

PERANAN BUTIR UNFAVORABEL DALAM MENGHASILKAN DIMENSI BARU DALAM PENGUKURAN PSIKOLOGI

Wahyu Widhiarso
Fakultas Psikologi - Universitas Gadjah Mada
wahyu_psy@ugm.ac.id

Abstrak

Meskipun penggunaan butir unfavorabel dalam skala psikologi disarankan dalam banyak referensi, namun beberapa penelitian menemukan rendahnya kualitas properti psikometris butir tersebut. Penelitian ini bertujuan untuk mereplikasi penelitian tersebut dengan menekankan pada validitas faktorial dan kriteria butir-butir unfavorabel. Data penelitian ini didapatkan dari pengukuran motivasi berprestasi pada siswa SMA (N=2986) dengan menggunakan Skala Motivasi Berprestasi. Hasil pengujian model pengukuran melalui pendekatan analisis faktor eksploratori menyimpulkan bahwa model 3 faktor menjelaskan data penelitian dengan tepat. Namun demikian, salah satu dari faktor di model tersebut berisi semua butir unfavorabel. Hasil ini menunjukkan bahwa faktor yang muncul dari analisis lebih diakibatkan oleh kesamaan metode dibanding dengan indikator. Hasil pengujian korelasi antara butir dan ketiga faktor hasil analisis sebelumnya menunjukkan hasil bahwa tidak ada perbedaan nilai validitas kriteria antara butir unfavorabel dan favorabel. Hasil yang sama juga muncul dari korelasi faktor dan kriteria. Kesimpulan yang didapatkan dari penelitian ini adalah melibatkan butir-butir unfavorabel dalam skala cenderung mengganggu validitas faktorial namun tidak pada validitas kriteria skor skala.

Kata Kunci : Butir Unfavorabel, Validitas Faktorial, Validitas Kriteria,

Although the use unfavorable item on the psychological scale was suggested by numerous authors, however, some studies found poor-quality psychometrics property on such items. This study aimed to replicate those studies with emphasis on factorial and criterion validity of unfavorable items. Research data was obtained from the measurement of achievement motivation among high-school students (N=2986) using Achievement Motivation Scale. Analysis results from testing measurement model by exploratory factor analysis approach suggest that three factors model was fitted with the data. However, one of these factors contain all unfavorable items. This finding indicated that the emerging factors from analysis were caused by the similarity of method rather than indicator content. Correlation between items/factors and criterion show no criterion validity differences between favorable and unfavorable items. Similar results were also shown from correlation between three factors and the criterion. This study concludes that using unfavorable items on the scale tend to interrupt factorial validity but not criterion validity.

Keywords : Unfavorable items, Factorial Validity, Criterion Validity.

Setelah mengamati beberapa hasil penelitian penyusunan skala psikologi, penulis menemukan beberapa temuan yang menarik untuk didiskusikan. Temuan tersebut adalah (a) banyaknya butir yang tidak lolos dalam seleksi butir, (b) tidak konsistensinya sebaran butir pada kisi-kisi teoritik skala dengan hasil pengujian secara empirik; (c) adanya korelasi antar eror (*error correlated*) dalam model pengukuran yang diuji oleh skala tersebut; dan (d) rendahnya nilai ketepatan model pengukuran ketika data dianalisis dengan menggunakan pendekatan analisis faktor konfirmatori (Widhiarso & Kožený, 2013). Penulis menduga bahwa salah penyebab keganjilan hasil penelitian tersebut adalah karakteristik butir yang dilibatkan dalam skala.

Berikut ini contoh-contoh penelitian tersebut. Pengembangan Skala Kompetensi Komunikasi Interpersonal yang dilakukan Fitri dan Zulkaida (2008) mendapatkan 77 persen butir yang gugur dalam proses analisis adalah butir *unfavorabel*. Tidak konsistennya antara kisi-kisi butir yang didasarkan oleh teori dan data di lapangan dilaporkan oleh Haynes, Miles, dan Clements (2000). Adanya kejanggalan sebaran butir dari hasil analisis faktor dapat disebabkan oleh eror varians yang dihasilkan oleh butir-butir *unfavorabel*.

Sementara itu dari model yang dikembangkan oleh Yamamura dkk. (2005) dan Pribadi (2010) yang di dalamnya berisi korelasi antar sesatan juga dapat disebabkan oleh butir *unfavorabel*. Korelasi antar eror menunjukkan bahwa varians unik yang tidak dijelaskan oleh model membentuk varians sistematis yang memiliki makna. Padahal seharusnya varians unik ini bersifat acak dan tidak memiliki korelasi dengan variabel apapun di luar model. Chen dkk. (2010) menemukan bahwa korelasi antar eror tersebut banyak terjadi pada butir yang sama-sama bersifat *unfavorabel*, sehingga dinamakan dengan 'efek metode'. Dua arah butir yaitu *favorabel* dan *unfavorabel* dilihat sebagai metode berbeda sehingga efek besarnya varians di luar model dilihat sebagai akibat dari metode yang berbeda.

Butir *unfavorabel* adalah butir yang memiliki karakteristik yang tidak mendukung atribut ukur skala. Keberadaan butir sudah ada sejak masa-masa awal pengembangan teori pengukuran psikologi melalui skala (Anastasi, 1982). Rensis Likert seperti yang dikutip dari Gliem dan Gliem (2003) menjelaskan bahwa skala *Likert* memuat seperangkat butir yang terdiri pernyataan yang jumlahnya seimbang antara yang mendukung dan tidak mendukung (*unfavorable*) terhadap objek sikap. Individu diminta untuk merespons setiap pernyataan tersebut berdasarkan derajat persetujuan atau ketidaksetujuan mereka sendiri. Tujuan peneliti melibatkan butir berlawanan arah ini kemudian meluas menjadi bagian dari upaya untuk mereduksi respons persetujuan monoton (DeVellis, 2011). Adanya butir memiliki arah terbalik (butir *unfavorabel*) diharapkan mengondisikan responden untuk membaca tiap butir pernyataan lebih seksama dan tidak asal-asalan dalam memberikan tanggapan.

Schriesheim dan Eisenbach (1995) menjelaskan bahwa secara implisit penggunaan butir unfavorabel didasarkan pada asumsi bahwa: (a) bias respons tersebut adalah ancaman serius terhadap validitas instrumen, (b) butir unfavorabel bisa dipakai jika tidak konsekuensi negatif yang serius misalnya butir tersebut tidak mengganggu properti psikometris instrumen, dan (c) kualitas psikometrik penyekoran pada favorabel dan unfavorabel memiliki kesetaraan. Poin menjelaskan bahwa dua jenis penegasian harus menghasilkan skor yang sama. Misalnya antara pasangan butir "Saya senang" dan "Saya tidak sedih" harus menghasilkan kualitas psikometri yang sama dengan pasangan butir "Saya tidak senang" dan "Saya sedih".

Banyak istilah yang dipakai untuk menjelaskan butir yang mengukur atribut psikologi pada arah yang berlawanan, misalnya tak mendukung (*unfavorable*), kalimat/penyekoran terbalik (*reverse-worded/scored*), butir arah negatif (*Negatively Worded*) (Chen et al., 2010), butir arah berlawanan (*opposite direction*) (Spector, Van Katwyk, Brannick, & Chen, 1997). Untuk menyederhanakan penggunaan kata dalam kalimat maka dalam tulisan ini butir yang memiliki karakteristik tersebut dinamakan dengan *butir unfavorabel*.

Melibatkan butir favorabel dan unfavorabel dalam skala harus memenuhi asumsi bahwa keduanya harus mengikuti pola bipolar (Friedman & Amoo, 1999). Dalam bidang psikologi dikenal berbagai jenis konstruk berdasarkan arah ukurnya. Jika dua konstruk merefleksikan sesuatu yang berlawanan dinamakan dengan konstruk bipolar (*bipolar construct*) sedangkan jika kedua konstruk adalah independen maka dinamakan dengan konstruk ortogonal (*orthogonal construct*) (Widhiarso, 2011). Jika dua konstruk memiliki hubungan bipolar, maka keduanya dapat diukur dalam satu alat ukur. Misalnya pusat kendali (*locus of control*) eksternal dan internal adalah konstruk bipolar sehingga diukur dalam satu variabel. Satu konstruk memuat butir favorabel sedangkan konstruk lainnya adalah unfavorabel (Allen & Lauterbach, 2007). Di sisi lain, jika dua konstruk adalah ortogonal maka keduanya diukur oleh dua alat ukur yang berbeda. Dalam domain afektif, faktor arah butir seringkali ditafsirkan mencerminkan konstruksi atribut ganda yang bersifat ortogonal. Misalnya, butir yang memanifestasikan emosi positif dan negatif ditafsirkan sebagai konstruk atribut psikologis yang independen yang diukur oleh dua alat ukur yang berbeda (Watson, Clark, & Tellegen, 1988). Dari penjelasan ini dapat disimpulkan bahwa skor yang dihasilkan butir favorabel dan unfavorabel, sebelum dilakukan pembalikan skor, diharapkan memiliki korelasi yang negatif.

Ketika asumsi bipolar tersebut tidak terpenuhi maka butir *favorabel* dan unfavorabel mengukur konstruk yang terpisah. van de Looij-Jansen dkk. (2011) menemukan bahwa penggunaan butir *unfavorabel* menjadi bermasalah karena cenderung mengukur atribut yang berbeda. Rodebaugh, Woods, dan Heimberg (2007) menemukan bahwa butir favorabel dan

unfavorabel pada pengukuran kecemasan sosial mengukur atribut yang berbeda. Dengan mengorelasikan skor kecemasan sosial dengan lima faktor kepribadian ditemukan temuan yang janggal. Butir-butir unfavorabel pengukuran kecemasan sosial memiliki korelasi yang lebih tinggi pada faktor ekstraversi dibanding dengan faktor neurotisisme. Hal ini berbeda dengan butir-butir favorabel yang memiliki korelasi lebih tinggi pada faktor neurotisisme. Temuan mereka membuktikan bahwa butir favorabel dan unfavorabel dapat mengukur atribut yang berbeda. Akibatnya, properti psikometris (i.e. validitas konvergen) yang dihasilkan butir unfavorabel lebih rendah dibanding butir favorabel.

Rodebaugh dkk. (2006) menguji model pengukuran pada skala kecemasan sosial yang memuat butir favorabel dan unfavorabel. Dari berbagai model yang diterapkan, model dua dimensi memiliki nilai ketepatan model lebih tinggi dibanding model unidimensi. Pada model dua dimensi, dimensi satu berisi butir-butir favorabel dan pada dimensi dua berisi butir-butir unfavorabel. Pada model satu dimensi, nilai muatan faktor butir-butir unfavorabel memiliki nilai muatan faktor yang rendah. Temuan ini menjelaskan bahwa butir unfavorabel mengganggu pengukuran kecemasan sosial. Wei dkk. (2010) juga menemukan hal yang relatif sama. Analisis faktor eksploratori yang diaplikasikan pada skala pengatasan masalah (*coping*) yang menghasilkan penggerombolan butir-butir unfavorabel dalam dua faktor dari lima faktor yang dihasilkan. Kedua faktor tersebut dinamakan dengan faktor resistansi dan lepasan yang merupakan faktor yang bersifat unfavorabel pada perilaku coping. Hasil pengujian model pengukuran yang mereka lakukan menunjukkan bahwa model pengukuran dengan memisahkan favorabilitas butir (*method factor*) memiliki ketepatan model lebih baik dibanding dengan model lainnya. Dua kasus ini menunjukkan bahwa butir-butir unfavorabel memiliki potensi untuk mengukur atribut tujuan ukur atau membentuk kelompok butir terpisah dengan butir favorabel.

Tujuan penelitian terbagi menjadi dua hal. Pertama, penelitian ini membandingkan properti psikometris antara pengukuran dengan menggunakan sub-skala yang berisi butir-butir favorabel dan unfavorabel. Properti psikometris yang diidentifikasi adalah korelasi butir total dan muatan faktor dalam model pengukuran. Kedua, penelitian ini menjawab pertanyaan apakah butir favorabel dan unfavorabel memiliki validitas kriteria yang setara dalam mengukur motivasi siswa. Pertanyaan ini dijawab dengan mengorelasikan skor total dari masing-masing jenis butir dengan skor yang didapatkan dari pengukuran motivasi yang sudah valid.

Metode

Responden

Responden penelitian adalah siswa kelas 2 SMA dari 30 sekolah yang berjumlah 3.000 orang dari 30 kota di Indonesia yang terbagi dalam 15 propinsi. Propinsi tersebut adalah Banten,

Bengkulu, D.I. Yogyakarta, Jawa Tengah, Jawa Timur, Kalimantan Timur, Lampung, Nusa Tenggara Barat, Sulawesi Selatan, dan Sumatera Selatan. Teknik pemilihan sampel yang dipakai adalah teknik purposif (non-acak) dengan mempertimbangkan proporsi lokasi sekolah (urban & sub urban). Dari jumlah tersebut data dari 2.986 orang dianalisis karena merespons alat ukur yang diberikan secara lengkap. Proporsi jenis kelamin responden adalah 36 persen laki-laki dan 63 persen perempuan.

Alat Ukur

Alat ukur yang dipakai dalam penelitian ini adalah Skala Motivasi Berprestasi yang dikembangkan oleh Hadjam dkk. (2004). Skala ini terdiri dari 40 butir pernyataan dengan menggunakan model Likert yang menyediakan empat opsi respons, dari sangat sesuai hingga sangat tidak sesuai. Responden diminta untuk melengkapi skala dengan menilai kesesuaian butir pernyataan dengan apa yang mereka alami dan rasakan. Skala ini terbagi menjadi tiga dimensi, yaitu dimensi prinsip dan pengangan, pengatasan hambatan dan aktivitas. Reliabilitas konsistensi internal pengukuran skala ini yang diestimasi pada masing-masing faktor menghasilkan nilai alpha yang cukup tinggi, yaitu 0,72, 0,82 dan 0,79. Dalam penelitian ini, instrumen pengukuran motivasi berprestasi yang dipakai sebagai kriteria untuk pengujian validitas adalah Skala Motivasi Berprestasi yang dikembangkan oleh Ancok dan Faturochman (1987). Skala ini memuat 42 butir yang menggunakan model pilihan ganda.

Tenik Analisis Data

Prosedur analisis data pada tulisan ini dibagi menjadi dua tahap. Tahap pertama adalah menguji validitas faktorial yang dilakukan dengan mengidentifikasi struktur faktor skala. Prosedur yang dipakai adalah analisis faktor eksploratori dengan menggunakan pendekatan pemodelan persamaan struktural (*exploratory structural equation modelling/ESEM*) (Asparouhov & Muthén, 2009). Dalam penelitian ini berbagai model pengukuran diuji dari satu dimensi hingga tiga dimensi. Dari berbagai model tersebut model yang memiliki nilai ketepatan tertinggi dan memenuhi kriteria ditetapkan model yang menjelaskan data dalam penelitian ini. Muatan faktor pada tiap butir diestimasi dan dikaitkan dengan konten butir tersebut. Tahap kedua adalah menguji validitas kriteria dengan mengorelasikan tiap butir dan faktor skala dengan kriteria. Faktor skala memuat total skor butir berdasarkan sebaran butir yang didapatkan dari analisis tahap sebelumnya. Analisis data pada tulisan ini dilakukan dengan menggunakan Program MPLUS 6.1

Hasil

Statistik Deskriptif

Rerata dan deviasi standar baik yang bersumber dari distribusi skor responden dan distribusi skor teoritik skala dipaparkan pada Tabel 1. Distribusi skor responden langsung didapatkan dari pengukuran sedangkan distribusi teoritik dari perkiraan distribusi di populasi hipotetik. Distribusi skor hipotetik ini didapatkan dari penghitungan yang memanfaatkan jumlah butir skala dan rentang penyekorannya. Sebagai contoh, pada Faktor-1 memuat 14 butir dengan penyekoran yang bergerak antara 1 hingga 4. Dengan demikian secara teoritik skor minimal yang dapat diperoleh responden adalah 14 (1 x 14 butir) dan skor maksimalnya adalah 56 (4 x 14 butir). Rerata teoritik adalah nilai tengah dari skor minimal dan maksimal sedangkan deviasi teoritik adalah rentang skor dibagi 6 yang mengikuti deviasi standar kurva normal.

Tabel 1. Deskripsi Statistik Skor Pengukuran Motivasi Belajar

Faktor	Jumlah Butir	Skor Minimal	Skor Maksimal	Rerata	Deviasi Standar
Faktor-1	14	20 (14)	55 (56)	45,99 (21)	4,09 (5,5)
Faktor-2	11	4 (11)	44 (44)	30,08 (16,50)	4,82 (5,5)
Faktor-3	15	24 (15)	60 (60)	43,56 (22,5)	4,74 (7,5)

Keterangan : Angka di dalam kurung menunjukkan statistik deskriptif yang bersumber dari distribusi teoritik/hipotetik

Dari hasil perbandingan antara distribusi empirik dan teoritik didapatkan bahwa rerata skor responden cenderung tinggi karena melebihi rerata teoritik. Di sisi lain, keragaman skor responden mendekati ideal karena nilainya mendekati nilai deviasi standar hipotetik.

Pengujian Model

Dari tiga model yang dieksplorasi melalui ESEM didapatkan hasil bahwa Model 3 Faktor memiliki nilai ketepatan yang lebih tinggi dibanding dengan model lainnya. Nilai indeks ketepatan komparatif (CFI) dan nilai indeks Tucker-Lewis (TLI) di atas kriteria yang dianjurkan para ahli, yaitu di atas 0.9. Nilai akar kuadrat residu (RMSEA) berada rentang yang dianjurkan, yaitu di bawah 0.08 (Tabel 2). Nilai kai-kuadrat yang dihasilkan oleh model menghasilkan nilai yang signifikan, namun nilai ini kurang tepat ketika diterapkan pada data dengan ukuran sampel yang besar (Schumacker & Lomax, 2004).

Tabel 2. Perbandingan Ketepatan Model

Alat Ukur		χ^2	Sig	CFI	TLI	RMSEA
Skala Motivasi	Model 1 Faktor	14807,60	p<0,01	0,69	0,67	0,08
	Model 2 Faktor	7667,23	p<0,01	0,46	0,82	0,06
	Model 3 Faktor	4194,93	p<0,01	0,92	0,91	0,04

Analisis faktor ini menghasilkan informasi mengenai faktor serta butir yang memanifestasikan faktor tersebut (Tabel 3). Dari hasil tersebut tampak bahwa butir-butir unfavorabel mengelompok menjadi satu faktor tersendiri. Ada satu butir unfavorabel yaitu butir ms-19 yang tidak masuk dalam ketiga faktor karena memiliki nilai muatan faktor di bawah 0,3 pada ketiga faktor. Namun demikian butir ini masuk dalam Faktor-1 karena nilai muatan faktor yang paling tinggi berada di Faktor-1. Kesimpulan yang didapatkan dari hasil analisis ini adalah butir-butir unfavorabel membentuk faktor baru.

Pengujian Validitas Kriteria

Uji validitas dilakukan dengan mengorelasikan skor butir-butir serta faktor skala dengan skor kriteria. Hasil korelasi menunjukkan bahwa nilai korelasi antara butir dengan kriteria baik pada butir favorabel dan unfavorabel tidak memiliki perbedaan. Rerata validitas kriteria butir favorabel (Faktor-1 & 2) adalah 0,42 dan 0,35. sedangkan pada butir unfavorabel adalah 0,31 (Tabel 3). Hasil korelasi antara Faktor-2 (yang berisi butir-butir unfavorabel) dengan kriteria juga setara dengan dengan Faktor-1 dan Faktor-3 (yang berisi butir-butir favorabel). Validitas kriteria yang dihasilkan oleh ketiga faktor ini sama-sama memenuhi kriteria validitas, yaitu di atas 0,3 (Tabel 4) (Azwar, 2007).

Tabel 3. Muatan Faktor antar Butir

Faktor	No.	Indikator Butir	F1	F2	F3	Validitas
1	ms-34	Menilai kegagalan sebagai penyemangat	0,36		0,61	0,32
	ms-35	Menilai hidup sebagai perjuangan			0,61	0,28
	ms-17	Berusaha bekerja optimal	0,43		0,58	0,31
	ms-33	Meningkatkan kemampuan	0,31		0,57	0,30
	ms-04	Belajar dari pengalaman			0,54	0,30
	ms-02	Menilai kegagalan bukan hambatan	0,38		0,53	0,33
	ms-13	Mempunyai prinsip dan pendirian	0,41		0,52	0,31
	ms-09	Menilai kegagalan sukses tertunda			0,51	0,29
	ms-29	Mengambil inisiatif jika mengalami jalan buntu	0,31		0,51	0,29
	ms-26	Menikmati aktivitas dilakukan	0,31		0,47	0,30
	ms-28	Dapat hidup dengan keterbatasan			0,39	0,26
	ms-39	Mempertimbangkan sebelum memutuskan			0,35	0,25
	ms-05	Orang lain sebagai mitra belajar			0,31	0,23
	ms-19	Tidak senang jika dikasihani (R)				0,23
	2	ms-37	Tidak mantap dengan diri (R)	0,34	0,68	
ms-21		Tidak berdaya menghadapi tugas berat (R)	0,31	0,65		0,33
ms-14		Tidak berdaya menghadapi masalah (R)		0,64	0,35	0,28
ms-36		Mudah cemas (R)		0,60		0,23

	ms-11	Kesulitan mengambil keputusan (R)		0,59		0,29
	ms-12	Menyusahkan orang lain (R)	0,36	0,59		0,28
	ms-06	Tidak berdaya mengatasi situasi sulit (R)		0,58		0,27
	ms-25	Merasa tidak berarti (R)		0,57	0,39	0,28
	ms-10	Merasa Tidak aman di lingkungan baru (R)		0,52		0,25
	ms-30	Sulit beradaptasi pada lingkungan baru (R)		0,52		0,27
	ms-38	Tidak puas dengan hasil kerja (R)		0,50		0,28
3	ms-20	Menilai keberhasilan adalah usaha pribadi	0,70			0,26
	ms-22	Dapat memanfaatkan kemampuan	0,65			0,28
	ms-15	Menilai kegagalan karena usaha diri	0,63			0,25
	ms-08	Pekerja keras dan pantang menyerah	0,55	0,35	0,42	0,35
	ms-07	Berusaha menyelesaikan sesuai target	0,53	0,31		0,29
	ms-16	Bersikap tegas	0,52	0,33	0,42	0,31
	ms-27	Tidak menunda-nunda pekerjaan	0,52	0,34		0,29
	ms-18	Menyelesaikan pekerjaan dengan baik	0,50		0,35	0,27
	ms-23	Menghasilkan sesuatu secara mandiri	0,48			0,22
	ms-24	Tidak pernah putus asa	0,48	0,34	0,44	0,31
	ms-03	Menghadapi masalah dengan tepat	0,47		0,33	0,28
	ms-40	Tidak mudah terpengaruh orang lain	0,44	0,35		0,28
	ms-01	Tidak tergantung pada orang lain	0,43			0,27
	ms-31	Mampu membaca peluang	0,43		0,39	0,30
	ms-32	Memenuhi kebutuhan diri sendiri	0,34			0,21

Keterangan : Nilai muatan faktor di bawah 0.3 tidak ditampilkan dalam tabel. Nilai validitas didapatkan dari korelasi antara butir dengan kriteria. Angka yang dicetak miring menunjukkan nilai muatan faktor yang dijadikan dasar memasukkan butir dalam faktor. Tanda (R) menunjukkan butir unfavorabel.

Tabel 4. Validitas Kriteria Butir

		1	2	3	4	5
1	Skor Total Skala	-	0,72	0,74	0,82	0,47
2	- Faktor-1		-	0,25	0,45	0,42
3	- Faktor-2			-	0,39	0,31
4	- Faktor-3				-	0,35
5	Kriteria					-

Keterangan : Semua korelasi pada matriks signifikan pada taraf 1% ($p < 0,01$)

Diskusi

Ada dua temuan dalam penelitian ini. Pertama, butir-butir unfavorabel membentuk faktor baru yang terpisah dari butir-butir favorabel. Kedua, butir-butir unfavorabel memiliki validitas yang sama dengan butir-butir favorabel. Hasil ini menunjukkan bahwa melibatkan butir unfavorabel dalam skala mengganggu validitas faktorial skor tetapi tidak mengganggu validitas kriteria. Hasil penelitian tentang munculnya faktor baru sejalan dengan penelitian-penelitian sebelumnya (Rodebaugh et al., 2007; van de Looij-Jansen et al., 2011). Sebaliknya, hasil penelitian ini yang menemukan tingginya nilai validitas butir unfavorabel berbeda dengan beberapa penelitian lain (Rodebaugh et al., 2006).

Dilihat dari struktur faktor yang dihasilkan dari analisis, makna faktor memuat butir-butir unfavorabel sulit untuk dijelaskan. Indikator yang dilibatkan dalam Faktor-2 tersebut sulit untuk diinterpretasikan karena mencakup indikator motivasi yang luas. Spector (1997) menamakan faktor seperti ini adalah faktor artifaktual (*artifactual*) yang dapat diartikan sebagai faktor akibat dari ketidakakuratan pengukuran atau faktor muncul atau teramati akan tetapi sebenarnya tidak ada. Ketika mendapati hal yang serupa, beberapa peneliti menginterpretasikan ini sebagai faktor efek metode (Percy, McCrystal, & Higgins, 2008; Timothy A, 2003) atau nama konstruk yang memiliki arah negatif. Contoh penelitian yang memberikan nama berarah negatif tersebut antara lain Hwang dkk. (2000) yang memberikan nama konflik teman atau keluarga untuk butir-butir unfavorabel dalam skala dukungan sosial dari teman dan keluarga. Wei dkk. (2010) juga melakukan hal yang serupa dengan memberikan nama resistensi dan lepasan pada butir-butir Skala Pengatasan Masalah. Berdasarkan pengalaman para peneliti tersebut Faktor-2 dalam penelitian diinterpretasikan sebagai faktor hambatan karena memanifestasikan indikator yang menunjukkan motivasi yang rendah.

Beberapa ahli menjelaskan beberapa kemungkinan munculnya faktor artifaktual butir-butir unfavorabel. Kemungkinan tersebut dapat diakibatkan oleh faktor proporsi butir unfavorabel yang cukup besar. Proporsi butir unfavorabel dalam jumlah yang kecil kemungkinan tidak akan memunculkan faktor baru. Hal ini dibuktikan oleh evaluasi struktur faktor skala yang konsisten dengan teori pendukungnya (Doran, Aldridge, Roesch, & Myers, 2011; GÜthlin & Walach, 2007). Selain faktor skala, munculnya faktor artifaktual ini juga dapat dikarenakan responden. Penelitian menunjukkan bahwa kurangnya kemampuan responden dalam memahami butir unfavorabel menyebabkan munculnya fenomena ini (Corderly & Sevastos, 1993). Selain kemampuan responden untuk memahami, kecerobohan dalam merespons butir juga menjadi faktor pendukung (Schmitt & Stults, 1985). Pendapat ini juga sama dengan apa yang dikatakan oleh Schreisheim dkk. (1991) yang menemukan bahwa butir unfavorabel cenderung kurang valid dibanding dengan favorabel yang disebabkan karena responden kesulitan untuk menafsirkan indikator butir.

Tingginya validitas kriteria pada butir-butir unfavorabel pada penelitian dapat dijelaskan melalui karakteristik indikator yang diukur butir unfavorabel. Butir unfavorabel pada skala penelitian ini memiliki sifat kebalikan (bipolar) dari butir favorabel. Pengujian homogenitas butir-butir dalam skala melalui uji konsistensi internal menghasilkan nilai yang memuaskan. Homogenitas ini menunjukkan kesetaraan domain ukur yang dibuktikan dengan kesetaraan nilai validitas kriteria semua butir dalam skala.

Berdasarkan temuan dari penelitian ini maka butir unfavorabel yang dilibatkan pada skala psikologi perlu dibuat dengan lebih seksama. Indikator yang diukur oleh butir unfavorabel harus benar-benar merepresentasikan rendahnya level atribut ukur dan bukan mengukur atribut lain. Hasil penelitian ini tetap mendukung gagasan Spector dkk. (1997) yang mengatakan bahwa meskipun butir unfavorabel berpotensi menghasilkan faktor baru, namun butir ini dapat dilibatkan jika diperlukan. Pernyataan ini juga didukung oleh Malhotra (2006) mengatakan bahwa butir unfavorabel tetap perlu dilibatkan meski pada penerapannya jumlahnya dapat lebih sedikit dibanding butir favorabel. Proporsi butir unfavorabel yang cukup banyak terbukti cenderung menghasilkan faktor baru. Oleh karena itu rekomendasi yang diberikan adalah agar penyusun skala melibatkan butir unfavorabel dalam jumlah yang optimal sesuai dengan tujuannya, yaitu mereduksi munculnya respons penyetujuan pernyataan secara monoton. Rekomendasi ini tidak berlaku pada skala sikap karena hasil penelitian ini dan penelitian-penelitian yang menemukan hasil yang sama dengan penelitian ini mengukur objek non sikap (ciri kepribadian).

DAFTAR PUSTAKA

- Allen, B., & Lauterbach, D. (2007). Personality characteristics of adult survivors of childhood trauma. *Journal of Traumatic Stress, 20*(4), 587-595.
- Anastasi, A. (1982). *Psychological testing*. New York: Macmillan.
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 397-438.
- Azwar, S. (2007). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Chen, Y.-H., Gobiolf, G. R., & Dedrick, R. F. (2010). Factorial invariance of a chinese self-esteem scale for third and sixth grade students: Evaluating method effects associated with positively and negatively worded items *The International Journal of Educational and Psychological Assessment, 6*(1), 21-35.
- Cordery, J. L., & Sevastos, P. P. (1993). Responses to the Original and Revised Job Diagnostic Survey - Is Education a Factor in Responses to Negatively Worded Items. *Journal of Applied Psychology, 78*(1), 141-143.
- DeVellis, R. F. (2011). *Scale development: Theory and applications*. Newbury Park: SAGE Publications, Inc.
- Doran, N., Aldridge, A., Roesch, S., & Myers, M. (2011). Factor structure and invariance of the Behavioral Undercontrol Questionnaire: Testing for differences across gender and

- ethnicity. *European Journal of Psychological Assessment*, 27(3), 145-152. doi: 10.1027/1015-5759/a000060
- Fitri, N., & Zulkaida, A. (2008). *Contribution of empathy to interpersonal communication competency level one students on*. (Skripsi), Fakultas Psikologi Universitas Gunadarma, Jakarta.
- Friedman, H. H., & Amoo, T. (1999). Rating the rating scales. *Journal of Marketing Management*, 9(3), 114-123.
- Gliem, J. A., & Gliem, R. R. (2003). Calculation, interpreting and reporting Cronbach's alpha reliability coefficient for Likert-Type Scales. *Midwest Research to Practice Conference in Adult, Continuing and Community Education*, 82-88.
- Güthlin, C., & Walach, H. (2007). MOS-SF 36: Structural Equation Modeling to Test the Construct Validity of the Second-Order Factor Structure. *European Journal of Psychological Assessment*, 23(1), 15-23. doi: 10.1027/1015-5759.23.1.15
- Hadjam, M. N. R., Widhiarso, W., & Widiana, H. S. (2004). *Studi Piloting Kompetensi Siswa Di Indonesia*. Laporan Penelitian. Jakarta: Departemen Pendidikan Nasional.
- Haynes, C. A., Miles, J. N. V., & Clements, K. (2000). A confirmatory factor analysis of two models of sensation seeking. *Personality and Individual Differences*, 29(5), 823-839. doi: 10.1016/s0191-8869(99)00235-4
- Hwang, W.-C., Chun, C.-A., Kurasaki, K., Mak, W., & Takeuchi, D. T. (2000). Factor Validity of Scores on a Social Support and Conflict Measure among Chinese Americans. *Educational and Psychological Measurement*, 60(5), 808-816. doi: 10.1177/00131640021970790
- Malhotra, N. K. (2006). Questionnaire design and scale development. In R. Grover & M. Vriens (Eds.), *The handbook of marketing research: Uses, misuses and future advances*. Thousand Oaks, CA: Sage Publications, Inc.
- Percy, A., McCrystal, P., & Higgins, K. (2008). Confirmatory factor analysis of the Adolescent Self-Report Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment*, 24(1), 43-48. doi: 10.1027/1015-5759.24.1.43
- Pribadi, A. A. (2010). *Pengaruh kualitas layanan mobile value added service terhadap kepuasan dan loyalitas pelanggan operator selular GSM*. Tesis. Fakultas Teknik Universitas Indonesia, Jakarta.
- Rodebaugh, T. L., Woods, C. M., & Heimberg, R. G. (2007). The Reverse of Social Anxiety Is Not Always the Opposite: The Reverse-Scored Items of the Social Interaction Anxiety Scale Do Not Belong. *Behavior Therapy*, 38(2), 192-206. doi: 10.1016/j.beth.2006.08.001

- Rodebaugh, T. L., Woods, C. M., Heimberg, R. G., Liebowitz, M. R., & Schneier, F. R. (2006). The factor structure and screening utility of the social interaction anxiety scale. *Psychological Assessment, 18*(2), 231-237.
- Schmitt, N., & Stults, D. M. (1985). Factors Defined by Negatively Keyed Items - the Result of Careless Respondents. *Applied Psychological Measurement, 9*(4), 367-373.
- Schriesheim, C. A., & Eisenbach, R. J. (1995). An Exploratory and Confirmatory Factor-Analytic Investigation of Item Wording Effects on the Obtained Factor Structures of Survey Questionnaire Measures. *Journal of Management, 21*(6), 1177-1193. doi: 10.1177/014920639502100609
- Schriesheim, C. A., Eisenbach, R. J., & Hill, K. D. (1991). The Effect of Negation and Polar Opposite Item Reversals on Questionnaire Reliability and Validity - an Experimental Investigation. *Educational and Psychological Measurement, 51*(1), 67-78.
- Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Spector, P. E., Van Katwyk, P. T., Brannick, M. T., & Chen, P. Y. (1997). When Two Factors Don't Reflect Two Constructs: How Item Characteristics Can Produce Artifactual Factors. *Journal of Management, 23*(5), 659-677. doi: 10.1177/014920639702300503
- Timothy A, B. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire: Multiple factors or method effects? *Behaviour Research and Therapy, 41*(12), 1411-1426. doi: 10.1016/s0005-7967(03)00059-7
- van de Looij-Jansen, P. M., Goedhart, A. W., de Wilde, E. J., & Treffers, P. D. A. (2011). Confirmatory factor analysis and factorial invariance analysis of the adolescent self-report Strengths and Difficulties Questionnaire: How important are method effects and minor factors? *British Journal of Clinical Psychology, 50*(2), 127-144. doi: 10.1348/014466510x498174
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.
- Wei, M., Alvarez, A. N., Ku, T.-Y., Russell, D. W., & Bonett, D. G. (2010). Development and validation of a Coping with Discrimination Scale: Factor structure, reliability, and validity. *Journal of Counseling Psychology, 57*(3), 328-344. doi: 10.1037/a0019969
- Widhiarso, W. (2011). Contoh konstruk psikologi : bipolar dan orthogonal. Retrieved 15 Januari 2012, 2012, from <http://widhiarso.staff.ugm.ac.id/wp/?p=1343>

- Widhiarso, W., & Kožený, J. (2013). Applying confirmatory factor analysis to evaluate the fit of measurement models in education and psychological assessment *Journal of Educational Sciences and Psychology*, 3(2), 78-88.
- Yamamura, S., Takehira, R., Kawada, K., Katayama, S., Nishizawa, K., Hirano, M., & Momose, Y. (2005). Structural equation modeling of qualification of pharmacists to improve subjected quality of life in cancer patients. *Journal of Pharmacy & Pharmaceutical Sciences*, 8(3).