factor analysis, respons bias.

Dampak *social desirability* pada hasil pengukuran telah lama menjadi kekhawatiran bagi para pengembang tes kepribadian (Leite & Cooper, 2010). *Social desirability* adalah salah satu bentuk bias respons yang terjadi ketika responden menjawab item *self-report*, bukan karena ingin melaporkan tentang keadaan dirinya yang sebenarnya melainkan ingin melindungi citra dirinya agar dinilai positif di mata sosial (Cohen & Swerdlik,

2005). Marlowe dan Crowne (1960) menyatakan bahwa *social desirability* merupakan bentuk motivasi di mana adanya kebutuhan dalam diri seseorang untuk memperoleh penilaian yang positif dari orang lain atau ingin memenuhi harapan sosial, dengan cara menampilkan perilaku-perilaku yang dianggap sesuai atau dapat diterima dalam kultur sosial.

Social desirability dapat mengakibatkan hasil pengukuran *self-*

report menjadi tidak akurat karena ada hal lain yang ikut terukur oleh tes yang tidak ada kaitannya dengan obyek perilaku yang hendak diukur (Marlowe & Crowne, 1960). Jika permasalahan ini tidak segera diatasi, maka validitas hasil pengukuran *self-report* tersebut akan diragukan. Oleh karena itu, para peneliti dan pengembang tes kepribadian telah melakukan banyak upaya untuk mengatasi ancaman bias respons *social desirability*. Beberapa metode atau strategi telah banyak dilakukan. Di antara metode yang sering dilakukan antara lain dengan menggunakan metode *demand reduction*, *rational technique* dan *covariate technique* (Paulhus, 1991).

Metode *demand reduction* dilakukan untuk mengurangi tekanan situasi saat pengesanan berlangsung yang dapat menyebabkan responden menjawab secara *socially desirable*. Bentuk pelaksanaan metode ini bervariasi misalnya: memberlakukan anonimitas pada identitas responden dengan tidak membubuhkan nama pada lembar jawaban, meminta secara langsung agar responden tidak memberi jawaban yang tidak sesuai dengan kondisi dirinya atau cara yang paling ekstrim adalah dengan menggunakan teknik *bogus pipeline* yaitu responden dihubungkan dengan sebuah mesin pendeteksi kebohongan palsu dan dikatakan bahwa mesin tersebut dapat mengetahui apabila responden berbohong (Paulhus, 1991). Tidak ada jaminan pasti bahwa penggunaan metode *demand reduction* ini berhasil mengatasi *social desirability* atau

mencegah responden untuk tidak menjawab item *socially desirable*.

Metode berikutnya adalah *rational technique* yang berupaya untuk mengatasi *social desirability* dengan cara mengubah format item ke dalam bentuk *forced-choice* di mana dua item yang dipasangkan dianggap memiliki tingkat *social desirability* sama, lalu responden menjawab dengan memilih salah satu dari item yang dipasangkan tersebut (Cohen & Swerdlik, 2005). Walau bagaimanapun juga, pada kenyataannya sangat sulit memasang item dengan tingkat *social desirability* yang setara (Lanyon & Goodstein, 1997). Selain itu, *missing response* juga sering terjadi akibat responden merasa kebingungan untuk memilih salah satu dari dua pilihan item yang tersedia (Paulhus, 1981). Oleh karena itu, metode dengan mengubah format ini belum juga efektif untuk mengatasi dampak *social desirability* pada tes kepribadian (Anastasi & Urbina, 1997; Lanyon & Goodstein, 1997; Domino & Domino, 1999; Cohen & Swerdlik, 2005).

Metode lain yang termasuk paling banyak digunakan yaitu *covariate technique*. *Covariate technique* dilakukan dengan mengukur kecenderungan *social desirability* dalam diri responden melalui alat ukur yang dibuat khusus untuk mengukur *social desirability*. Alat ukur *social desirability* yang paling banyak digunakan hingga saat ini antara lain adalah *Marlowe-Crowne Social Desirability Scale* (MCSDS) (Marlowe & Crowne, 1960) dan *Balanced Inven-*

tory of Desirable Responding (BIDR) (Paulhus, 1984). Bahkan di beberapa inventori kepribadian seperti MMPI juga memiliki sub-skala tersendiri untuk mendeteksi manipulasi respon seperti *L-Scale* dan *F-Scale* (Paulhus, 1991). Alat ukur tersebut sering disebut juga sebagai skala validitas karena berfungsi untuk menguji validitas hasil pengukuran dari inventori yang digunakan. Skor yang dihasilkan alat ukur *social desirability* itu nantinya akan dikorelasikan dengan hasil pengukuran *self-report*. Jika terdapat korelasi yang signifikan antara skor *self-report* dengan skor alat ukur *social desirability*, maka *self-report* tersebut dinyatakan tidak valid (Paulhus, 1991).

Dari ketiga metode yang telah disebutkan di atas, pada dasarnya ketiganya tergolong sebagai metode yang konvensional. Ketiga metode tersebut juga dapat dikatakan kurang efektif karena tidak bisa memberi bukti empirik bahwa responden benar-benar telah berbohong atau memanipulasi jawabannya. Bahkan metode *covariate technique* yang paling banyak ditemukan di berbagai literatur penelitian sebagai metode untuk mengoreksi hasil pengukuran *self-report* telah menuai kritik karena dianggap merugikan responden dengan menuduh mereka berbohong padahal kenyataannya bisa saja tidak (Zickar & Drasgow, 1996). Jika demikian, maka selama cara yang ditempuh tetap berfokus pada deteksi atau pengukuran *social desirability* sebagai bagian dari karakteristik yang ada dalam diri seseorang, maka cara

tersebut tidak akan pernah menjadi cara yang efektif.

Social desirability sesungguhnya tidak hanya dipandang sebagai karakteristik orang, melainkan dapat pula dianggap sebagai karakteristik dari item tes (Marlowe & Crowne, 1960). Item yang dimaksud memiliki karakteristik *social desirability* adalah item yang dianggap memiliki potensi untuk direspon secara *socially desirable*. Sesuai dengan pernyataan Anastasi dan Urbina (1997) bahwa isi item suatu *self-report* dapat memiliki kemungkinan untuk mendorong responden memberikan jawaban yang tidak sebenarnya. Oleh karena itu, bias respons hasil pengukuran *self-report* akibat *social desirability* sesungguhnya terjadi akibat adanya interaksi antara item dengan orang di mana item dapat menstimulasi kecenderungan *social desirability* yang ada dalam diri responden. Oleh karena itu, dengan adanya pandangan ini, yaitu jika item dianggap bisa memicu timbulnya kecenderungan *social desirability*, maka cara lain yang mungkin lebih efektif dengan tidak melibatkan tes seperti MCSDS adalah dengan memfokuskan upaya deteksi *social desirability* pada level item dan bukan pada orang.

Ada sebuah cara yang dapat diharapkan lebih akurat dalam mendeteksi ada tidaknya *social desirability* pada respons terhadap item yaitu dengan metode analisis faktor. Cara ini pernah diperkenalkan oleh Paulhus (1981) dan Ferrando (2005). Perbedaan di antara keduanya adalah Paulhus (1981) menggunakan pen-

dekatan *Principal Component Analysis* (PCA) sedangkan Ferrando (2005) lebih menyarankan untuk menggunakan pendekatan *Confirmatory Factor Analysis* (CFA). Menurut Ferrando (2005), pendekatan PCA kurang praktis digunakan dibandingkan pendekatan CFA. Metode CFA dilakukan dengan asumsi bahwa hasil pengukuran *self-report* mengandung dua faktor yaitu faktor konstruk yang hendak diukur dan faktor *social desirability* sebagai faktor lain yang ikut terukur dan tak ada kaitannya dengan konstruk yang hendak diukur. Meski metode CFA yang diperkenalkan oleh Ferrando (2005) lebih mudah digunakan namun metode tersebut masih melibatkan item-item dari alat ukur *social desirability* (seperti MCSDS) sebagai indikator yang mewakili faktor *social desirability* dalam model.

Bagaimana pun juga, pengukuran *social desirability* sebagai karakteristik orang adalah hal yang ingin dihindari pada studi ini. Tujuan studi ini adalah menemukan upaya alternatif untuk mendeteksi *social desirability* pada level item dengan pendekatan CFA, namun tanpa melibatkan pengukuran *social desirability* dalam diri seseorang. Oleh karena itu, dalam studi ini, indikator dari faktor *social desirability* seharusnya ditentukan dari item-item yang berpotensi rentan terhadap *social desirability*. Dengan demikian, tuduhan tidak lagi mengarah pada diri responden melainkan pada item karena item memiliki potensi untuk direspons secara *socially desirable*.

Metode

Metode yang digunakan dalam studi ini adalah pendekatan *Confirmatory Factor Analysis* (CFA) yang diaplikasikan pada data dari hasil pengukuran sebuah *self-report* yang memiliki format jawaban dalam skala Likert. Melalui pendekatan CFA ini, *social desirability* akan dimodelkan sebagai sebuah faktor. Analisis model CFA ini dilakukan dengan menggunakan *software LISREL 8.8* (Joreskog & Sorbom, 2006). Selain untuk membuktikan apakah item-item *self-report* mengukur konstruk yang memang hendak diukur, metode CFA juga digunakan untuk menguji apakah item-item yang dinilai berpotensi rentan terhadap *social desirability* memang tergolong ke dalam satu faktor yang sama yaitu faktor *social desirability*.

Oleh karenanya, sebelum menganalisis model dengan pendekatan CFA, potensi kerentanan item *self-report* terhadap *social desirability* diidentifikasi dari hasil *rating* secara kualitatif oleh sejumlah *rater*. Hal ini sesuai dengan pernyataan Paulhus (1981) bahwa nilai kerentanan item terhadap *social desirability* (*Social Desirability Scale Value* atau SDSV) dapat diperoleh dari hasil *rating* selain dari hasil pengukuran tes seperti MCSDS. Format penilaian skala *rating* SDSV dalam studi ini disusun dengan empat pilihan penilaian yaitu mulai dari kondisi “tidak rentan” (Skor 1), “agak rentan” (Skor 2), “rentan” (Skor 3) hingga “sangat rentan” (Skor 4). SDSV item diperoleh dari perhitungan nilai rata-rata hasil

rating para *rater* di setiap item. Item-item yang memiliki nilai rata-rata SDSV tinggi yaitu sebesar 3.00 hingga 4.00 dinilai sebagai item yang berpotensi terhadap bias respon *social desirability*. Meski demikian, keputusan untuk menentukan item-item yang rentan tidak hanya dilakukan berdasarkan nilai rata-rata SDSV yang tinggi saja, tapi dipilih juga berdasarkan nilai varians yang kecil. Dari nilai varians skor *rating* tiap item yang cenderung kecil, dapat diketahui bahwa nilai *rating* yang diberikan oleh para narasumber tidak begitu bervariasi. Dengan demikian, hanya item-item yang memiliki nilai rata-rata *rating* tinggi dan varians kecil yang dijadikan kriteria dalam analisis CFA sebagai item-item yang rentan terhadap bias respons *social desirability* (Ciptadi, 2011).

Oleh karena metode CFA digunakan sebagai metode pengujian hipotesis (Joreskog & Sorbom, 1993; Byrne, 1998), maka uji hipotesis akan dilakukan dengan menilai model fit secara keseluruhan yaitu ketika *social desirability* tidak diikutsertakan sebagai salah satu faktor (model unidimensional), dan ketika dilibatkan sebagai faktor dalam model (*bi-factor model*). Hal ini dilakukan untuk membuktikan dugaan bahwa model dengan *social desirability* sebagai salah satu faktor (*bi-factor model*) adalah model yang lebih fit. Jika model dengan melibatkan faktor *social desirability* terbukti lebih fit maka hasil ini bisa mendukung asumsi yang mengatakan bahwa pengukuran *self-report* akan mengandung faktor lain

selain faktor konstruk yang hendak diukur yaitu faktor *social desirability*.

Penilaian model fit secara keseluruhan dilakukan dengan cara melihat besaran indeks kecocokan model (*goodness of fit index*) dari *output LISREL*. Kriteria penilaian model fit dapat dilihat berdasarkan nilai *Chi-Square* (χ^2), *Root Mean Square Error of Approximation* (RMSEA), dan masih ada beberapa indeks lain yang dihasilkan oleh *output LISREL* yang bisa dijadikan pedoman untuk menilai model fit. Jika penilaian model fit dilakukan berdasarkan nilai χ^2 maka diharapkan *p-value* dari nilai χ^2 yang tidak signifikan (*p-value* > 0.05). Tidak signifikannya nilai χ^2 (tidak ditolaknya hipotesis nihil $S-\Sigma=0$, di mana S adalah matriks kovarians dari data dan Σ adalah matriks kovarians menurut model) mengindikasikan bahwa tidak ada perbedaan signifikan antara model dengan data (Joreskog & Sorbom, 1993). Sedangkan jika penilaian model fit dilakukan berdasarkan RMSEA, maka besaran nilai RMSEA diharapkan ≤ 0.05 untuk bisa dinyatakan bahwa model adalah model yang *close fit*, atau nilai RMSEA ≤ 0.08 untuk dinilai sebagai model yang *good fit* (Browne & Cudeck, 1993). Uji signifikansi parameter tiap item juga dilakukan untuk menguji kemampuan item-item dalam mengukur faktor yang diwakilinya. Uji signifikansi kontribusi dari tiap-tiap item dilakukan dengan cara melihat besaran *t-value* yang diharapkan lebih besar dari 1.96 agar item bisa dinyatakan sebagai item yang berkontribusi secara signifikan terhadap faktor yang

diwakilinya (Byrne, 1998). Dari hasil uji signifikansi kontribusi item ini dapat pula diketahui apakah item-item yang dijadikan kriteria sebagai item bias terbukti memiliki kontribusi yang signifikan pada faktor *social desirability*. Jika analisis CFA menunjukkan model fit dan item dengan SDSV tinggi terbukti sebagai item yang selain mampu mengukur konstruk yang ingin diukur, juga mampu mengukur *social desirability*, maka metode CFA dinyatakan berhasil mengidentifikasi item-item *self-report* yang bias.

Contoh Analisis Data

Data dan Model:

Pada contoh ini akan ditunjukkan bagaimana penerapan metode CFA dalam menganalisis item-item *self-report inventory* yang rentan terhadap bias respons *social desirability*. Data diperoleh dari 195 orang mahasiswa S1 yang mengisi inventori kepribadian *Values in Action - Inventory of Strength (VIA-IS)*. Pengumpulan data tersebut dilakukan dalam rangka penyusunan tesis di Universitas Indonesia (Ciptadi, 2011). Inventori VIA-IS yang digunakan adalah hasil adaptasi ke dalam Bahasa Indonesia dan sebelumnya juga pernah digunakan dalam penelitian skripsi/tesis di Universitas Indonesia (Lestari, 2006; Wijayanti, 2008).

Inventori VIA-IS merupakan salah satu *self-report inventory* yang dibuat untuk mengukur kekuatan karakter (*character strength*) positif yang ada dalam diri manusia dan dipakai dalam

penelitian di bidang psikologi positif (Peterson & Seligman, 2004). Terdapat 24 faktor *character strengths* yang diukur melalui inventori ini dimana pada dasarnya setiap *character strength* mewakili suatu faktor laten yang lebih tinggi dan bersifat abstrak yang disebut *virtue*. *Virtue* merupakan suatu nilai moral tertinggi yang hanya dapat tercapai ketika seseorang mengembangkan *character strength* yang ada dalam dirinya (Carr, 2004).

Menurut Peterson dan Seligman (2004) ada enam *virtue* yang ingin dijunjung oleh manusia yaitu *wisdom* (kebijaksanaan), *courage* (keberanian), *humanity* (berperikemanusiaan), *justice* (keadilan), *temperance* (kesederhanaan) dan *transcendence* (transendensi). Masing-masing *virtue* terwakili oleh *character strength* yang berbeda, seperti *virtue* „wisdom’ yang diwakili oleh lima *character strengths* seperti *creativity*, *curiosity*, *love of learning*, *open-mindedness* dan *perspective*. Begitu pula *virtue* lainnya yang masing-masing diwakili oleh *character strength* lainnya.

Setiap skala *character strength* memiliki 10 item, karena ada 24 skala yang mengukur *strength* sehingga total keseluruhan item inventori VIA-IS berjumlah 240 item (Tabel 1). Seluruh item memiliki format respons dengan menggunakan skala Likert dengan enam pilihan jawaban yaitu dari pilihan „sangat tidak sesuai“ (skor 1), „tidak sesuai“ (skor 2), „agak tidak sesuai“ (skor 3), „agak sesuai“ (skor 4), „sesuai“ (skor 5) dan „sangat sesuai“ (skor 6). Tidak ada *unfavorable item* dalam inventori VIA-IS berbahasa

Indonesia ini sehingga tidak ada item yang perlu diskor terbalik (*reversed*).

Beberapa contoh item yang mengukur *character strength* dalam inventori tersebut adalah “Saya selalu mengidentifikasi alasan-alasan di balik tindakan saya.” (mengukur *strength* „*open-mindedness*“ yang mewakili *virtue* „*wisdom*“), “Saya selalu menghadapi ketakutan-ketakutan saya” (mengukur *strength* „*bravery*’ yang mewakili *virtue* „*courage*“), “Apapun yang dilakukan oleh keluarga dan teman-teman dekat saya, tidak akan membuat saya berhenti mencintai mereka” (mengukur *strength* „*love*“ yang mewakili *virtue* „*humanity*“), “Saya memperlakukan semua orang sama rata tanpa mempedulikan siapa orang itu” (mengukur *strength* „*fairness*’ yang mewakili *virtue* „*justice*“), “Saya selalu berhati-hati dalam menentukan pilihan” (mengukur *strength* „*prudence*’ yang mewakili *virtue* „*temperance*“) dan “Paling tidak sekali dalam sehari, saya berhenti dari rutinitas & bersyukur atas segala anugerah dari Tuhan” (mengukur *strength* „*gratitude*’ yang mewakili *virtue* „*transcendence*“).

Pengujian reliabilitas hasil pengukuran inventori VIA-IS dalam versi asli berbahasa Inggris dilaporkan menunjukkan hasil yang memuaskan di setiap skala *strength* dengan rata-rata besaran nilai *Cronbach’s Alpha* > 0.7. Begitu pula hasil pengujian kembali pada sampel yang sama (*test-retest reliability*) dalam rentang waktu selama empat bulan menunjukkan tingkat kestabilan yang cukup memuaskan

($r_{xx} > 0.7$) pada semua skala (Peterson & Seligman, 2004).

Laporan hasil pengujian reliabilitas pada inventori VIA-IS hasil adaptasi ke dalam Bahasa Indonesia juga menunjukkan tingkat konsistensi internal yang cukup memuaskan dengan nilai *Cronbach’s Alpha* antara 0.6 hingga 0.8 di semua skala *strength* (Lestari, 2006; Wijayanti, 2008).

Pemilihan inventori VIA-IS sebagai inventori yang digunakan sebagai contoh analisis terkait dengan isu kerentanan item-item inventori VIA-IS terhadap respons yang bersifat *socially desirable* di mana pernyataan item inventori ini diduga rentan mengundang jawaban yang tidak sebenarnya dari responden sehingga responden yang memiliki kecenderungan *social desirability* cenderung melaporkan dirinya seolah-olah memiliki karakteristik diri yang positif sesuai apa yang ingin diukur oleh inventori VIA-IS (Osin, dalam Freire, 2009). Kekhawatiran akan hal ini ternyata diwaspadai oleh Peterson dan Seligman (2004) selaku pengembang inventori VIA-IS, sehingga pengujian tingkat kerentanan item-item terhadap *social desirability* pun dilakukan dengan cara menghitung besaran korelasi antara hasil pengukuran VIA-IS dengan MCSDS. Menurut data psiko-metrik yang dilaporkan, korelasi hasil pengukuran VIA-IS dengan MCSDS menunjukkan hubungan yang tidak signifikan ($p > 0.05$) di semua skala *strength* kecuali dua skala yaitu skala yang mengukur *strength* „*prudence*’ ($r = 0.44$) dan „*spirituality*’ ($r = 0.30$).

Tabel 1. Virtue, Strength dan Jumlah Item dalam VIA-IS

Virtues	Character Strengths	Jumlah Item
<i>Wisdom</i> (Kearifan)	<i>Creativity</i> (Kreativitas)	10
	<i>Curiosity</i> (Keingintahuan)	10
	<i>Love of learning</i> (Kecintaan belajar),	10
	<i>Open-mindedness</i> (Keterbukaan pikiran)	10
<i>Courage</i> (Keteguhan hati)	<i>Perspective</i> (Perspektif)	10
	<i>Bravery</i> (Keberanian)	10
	<i>Persistence</i> (Ketekunan)	10
	<i>Integrity</i> (Integritas)	10
<i>Humanity</i> (Berperikemanusiaan)	<i>Vitality</i> (Vitalitas)	10
	<i>Kindness</i> (Kebaikan)	10
	<i>Love</i> (Cinta)	10
<i>Justice</i> (Keadilan)	<i>Social intelligence</i> (Kecerdasan sosial)	10
	<i>Citizenship</i> (Keanggotaan dalam kelompok)	10
	<i>Fairness</i> (Keadilan dan persamaan)	10
<i>Temperance</i> (Kesederhanaan)	<i>Leadership</i> (Kepemimpinan)	10
	<i>Self-regulation</i> (Regulasi diri)	10
	<i>Prudence</i> (Kebijaksanaan)	10
	<i>Humility</i> (Kerendahan hati)	10
<i>Transcendence</i> (Transendensi)	<i>Forgiveness</i> (Memaafkan)	10
	<i>Appreciative of beauty and excellence</i> (Apresiasi terhadap keindahan & kesempurnaan)	10
	<i>Hope</i> (Harapan)	10
	<i>Gratitude</i> (Bersyukur)	10
	<i>Humor</i> (Humor)	10
	<i>Spirituality</i> (Spiritualitas)	10

Dari hasil tersebut Peterson dan Seligman (2004) berusaha membuktikan bahwa sebagian besar item-item VIA-IS dianggap tidak rentan terhadap bias respons *social desirability*. Bagaimana pun juga, seperti yang telah dipaparkan sebelumnya bahwa

deteksi kerentanan item terhadap bias respons *social desirability* dengan melibatkan tes seperti MCSDS berpotensi melanggar etika penelitian karena terkesan menuduh responden berbohong (Zickar dan Drasgow, 1996).

Oleh karena itu, dalam studi ini akan dilakukan analisis terhadap kerentanan item VIA-IS dari dampak bias respons dengan menerapkan metode analisis CFA tanpa melibatkan satupun alat ukur khusus mengukur *social desirability*. Meski demikian, sebelum melakukan analisis CFA, perlu terlebih dahulu dilakukan identifikasi tingkat kerentanan item terhadap *social desirability* (SDSV) dengan menggunakan skala rating. Berpotensi atau tidaknya item-item terhadap jawaban yang *socially desirable* dilakukan berdasarkan penilaian (*judgement*) dari 10 orang narasumber (*raters*) yang memiliki latar belakang sebagai psikolog dan sarjana psikologi dan memiliki pengalaman menggunakan inventori VIA-IS serta paham mengenai konstruk yang hendak diukur.

Berdasarkan perhitungan nilai rata-rata SDSV hasil *rating* dari setiap item maka diketahui bahwa dari 240 item yang ada dalam inventori VIA-IS terdapat 25 item yang dinilai para *raters* tergolong rentan terhadap jawaban *socially desirable* (nilai SDSV ≥ 3.00). Item-item tersebut adalah item-item dari 15 skala yang mengukur *character strength* seperti *creativity*, *perspective*, *persistence*, *integrity*, *kindness*, *love*, *social intelligence*, *citizenship*, *fairness*, *self-regulation*, *humility*, *forgiveness*, *hope*, *gratitude* dan *humor* (Tabel 2).

Jika dilakukan telaah secara kualitatif, maka isi pernyataan dari 25 item yang dinilai rentan tersebut memang cenderung dapat diinterpretasi sebagai bentuk-bentuk perilaku

yang positif dan diharapkan oleh norma sosial, setidaknya bagi kultur orang Indonesia. Salah satu contoh pernyataan item yang memiliki SDSV tinggi tersebut misalnya “Saya selalu mengaku bila saya salah” (item dari skala *strength* ‘*fairness*’ yang mewakili *virtue* ‘*justice*’), “Saya selalu mendengarkan jika orang lain menceritakan masalah mereka kepada saya” (item dari skala *strength* ‘*kindness*’ yang mewakili *virtue* ‘*humanity*’) dan “Saya tidak pernah menyombongkan prestasi-prestasi saya” (item dari skala *strength* ‘*humility*’ yang mewakili *virtue* ‘*temperance*’). Secara umum perilaku yang ada pada pernyataan kedua item tersebut merupakan bentuk perilaku yang terkait dengan harapan sosial di mana sebagian besar orang diharapkan melakukannya dalam kehidupan sosial sehari-hari.

Dari kedua item contoh tersebut, tentu saja tidak ada orang yang tidak ingin dianggap sebagai orang yang tidak berani bertanggung jawab atas kesalahannya sendiri atau tidak menjadi pendengar yang baik atau menjadi orang yang terlalu membanggakan dirinya sendiri di hadapan orang lain. Oleh karena itu isi item yang mencerminkan bentuk perilaku seperti yang dicontohkan, dinilai oleh para *raters* layak dianggap sebagai item yang berpotensi akan direspons secara *socially desirable*.

Setelah diketahui item-item dari skala mana saja yang dinilai rentan, maka tahap selanjutnya, item-item tersebut akan dianalisis melalui metode CFA dan dijadikan kriteria

sebagai item yang diduga bias sehingga dapat terbukti apakah item-item tersebut memang tergolong ke dalam satu faktor yang sama yaitu faktor *social desirability*. Mengingat konstruk *strength* yang diukur oleh 25 item tersebut adalah item-item yang mewakili enam *virtue*, atau dengan kata lain setiap *virtue* memiliki *strength* yang item-itemnya dinilai rentan oleh *raters*, maka akan ter-

dapat enam studi yang masing-masing menganalisis satu model *virtue*. Selain itu, berdasarkan model hubungan antara faktor *virtue*, *character strength* dan item-item yang mengukur *strength* maka model CFA yang dianalisis akan mengikuti model *Second-Order CFA* di mana faktor *virtue* dijadikan sebagai *second-order factor* dan faktor *strength* dijadikan sebagai *first-order factor*.

Tabel 2. Virtue, Strength dan Item dengan SDSV Tinggi

Virtue	Character strength	Jumlah Item dengan SDSV Tinggi
Wisdom	<i>Creativity</i>	1
	<i>Perspective</i>	2
Courage	<i>Persistence</i>	1
	<i>Integrity</i>	2
Humanity	<i>Kindness</i>	3
	<i>Love</i>	1
	<i>Social Intelligence</i>	1
Justice	<i>Citizenship</i>	3
	<i>Fairness</i>	3
Temperance	<i>Self-regulation</i>	1
	<i>Humility</i>	1
	<i>Forgiveness</i>	1
Transcendence	<i>Hope</i>	1
	<i>Gratitude</i>	3
	<i>Humor</i>	1

Meskipun terdapat enam model *virtue* yang dianalisis secara terpisah, namun masing-masing pengujian model *virtue* memiliki tahapan pengujian hipotesis yang sama. Pengujian hipotesis dengan menilai model fit secara keseluruhan dan uji signifikansi kontribusi item pada faktor yang di-

wakilnya akan dilakukan di setiap model *virtue* dengan tiga model pengujian hipotesis yaitu: *Model 1*: ketika *social desirability* belum dilibatkan sebagai faktor (ilustrasi model pada Gambar 1); *Model 2*: ketika item-item dengan SDSV tinggi tidak dilibatkan dalam model (ilustrasi model pada

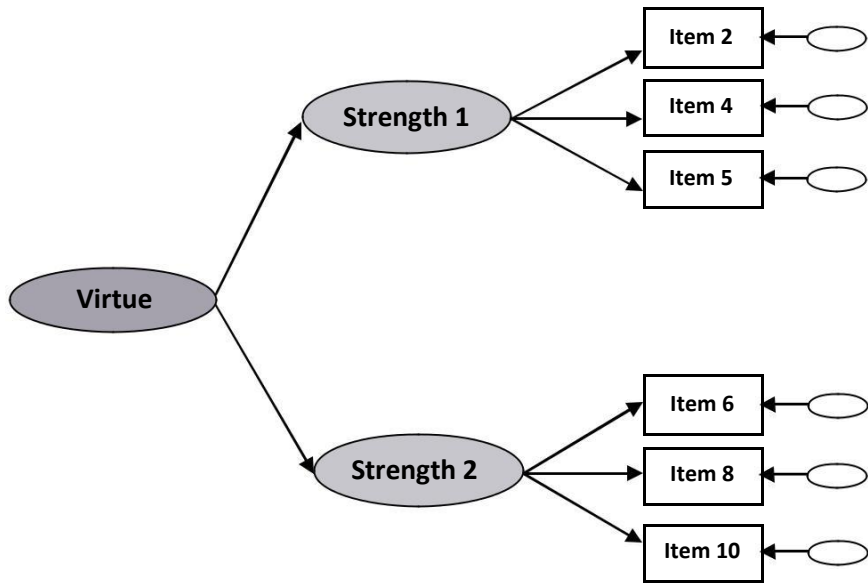
Gambar 2); dan *Model 3* atau *bi-factor model* ketika faktor *social desirability* dilibatkan sebagai salah satu faktor (ilustrasi model pada Gambar 3). Khusus untuk *Model 3*, di mana *social desirability* akan menjadi faktor lain yang dilibatkan dalam model (*bi-factor model*), maka faktor *social desirability* tidak akan memiliki garis hubungan langsung dengan faktor *virtue* melainkan berhubungan langsung dengan item-item yang dijadikan criteria sebagai item bias.

Dalam setiap model *virtue* yang dianalisis, faktor *strength* yang dijadikan *first-order factor* hanyalah faktor yang memiliki item dengan SDSV yang tinggi saja. Model 1 dapat dikatakan sebagai model pengujian validitas konstruk di mana dalam model pengujian tersebut akan diketahui apakah model yang dianalisis fit dengan data dan seluruh item terbukti mampu mengukur konstruk *strength* yang memang hendak diukur. Sedangkan pada Model 2, penilaian model fit dilakukan untuk melihat apakah model menjadi semakin fit ketika item-item dengan SDSV tinggi tidak dilibatkan. Lalu, Model 3 adalah model ketika *social desirability* dilibatkan sebagai faktor di mana seluruh item dari faktor *strength* dilibatkan namun item yang dinilai memiliki SDSV tinggi dari hasil *rating* dibebaskan mengarah selain pada faktor *strength* juga pada faktor *social desirability*.

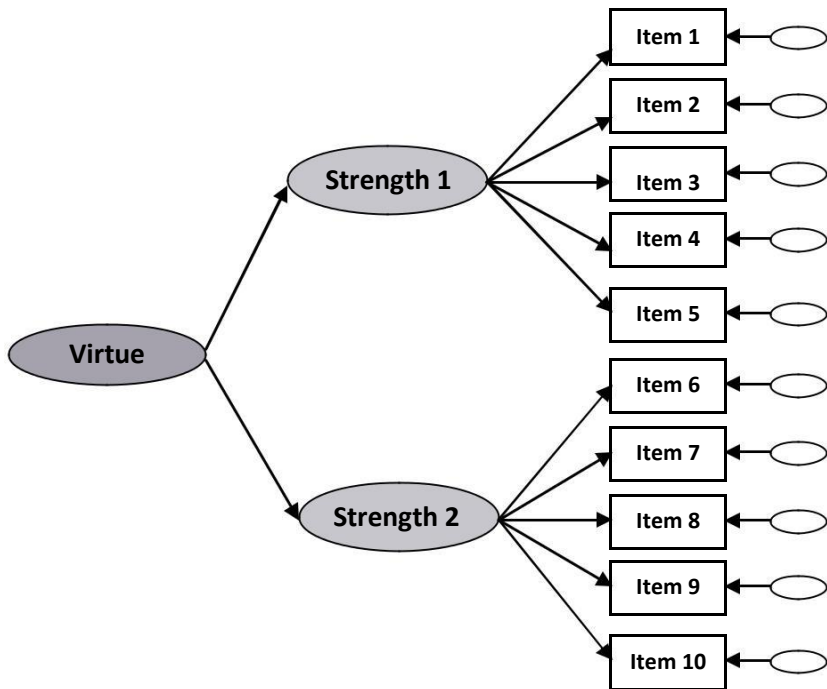
Perbandingan antar-model juga dilakukan dengan tujuan untuk melihat model mana yang paling fit di antara ketiga model tersebut. Terutama pada Model 1 dan Model 3 di mana diharapkan Model 3 jauh lebih fit dibandingkan Model 1 serta untuk membuktikan apakah item-item yang bias dapat terdeteksi.

Hasil Analisis:

Hasil penilaian model secara keseluruhan menunjukkan bahwa Model 1 dan Model 2 di setiap model *virtue* terbukti fit dengan data (RMSEA < 0.08). Jika membandingkan nilai *Chi-Square* (χ^2) antara Model 1 dan Model 2 untuk setiap model *virtue* yang dianalisis, terdapat kesamaan hasil di semua model *virtue* yaitu nilai χ^2 pada Model 2 menjadi lebih kecil dibandingkan Model 1. Semakin mengecilnya nilai χ^2 pada Model 2 dibandingkan Model 1 menandakan bahwa model menjadi semakin fit ketika item-item dengan SDSV tinggi tidak dilibatkan. Hasil uji signifikansi kontribusi tiap item pada Model 1 dan Model 2 pada masing-masing model *virtue* menunjukkan bahwa item-item pada Model 1 dan Model 2 secara keseluruhan terbukti mampu mengukur faktor *strength* yang diwakilinya ($t > 1.96$) (Ciptadi, 2011).

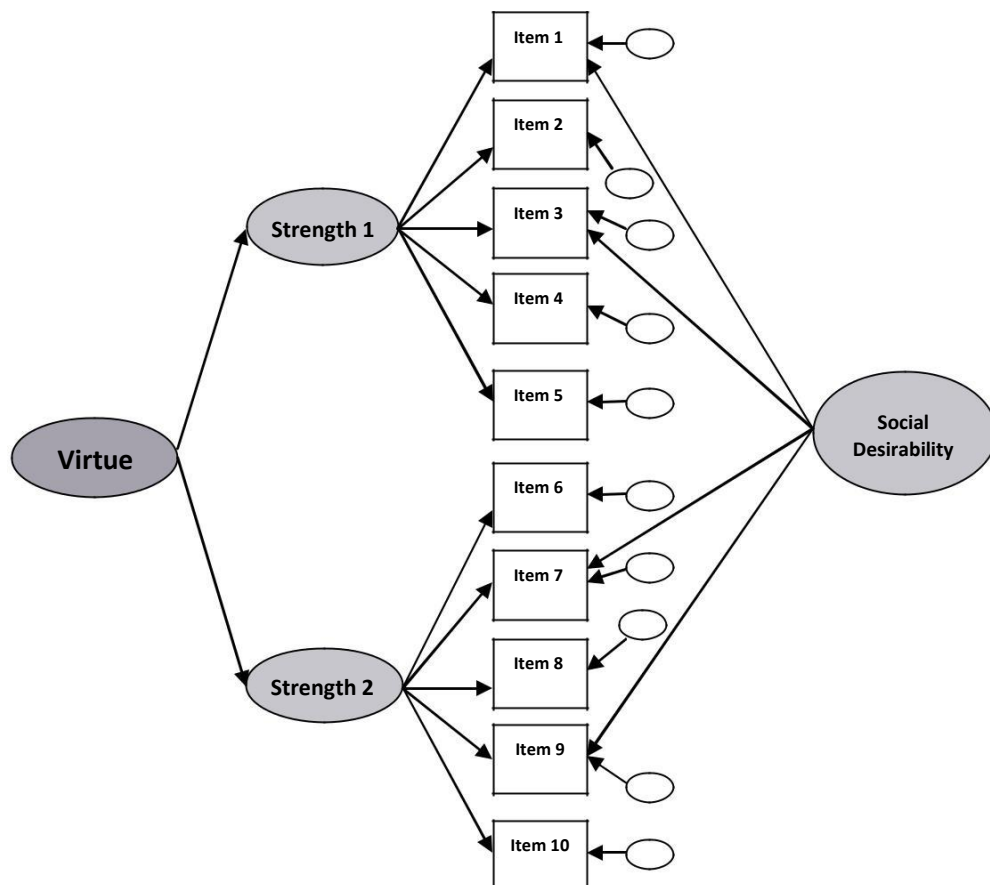


Gambar 1. Ilustrasi Model 1 (uni-dimensional model)



Gambar 2. Ilustrasi Model 2 (uni-dimensional model)

Misalkan Item 1, 3, 7 dan 9 memiliki nilai SDSV tinggi maka item-item tersebut tidak dilibatkan dalam model



Gambar 3. Ilustrasi Model 3 (*bi-factor model*)

Item 1, 3, 7 dan 9 dilibatkan kembali dalam model serta menjadi indikator bagi faktor *strength* dan *social desirability*

Setelah diketahui bahwa model menjadi semakin fit ketika item-item dengan SDSV tinggi tidak diikutsertakan maka tahapan selanjutnya adalah membandingkannya dengan model ketika *social desirability* dijadikan sebagai faktor lain dalam model. Berdasarkan penilaian model secara keseluruhan, Model 3 untuk setiap model *virtue* secara keseluruhan me-

nunjukkan nilai RMSEA < 0.08 atau model yang fit. Jika membandingkan nilai χ^2 Model 1 dengan Model 3, nilai χ^2 Model 3 pada setiap model *virtue* menunjukkan hasil yang serupa yaitu besaran nilai χ^2 Model 3 selalu lebih kecil dibandingkan nilai χ^2 Model 1 (Ciptadi, 2011). Hal ini mengindikasikan bahwa ketika faktor *social desirability* dilibatkan sebagai

faktor lain dalam model yang di-analisis, model CFA menjadi semakin fit.

Hasil uji signifikansi kontribusi tiap item pada Model 3 di setiap model *virtue* menunjukkan bahwa seluruh item terbukti mampu mengukur faktor *strength* namun untuk item-item yang dinilai memiliki SDSV tinggi dari hasil *rating*, hanya tujuh item dari total 25 item yang dijadikan kriteria sebagai item bias terbukti selain mampu mengukur faktor *strength* juga mampu mengukur *social desirability*. Hal ini ditunjukkan dari besaran *t-value* dari tujuh item tersebut yang signifikan (*t-value* > 1.96) pada kedua faktor yaitu pada faktor *strength* dan *social desirability*.

Dari 3 item skala *character strength* yang mewakili *virtue* „wisdom’ tidak ada satupun item yang terbukti signifikan selain mengukur faktor *strength* juga mengukur faktor *social desirability*. Sedangkan item-item dari skala *character strength* yang mewakili *virtue* „courage’, „humanity’, „justice’ dan „temperance’, masing-masing terbukti memiliki 1 item yang bias. Lalu dari 5 item *character strength* yang mewakili *virtue* „transcendence’ terdapat 3 item yang terbukti sebagai item yang bias terhadap *social desirability*.

Dari sejumlah item yang dijadikan kriteria sebagai item yang rentan direspon secara *socially*

desirable, item-item yang mengukur *character strength* dari *virtue transcendence* terbukti memiliki lebih banyak item yang terbukti sebagai item yang bias dibandingkan dengan item-item *character strength* lainnya. Contoh pernyataan item dari *virtue transcendence* yang terbukti bias tersebut yaitu “Saya selalu melihat segala hal dari sisi positifnya” (mengukur *character strength* „hope’), “Saya adalah orang yang sangat bersyukur” (mengukur *character strength* „gratitude’) dan “Saya mencoba untuk menikmati segala situasi” (mengukur *character strength* „humor’).

Dengan demikian, bentuk perilaku yang ada dalam pernyataan item tersebut selain mampu merepresentasikan *strength* positif tertentu yang ada dalam diri seseorang, juga mampu menstimulasi kecenderungan dalam diri seseorang untuk merespon pernyataan tersebut secara *socially desirable*. Jika melihat hasil analisis yang telah dilakukan, di mana tujuh dari 25 item yang dijadikan kriteria sebagai item bias berdasarkan hasil *rating* terbukti sebagai item yang rentan terhadap bias respons *social desirability*, dapat disimpulkan bahwa metode CFA yang diterapkan dalam analisis ini berhasil mendeteksi item-item berpotensi direspon secara *socially desirable*.

Tabel 3. Item yang Terbukti Bias

Virtue	Character strength	Jumlah Item dengan SDSV Tinggi	Jumlah Item yang Terbukti Bias*
Wisdom	<i>Creativity</i>	1	-
	<i>Perspective</i>	2	-
Courage	<i>Persistence</i>	1	1
	<i>Integrity</i>	2	-
Humanity	<i>Kindness</i>	3	1
	<i>Love</i>	1	-
	<i>Social Intelligence</i>	1	-
Justice	<i>Citizenship</i>	3	-
	<i>Fairness</i>	3	1
Temperance	<i>Self-regulation</i>	1	-
	<i>Humility</i>	1	1
	<i>Forgiveness</i>	1	-
Transcendence	<i>Hope</i>	1	1
	<i>Gratitude</i>	3	1
	<i>Humor</i>	1	1

**t*-value > 1.96 di faktor strength dan social desirability

Tabel 4. Pernyataan Item yang Terbukti Bias

Virtue	Character strength	Pernyataan Item
<i>Courage</i>	<i>Persistence</i>	Saat saya memperoleh apa yang saya inginkan, itu karena saya berusaha keras untuk mendapatkannya.
<i>Humanity</i>	<i>Kindness</i>	Saya selalu mendengarkan jika orang lain menceritakan masalah mereka kepada saya.
<i>Justice</i>	<i>Fairness</i>	Saya selalu mengaku bila saya salah.
<i>Temperance</i>	<i>Humility</i>	Saya tidak pernah menyombongkan prestasi-prestasi saya.
<i>Transcendence</i>	<i>Hope</i>	Saya selalu melihat segala hal dari sisi positifnya.
	<i>Gratitude</i>	Saya adalah orang yang sangat bersyukur.
	<i>Humor</i>	Saya mencoba untuk menikmati segala situasi.

Diskusi

Berdasarkan contoh penerapan metode CFA terhadap data hasil pengukuran inventori VIA-IS dapat disimpulkan bahwa metode CFA sangat efektif untuk mengidentifikasi item-item yang rentan terhadap bias respon *social desirability*. Dalam model CFA yang diuji, beberapa item yang dinilai rentan dari hasil *rating* terbukti membentuk satu faktor yang sama yaitu faktor *social desirability*.

Metode CFA yang diterapkan dalam studi ini dapat dijadikan sebagai metode alternatif dari sejumlah metode yang telah ada, bahkan dapat pula dikatakan sebagai pengembangan dari metode CFA yang pernah diperkenalkan oleh Ferrando (2005). Jika dibandingkan dengan metode-metode lain yang konvensional, metode CFA yang diterapkan dalam studi ini memiliki beberapa kelebihan di antaranya adalah tidak lagi dibutuhkan alat ukur yang khusus mengukur *social desirability* karena upaya deteksi tidak mengarah pada kecenderungan *social desirability* yang ada di dalam diri responden melainkan fokus pada item.

Dengan demikian maka kemungkinan terjadinya pelanggaran etika, seperti menuduh responden berbohong, dapat dihindari. Bukti empirik yang bisa diperoleh melalui metode CFA ini adalah ditemukannya item-item *self-report inventory* yang bias terhadap respon *social desirability*. Dengan demikian, para peneliti atau pengembang tes kepribadian yang menguasai metode CFA dapat

menjadikan metode CFA sebagai bagian dari kegiatan analisis item dalam tahapan konstruksi tes. Melalui metode ini, item-item yang terbukti berpotensi sebagai item yang bias dapat terdeteksi sejak dini sehingga dapat direvisi atau tidak dilibatkan dalam pengukuran. Hal ini terkait dengan upaya meningkatkan validitas *self-report* di mana metode CFA menjadi metode yang selain untuk menguji unidimensionalitas, juga mampu digunakan untuk membebaskan potensi ancaman dampak bias respon *social desirability* dari hasil pengukuran *self-report*. Jika item-item yang bias dapat terdeteksi sejak dini melalui metode CFA, maka metode CFA bisa menjadi metode yang jauh lebih efisien dibandingkan metode lainnya terutama *covariate technique* yang harus melibatkan alat ukur khusus mengukur *social desirability* serta melakukan koreksi terhadap hasil pengukuran *self-report* dari dampak bias respon tersebut. Metode CFA sebagai metode analisis item mampu meminimalisasi kemungkinan terjadinya potensi bias respon sebelum *self-report* diberikan pada responden. Hanya item-item yang tidak berisi pernyataan yang mengundang respon bersifat *socially desirable* yang tersaji dalam *self-report*.

Kualitas isi item tentunya terkait pula dengan kemampuan para pengembang tes dalam menulis item itu sendiri. Bagaimanapun juga, meski pengembang tes telah memiliki kompetensi yang baik dalam membuat item, proses analisis item melalui

CFA tetap diperlukan untuk membuktikan bahwa item tersebut tergolong sebagai item yang bias atau tidak.

Kelebihan lain dari metode CFA adalah dilakukannya analisis secara menyeluruh yaitu mulai dari analisis terhadap model secara keseluruhan hingga uji parameter dari setiap item untuk melihat kemampuan item tersebut berkontribusi pada konstruk yang hendak diukur. Pada studi ini, penilaian model secara keseluruhan dengan CFA pada dasarnya dilakukan untuk menguji validitas konstruk di mana akan diperoleh informasi mengenai unidimensionalitas dari model pengukuran serta fit atau tidaknya model ketika *social desirability* dilibatkan sebagai salah satu faktor. Oleh karena itu perlunya melakukan tiga tahapan pengujian model hipotesis seperti yang terpapar dalam contoh analisis data.

Informasi awal mengenai dampak *social desirability* pada model sudah dapat diperoleh dari perbandingan antara dua model yang keduanya belum melibatkan *social desirability* sebagai faktor. Dalam contoh analisis ditunjukkan pada pengujian Model 1 dan Model 2. Kedua model tersebut pada dasarnya menguji unidimensionalitas dari *self-report* namun pada salah satu model, item-item yang dijadikan kriteria sebagai item bias berdasarkan *rating* tidak disertakan. Jika indeks kecocokan model pada model pengujian kedua menjadi semakin fit dibandingkan model pengujian pertama maka dari hasil tersebut sudah bisa diperoleh informasi bahwa

item-item bias tersebut lebih baik tidak usah disertakan dalam *self-report*. Namun belum cukup bukti empirik yang bisa ditunjukkan mengenai kebenaran item-item tersebut tergolong dalam faktor *social desirability*. Oleh karena itu, perlunya pengujian model ketiga. Apabila *social desirability* dijadikan sebuah faktor dalam Model 3, maka diharapkan Model 3 akan menjadi lebih fit dibandingkan dengan Model 1. Dengan demikian, jika *social desirability* dimodelkan menjadi sebuah faktor maka seharusnya model tersebut menjadi lebih fit dibandingkan model dimana *social desirability* belum dilibatkan sebagai faktor. Berdasarkan contoh analisis yang telah dipaparkan, perbandingan indeks kecocokan model secara keseluruhan antara model tanpa faktor *social desirability* dengan model yang menjadikan *social desirability* sebagai faktor telah mendukung prinsip tersebut. Model menjadi lebih fit ketika *social desirability* dijadikan sebagai faktor tambahan yang dianalisis. Hasil ini juga mendukung asumsi model CFA di mana hasil pengukuran setiap *self-report* dikatakan mengandung dua hal yaitu konstruk aspek kepribadian yang hendak diukur dan *social desirability* sebagai faktor lain yang ikut terukur.

Meski berbagai kelebihan dari metode CFA telah disebutkan namun ada beberapa hal penting yang perlu diperhatikan, terutama terkait penggunaan skala *rating* untuk menentukan kriteria item yang bias. Sesuai dengan pernyataan Paulhus (1981)

bahwa valid tidaknya penilaian item-item yang dijadikan kriteria sebagai item bias melalui hasil *rating* sangat tergantung dari kredibilitas para ahli yang dilibatkan sebagai *raters*. Para *raters* sebaiknya memiliki pengalaman dan pengetahuan yang cukup mendalam mengenai konstruk yang hendak diukur *self-report*. Disamping itu, sebaiknya dilibatkan pula beberapa orang ahli atau peneliti yang menguasai tentang konsep *social desirability*. Pentingnya kredibilitas *raters* juga terkait dengan upaya mencegah terjadinya *over-rated* atau *under-rated* pada hasil *rating* tiap item. *Over-rated* bisa terjadi ketika *rater* menilai bahwa suatu item tertentu merupakan item yang bias namun sesungguhnya tidak. Sebaliknya, *under-rated* terjadi ketika *rater* menilai suatu item bukan sebagai item yang bias padahal nyatanya item tersebut bias.

Selain hal-hal yang telah disebutkan, hal lain terkait *social desirability* di luar pembahasan metode CFA namun sering didiskusikan adalah mengenai harapan responden secara pribadi terhadap hasil pengukuran *self-report*. Tujuan dan harapan responden dalam menjawab *self-report* dapat saja tidak berhubungan dengan potensi item yang bias. Jika responden bertujuan ingin menggunakan skor hasil pengukuran *self-report* sebagai sumber evaluasi diri maka kemungkinan responden untuk tidak menjawab secara *socially desirable* menjadi sangat kecil meskipun item-item pada *self-report* itu tergolong bias. Sebaliknya, jika tujuan respon-

den menjawab dilandasi kesadaran untuk melindungi citra dirinya dari penilaian orang lain maka sangat besar kemungkinannya *social desirability* terjadi. Bagaimanapun juga, meski kondisi tersebut memang perlu untuk diketahui, namun tidak pernah ada cara yang dapat dilakukan untuk memastikan apa tujuan dan harapan dari masing-masing responden ketika mengisi *self-report*. Hal ini sama saja seperti menuduh responden akan berbohong atau tidak. Oleh karena itu, terlepas dari apapun tujuan dan harapan responden dalam mengisi *self-report*, upaya mendeteksi item yang berpotensi bias melalui metode CFA tetap layak untuk dilakukan.

Singkatnya, metode CFA memiliki banyak keuntungan dibandingkan metode lain dalam mengatasi *social desirability*. Metode CFA ini pun bisa dan masih menjadi metode yang paling canggih, efektif dan aman dari tindak pelanggaran etika penelitian. Oleh karena itu metode CFA layak diperhitungkan sebagai salah satu metode untuk mendeteksi kerentanan item yang bias terhadap *social desirability*.

Daftar Pustaka

- Anastasi, Anne & Urbina, Susana. (1997). *Psychological testing*. 7th edition. Pearson Prentice Hall.
- Browne, M.W & Cudeck, R. (1993). *Alternative ways of assessing model fit*. In: Bollen, Kenneth & Long, J. Scott. (editors). (1993). *Testing structural equation model*. Sage Publication.

- Byrne, Barbara M. (1998). *Structural equation modeling with LISREL, PRELIS and SIMPLIS: Basic concept, applications and programming*. New Jersey: Lawrence Erlbaum Associates.
- Carr, Alan. (2004). *Positive psychology: The science of happiness and human strengths*. NY: Brunner-Routledge.
- Ciptadi, Bina. (2011). *Studi tentang cara mendeteksi item-item self-report inventory yang rentan terhadap bias respon social desirability: Sebuah pendekatan confirmatory factor analysis [A study on the method to detect self-report inventory items which are vulnerable to social desirability response bias: A confirmatory factor analysis approach]*. Thesis. Unpublished manuscript, UI, Depok, Indonesia.
- Cohen, Ronald & Swerdlik, Mark. (2005). *Psychological testing and assessment: An introduction to tests and measurement*. 6th edition. McGraw-Hill International.
- Domino, George & Domino, Marla. (1999). *Psychological testing: An introduction*. Prentice Hall.
- Ferrando, Pere. (2005). *Factor analytic procedures for assessing social desirability in binary items*. *Multivariate Behavioral Research*, 40 (3), 331-349. Lawrence Erlbaum Associates, Inc.
- Joreskog, K.G & Sorbom, D. (1993). *Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software International.
- Joreskog, K.G. & Sorbom, D. (2006). *LISREL 8.80 for windows*. Scientific Software International, Inc. Lincolnwood, IL.
- Lanyon, Richard & Goodstein, Leonard. (1997). *Personality assessment*. 3rd edition. NY: John Wiley & Sons, Inc.
- Leite, Walter & Cooper, Lou Ann. (2010). *Detecting social desirability bias using factor mixture models*. *Multivariate Behavioral Research*, 45: 271-293. Taylor & Francis Group, LLC.
- Lestari, M.D. (2006). *Adaptasi alat ukur values in action inventory strengths pada perawat di rumah sakit Cengkareng*. Tesis. Unpublished manuscript, Universitas Indonesia, Depok, Indonesia.
- Marlowe, D & Crowne, DP. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Osin, Evgeny. (2009). *Social desirability in positive psychology: Bias or desirable sociality?*. In Freire, Teresa (editor). (2009). *Understanding positive life, research and practice on positive psychology*. CLIMEPSI.
- Paulhus, DL. (1981). Control of social desirability in personality inventories: Principal-factor deletion. *Journal of Research in Personality*, 15, 383-388.
- Paulhus, DL. (1984). Two-component models of social desirable responding. *Journal of personality and social psychology*, 1984, Vol. 46, No. 3, 598-609.

- Paulhus, DL. (1991). Measurement and control of response bias. In JP Robinson, PR Shaver & LS Wrightsman, *Measures of personality and social psychological attitudes*, pp. 17-59, San Diego: Academic Press Inc.
- Wijayanti, Herlani. (2008). *Hubungan antara kekuatan karakter dan kebahagiaan pada suku jawa*. Skripsi (*Undergraduate Thesis*). Unpublished manuscript, UI, Depok, Indonesia.
- Peterson, C. & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Washington DC: APA.
- Zickar, M.J & Drasgow, F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied Psychological Measurement*, Vol. 20, No.1. 71-87.

