

PERBANDINGAN SOSIAL DESIRABILITAS BENTUK SOAL SKALA LIKERT, SEMANTIK DIFERENSIAL, DAN *FORCED-CHOICE* DALAM PENGUKURAN *PROSOCIAL TENDENCIES*

Farah Fauzia

farahvictory@gmail.com

Fakultas Psikologi UIN Jakarta

Abstrak:

Self report Prosocial Tendencies Measure (PTM) dibuat kedalam tiga format yang berbeda (Likert scale, Semantic Differential, dan Forced-Choice). Ketiga format PTM tersebut dianalisis untuk dilihat tingkat validitasnya dan kerentanannya terhadap social desirability response (SDR). Masing-masing format self-report terdiri dari 30 item pernyataan. Marlowe-Crowne Social Desirability Scale (MCSDS) digunakan untuk melihat kerentanan masing-masing format PTM terhadap SDR. Data yang digunakan dalam penelitian ini diperoleh dari hasil pengukuran terhadap 311 orang mahasiswa (219 perempuan dan 92 laki-laki). Hasil menunjukkan bahwa bentuk Semantic Differential memiliki tingkat validitas yang tinggi akan tetapi sangat rentan terhadap SDR. Bentuk Likert scale juga memiliki tingkat validitas yang tinggi setelah bentuk Semantic Differential akan tetapi masih rentan terhadap SDR. Sedangkan bentuk Forced-Choice sangat kurang rentan terhadap SDR. Analisis tambahan menunjukkan bahwa bentuk Likert scale –lah yang paling baik dalam mengukur PTM. Bentuk Semantic Differential dan Likert scale dilaporkan selain mengukur PTM juga mengandung bias social desirability. Meskipun bentuk Forced-Choice juga dilaporkan mengukur PTM akan tetapi mengandung bias social undesirability.

Kata kunci: Validitas konstruk, *social desirability*, *prosocial tendencies measures*, *Likert scale*, *semantic differential scale*, *forced-choice format*

Pendahuluan

Tes psikologi digunakan sebagai sarana untuk memperoleh data penelitian. Terdapat dua jenis tes psikologi berdasarkan tipe perilaku yang diukur, yaitu tes kemampuan dan tes kepribadian. Tes kemampuan

merupakan tes untuk mengukur kemampuan seseorang mengenai tingkat prestasi, bakat, minat dan intelegensinya. Sedangkan tes kepribadian merupakan tes untuk mengukur ciri atau karakteristik khusus yang ada dalam diri seseorang mengenai sikap, *trait*, dan tempramen

(Anastasi & Urbina, 1997). Tes kepribadian juga terbagi dalam kedua bentuk, yaitu tipe tes berstruktur dan tipe tes tak berstruktur. *Self report inventory* merupakan contoh dari tipe tes tak berstruktur ini (Cohen & Swerdlik, 2005). Penggunaan tes kepribadian tipe ini banyak sekali ditemukan dalam bidang psikologi industri dan organisasi. Tes kepribadian disebut-sebut sebagai prediktor yang valid mengenai kinerja pekerjaan (Ellingson et al, 2001). Survei yang dilakukan terhadap perekrutan pada tahun 2003 menunjukkan bahwa 30% dari perusahaan Amerika menggunakan tes kepribadian guna untuk menyaring atau menyeleksi para pelamar pekerjaan (Rothstein & Goffin, 2006). Dalam penyeleksian pegawai, tes kepribadian tersebut dibagikan kepada calon pegawai untuk ditanggapi sesuai dengan pernyataan yang diberikan. Data penelitian yang benar-benar tepat dalam menjelaskan kondisi empirik tidak begitu saja didapatkan, akan tetapi harus melalui sejumlah tantangan. Hal yang menyangkut penggunaan tes kepribadian adalah bahwa para pelamar dapat “memalsukan” respon mereka sehingga hal ini secara serius dapat mengurangi skor kepribadian (Rothstein & Goffin, 2006). Salah satu tantangan yang harus dihadapi adalah masalah ketepatan subjek dalam memberikan informasi yang sesuai dengan kondisinya. Keadaan ini terjadi saat dimana subjek memberikan jawaban menipu untuk menyembunyikan

identitasnya dan demi keamanan privasinya. Penelitian yang dilakukan Joinson (1999) mengenai *social desirability*, anonimitas, dan kuesioner yang berdasarkan pada internet ditemukan bahwa responden memiliki kecemasan sosial yang rendah, *social desirability* yang rendah dan *self esteem* yang tinggi ketika mereka menggunakan anonimitas. Lebih jauh lagi, responden dilaporkan memiliki kecemasan sosial dan *social desirability* yang rendah ketika mereka menggunakan internet daripada menggunakan metode manual. *Social desirability response* adalah kecenderungan responden untuk memberikan jawaban agar terlihat baik (Paulhus, 1991).

Deteksi respon desirabilitas sosial merupakan isu penting dalam perekrutan pegawai, karena umumnya diasumsikan bahwa pelamar memiliki motivasi yang kuat untuk menampilkannya diri dalam keterangan yang paling positif. Dengan demikian, responden dapat membuat diri mereka terlihat lebih baik pada item-item tertentu (Barrick & Mount, 1996). Respon yang distorsif sulit untuk diminimalisir dalam pengambilan data jika peneliti tidak melakukan tindakan antisipasi dengan melakukan beberapa cara. Zickar dan Drasgow (1996) memberikan tiga metode untuk mendeteksi *faking* pada pengukuran kepribadian, diantaranya adalah (1) menulis item tes yang sulit bagi responden untuk memalsukan respon mereka, (2) menyertakan skala validitas yang dirancang untuk

mendeteksi respon yang aneh, dan (3) mendeteksi respon distorsi yang didasarkan pada respon latensi. Sedangkan Messick menyarankan untuk menggunakan metode statistik untuk mengontrol bias respon (Mills & Kroner, 2005).

Paulhus (1991) memaparkan beberapa metode yang dapat digunakan untuk mengontrol *social desirability response*. Pertama adalah *rational technique*, teknik ini dilakukan ketika hendak membuat instrumen *self-report*. Teknik ini menyarankan untuk menggunakan *self-report* yang sulit bagi responden untuk berbohong. Contohnya adalah menggunakan format item *forced-choice*. Kedua adalah *factor analytic technique*, membuang pengaruh desirabilitas sosial dari skor skala dengan menggunakan teknik korelasi partial atau regresi (Paulhus, 1981). Teknik ini menganjurkan untuk menghilangkan item yang tidak valid pada saat analisis faktor. Ketiga adalah *covariate technique*, teknik ini menyarankan untuk menggunakan skala *social desirability* yang kemudian dikorelasikan dengan konstruk yang akan diukur. Yang keempat adalah *demand reduction*, merupakan teknik untuk mengurangi tekanan situasional. Salah satu caranya adalah menggunakan anonimitas pada responden.

Penelitian kali ini membahas mengenai peluang munculnya respon distorsif dengan membandingkan penyusunan bentuk item pernyataan. Penelitian kali ini sangat berguna

sekali dalam upaya meminimalisir tipuan respon. Karena salah satu hal yang mendukung subjek kesulitan untuk memberikan respon yang menipu adalah format dari pernyataan item itu sendiri. Format item yang tepat, diharapkan akan dapat mengurangi munculnya respon yang distorsif. Untuk mendukung penelitian ini, peneliti memerlukan *self-report* yang dapat dijadikan objek analisis. *Self report* yang dimaksud adalah *Prosocial Tendencies Measures* (PTM), yang mengukur perilaku *prosocial* yang spesifik, yaitu kecenderungan perilaku untuk menguntungkan orang lain (Carlo & Randall, 2002).

Metode

Jumlah sampel dalam penelitian ini sebanyak 311 orang, terdiri dari 219 wanita dan 92 laki-laki. Pengambilan sampel ini dilakukan secara *non probability sampling*. Adapun cara yang ditempuh adalah memilih responden sesuai karakteristik sampel yang dibutuhkan dan mudah untuk ditemukan serta menyertakan kesediaannya untuk dilibatkan dalam penelitian ini (*accidental sampling*). Karakteristik dari sampel yang dibutuhkan adalah seseorang pada masa perkembangan remaja akhir atau dewasa awal (18-21 tahun), dalam hal ini adalah mahasiswa.

Prosocial Tendencies Measures (PTM) terdiri dari 23 item pernyataan yang mengukur perilaku *Prosocial* yang spesifik, yaitu kecenderungan

perilaku untuk menguntungkan orang lain (Carlo & Randall, 2002). Terdapat enam tipe perilaku *prosocial* dalam PTM, yaitu: *Altruism*, *Compliant*, *Emotional*, *Public*, *Anonymous*, dan *Dires*. Setelah diterjemahkan kedalam bahasa Indonesia, *self report* PTM ini dibuat kedalam tiga bentuk format skala yang berbeda, yaitu: *Likert scale*, *semantic differential* dan *forced-choice*. Sebanyak 8 item ditambahkan dan 1 item tidak digunakan. Penambahan item ini agar masing-masing dimensi PTM memiliki 5 item.

Marlowe-Crowne Social Desirability Scale (MCSDS) (Crowne & Marlowe, 1960) terdiri dari 33 item yang terbagi dalam dua perilaku; (a) *desirable but uncommon behavior* dan (b) *undesirable but common behavior*. Adapun jumlah item yang digunakan dalam penelitian ini adalah 32 item, karena satu item tidak digunakan karena tidak cocok jika digunakan di Indonesia. Pada *self-report* ini responden diminta untuk merespon “benar” atau “salah” pada setiap item pernyataan yang disajikan.

Setelah data dikumpulkan melalui 1 set *self-report* (terdiri dari 3 format PTM dan MCSDS), maka data tersebut selanjutnya akan dianalisis. Dalam analisis penelitian ini terdapat 3 tahapan. Pertama, analisis faktor untuk mendapatkan validitas masing – masing bentuk soal dari *Prosocial Tendencies Measures* (PTM) dengan CFA *First Order* (tingkat item) maupun CFA *Second Order* (tingkat dimensi). Selain itu juga analisis

faktor untuk mendapatkan validitas dari *Marlowe-Crowne Social Desirability Scale* (MCSDS). Kedua, mengkorelasikan hasil analisis CFA *Second Order* (tingkat dimensi) masing-masing bentuk PTM dengan MCSDS. Data yang digunakan dalam korelasi ini adalah data berupa *true score* (faktor skor) masing-masing dimensi pada berbagai bentuk PTM dan *true score* skala MCSDS. Ketiga, analisis CFA *Second Order* masing-masing bentuk soal (*Likert scale*, *Semantic Differential* dan *Forced-Choice*) yang sama-sama mengukur *Prosocial Tendencies*, agar dapat dilihat skor PTM dari bentuk soal mana yang lebih tinggi *loading*-nya dalam mengukur PTM. Selain itu juga mengkorelasikan *true score* hasil dari masing–masing bentuk soal PTM guna mendapatkan validitas konvergen. Validitas konvergen ini berasumsi bahwa korelasi metode-metode yang berbeda tetapi mengukur konstruk yang sama seharusnya memiliki korelasi yang tinggi satu sama lain. Untuk melakukan ketiga analisis ini, maka penulis memutuskan untuk menggunakan pendekatan *Confirmatory Factor Analysis* (CFA) yang dikembangkan oleh Joreskog & Sorbom (2004).

Hasil

Setelah diperoleh model fit dimana nilai *chi-square* tidak signifikan ($p > 0.05$), maka dapat dilakukan analisis selanjutnya untuk mengetahui item-item mana yang valid dalam mengukur *Prosocial*

Tendencies. Pengujian model *Pro-social Tendencies* pertama adalah pada bentuk Likert. Pengujian model *Prosocial Tendencies* bentuk Likert pada dimensi *Altruism* dilaporkan fit ($chi-square = .89$, $df=4$, $p-value = .930$) dan terdapat 3 item yang valid. Model fit *Prosocial Tendencies* bentuk Likert pada dimensi *Compliant* ($chi-square = 5.54$, $df=3$, $p-value = .136$) terdapat 3 item yang valid. Model fit *Prosocial Tendencies* bentuk Likert pada dimensi *Emotional* ($chi-square = 3.36$, $df=4$, $p-value = .498$) menghasilkan 5 item yang valid. Model fit *Prosocial Tendencies* bentuk Likert pada

dimensi *Public* ($chi-square = 7.60$, $df=4$, $p-value = .107$) menghasilkan 5 item yang valid. Model fit *Prosocial Tendencies* bentuk Likert pada dimensi *Anonymous* ($chi-square = 6.58$, $df=3$, $p-value = .086$) menghasilkan 5 item yang valid. Model fit *Prosocial Tendencies* bentuk Likert pada dimensi *Dire* ($chi-square = 5.68$, $df = 3$, $p-value = .128$) menghasilkan 3 item yang valid. Tabel 3.1 menjelaskan muatan faktor, *standard error*, *t-value* dan *R-square* masing-masing item yang mengukur 6 dimensi PTM.

Tabel 3.1
Informasi mengenai item-item yang valid pada bentuk Likert

Dimensi	Item	Factor Loading	Std. Error	t-value	R-square
<i>Altruism</i>	4	0.55	0.07	7.49	0.30
	15	0.47	0.07	6.77	0.22
	19	0.84	0.09	9.35	0.70
<i>Compliant</i>	7	1.00 ^a	0.26	-	1.00
	24	0.24	0.05	3.54	0.06
	30	0.22	0.06	4.06	0.05
<i>Emotional</i>	2	0.41	0.06	6.30	0.17
	11	0.37	0.07	5.59	0.13
	16	0.67	0.07	9.98	0.45
	20	0.75	0.07	10.87	0.57
	26	0.34	0.07	5.19	0.12
<i>Public</i>	1	0.63	0.07	8.58	0.39
	3	0.67	0.07	8.92	0.44
	5	0.53	0.07	7.67	0.28
	12	0.29	0.07	3.96	0.08
	25	0.14	0.07	2.06	0.02
<i>Anonymous</i>	8	0.70	0.06	11.40	0.50
	10	0.42	0.07	5.96	0.18
	14	0.49	0.06	7.78	0.24
	18	0.59	0.06	9.78	0.35
	21	0.69	0.06	11.43	0.47
<i>Dire</i>	9	0.19	0.08	2.29	0.03
	13	0.47	0.08	5.65	0.22
	27	0.63	0.10	6.54	0.40

^afactor loading item ditetapkan bernilai 1

Setelah pengujian model *first order*, dengan menggunakan item-item yang valid maka dilakukan pengujian model fit *second order* untuk melihat muatan faktor *loading* masing-masing dimensi (*Gamma*). Setelah model dikatakan fit (*chi-square*= 187.19, *df*= 161, *p-value*=

.077) maka dapat diketahui mengenai informasi *factor loading (gamma)*, *standard error*, *t-value* dan *R-square* masing-masing dimensi dalam hal mengukur konstruk *Prosocial Tendencies*. Untuk lebih jelasnya, dapat dilihat pada tabel 3.2.

Tabel 3.2.

Koefisien muatan faktor setiap dimensi PTM dari CFA Second order untuk bentuk Likert

Dimensi	<i>Gamma</i>	<i>Std. Error</i>	<i>t-value</i>	<i>R-square</i>
<i>Altruism</i>	0.22	0.05	4.12	0.10
<i>Compliant</i>	0.67	0.05	12.57	0.45
<i>Emotional</i>	0.58	0.05	11.10	0.74
<i>Public</i>	-0.08 ^b	0.04	-2.28	0.03
<i>Anonymous</i>	0.47	0.06	8.42	0.80
<i>Dire</i>	0.75	0.05	14.46	1.07 ^c

^bmuatan faktor negatif

^c*Heywood case*

Berdasarkan hasil analisis CFA *second order* PTM bentuk Likert diperoleh informasi mengenai muatan faktor (*Gamma*) masing-masing dimensi. Pada bentuk Likert, dimensi yang paling besar sumbangnya dalam hal mengukur *Prosocial Tendencies* adalah dimensi *Dire*. Pada dimensi *Public*, muatan faktor (*Gamma*) bernilai negatif. hal ini dikarenakan item-item pada dimensi ini memiliki nilai yang negatif jika dibandingkan dengan item yang lain. Dengan kata lain, semakin tinggi prososial seseorang maka semakin rendah skor yang diberikan pada dimensi *Public*.

Pengujian model *Prosocial Tendencies* selanjutnya adalah pada bentuk *semantic differential*. Pengujian model fit pertama dilakukan adalah pengujian model fit CFA *first order* setelah itu baru CFA *second order*. Model fit *Prosocial Tendencies* bentuk *semantic differential* pada dimensi *Altruism* (*chi-square*=5.73, *df*=4, *p-value*=.22) menghasilkan 5 item yang valid. Model fit *Prosocial Tendencies* bentuk *semantic differential* pada dimensi *Compliant* (*chi-square*=3.45, *df*=2, *p-value*=.177) menghasilkan 4 item yang valid. Model fit *Prosocial Tendencies* bentuk *semantic differential* pada

dimensi *Emotional* ($chi-square=0.39$, $df=3$, $p-value=.943$) menghasilkan 5 item yang valid. Model fit *Prosocial Tendencies* bentuk *semantic differential* pada dimensi *Public* ($chi-square=5.69$, $df=4$, $p-value=.223$) menghasilkan 5 item yang valid. Model fit *Prosocial Tendencies* bentuk *semantic differential* pada dimensi *Anonymous* ($chi-square= 5.79$, $df=3$, $p-value=.122$) menghasilkan 5 item yang valid. Model fit *Prosocial Tendencies* bentuk *semantic differential* pada dimensi *Dire* ($chi-square=5.69$, $df=3$, $p-value=.128$) menghasilkan 5 item yang valid. Berbeda dari hasil CFA PTM bentuk Likert, bentuk *semantic differential* menghasilkan jumlah item valid yang lebih banyak. Tabel 3.3 menjelaskan muatan faktor, *standard error*, *t-value*

dan *R-square* masing-masing item yang mengukur dimensi PTM. Item nomer 25 tidak disertakan dalam analisis *second order* selanjutnya, meskipun dilaporkan valid, karena item ini memiliki korelasi kesalahan pengukuran dengan lebih dari tiga item lainnya. Berdasarkan uji model fit *first order* CFA PTM bentuk *semantic differential* menghasilkan 28 item yang valid. Item-item yang valid tersebut diujikan lagi dengan model *second order* untuk mendapatkan *factor loading* (*Gamma*) masing-masing dimensi dari PTM. Tabel 3.4 menjelaskan mengenai muatan faktor (*Gamma*), *standard error*, *t-value*, dan *R-square* masing-masing dimensi yang mewakili PTM.

Tabel 3.3
Informasi mengenai item-item yang valid pada bentuk *semantic differential*

Dimensi	Item	Factor Loading	Std. Error	t-value	R-square
<i>Altruism</i>	5	0.52	0.06	8.80	0.27
	9	0.61	0.06	10.18	0.38
	21	0.90	0.07	13.53	0.82
	22	0.74	0.07	10.40	0.55
	23	0.30	0.06	5.14	0.09
<i>Compliant</i>	8	0.60	0.05	11.15	0.35
	9	0.89	0.05	18.29	0.80
	10	0.94	0.05	19.72	0.89
	25	0.14	0.06*	2.36	0.02
<i>Emotional</i>	2	0.42	0.06	6.53	0.17
	15	0.73	0.06	12.63	0.53
	20	0.68	0.06	11.97	0.57
	29	0.77	0.06	13.12	0.47
	30	0.76	0.06	12.91	0.59

	1	0.74	0.05	13.46	0.55
	3	0.79	0.05	14.64	0.63
<i>Public</i>	4	0.62	0.06	10.80	0.38
	16	0.49	0.06	8.14	0.24
	24	0.62	0.06	10.88	0.39
	11	0.67	0.08	8.83	0.45
	12	0.66	0.08	8.65	0.43
<i>Anonymous</i>	14	0.71	0.08	9.28	0.50
	17	0.39	0.07	5.79	0.15
	18	0.46	0.07	6.98	0.21
	6	0.50	0.06	8.43	0.29
	7	0.56	0.06	9.62	0.31
<i>e</i>	13	0.30	0.06	15.08	0.09
	27	0.68	0.06	11.72	0.47
	28	0.92	0.06	15.51	0.85

*item yang memiliki korelasi kesalahan 3 atau lebih dengan item lain

Hasil *CFA Second order* pada bentuk *semantic differential* menunjukkan hasil yang sama dengan bentuk Likert. Pada analisis CFA ini ditemukan adanya dimensi PTM yang memiliki koefisien muatan faktor (*Gamma*) yang negatif. Dimensi tersebut adalah dimensi *Public*. Item-item pada dimensi ini memiliki nilai yang bersifat negatif jika disebanding-

kan dengan item-item pada dimensi lain. Dengan demikian memberikan penjelasan bahwa semakin tinggi tendensi prososial seseorang maka semakin rendah skor pada dimensi *Public*. Dimensi PTM yang paling besar sumbangannya dalam mengukur *Prosocial Tendencies* adalah dimensi *Dire*.

Tabel 3.4

Koefisien muatan faktor setiap dimensi PTM dari CFA Second order untuk bentuk Semantik

Dimensi	<i>Gamma</i>	<i>Std. Error</i>	<i>t-value</i>	<i>R-square</i>
<i>Altruism</i>	0.46	0.05	9.58	0.72
<i>Compliant</i>	0.65	0.05	13.19	0.57
<i>Emotional</i>	0.70	0.05	14.36	0.90
<i>Public</i>	-0.25 ^b	0.04	-6.32	0.37
<i>Anonymous</i>	0.58	0.05	11.05	0.58
<i>Dire</i>	0.68	0.05	13.57	1.17 ^c

^bmuatan faktor negatif

^c*Heywood case*

Pengujian model *Prosocial Tendencies* selanjutnya adalah pada bentuk *forced-choice*. Pada bentuk ini hanya dilakukan pengujian model fit CFA *first order* akan tetapi pengujiaannya setara dengan CFA *Second order*, karena yang dianalisis adalah langsung skor level dimensi (bukan level item). Pada bentuk ini subjek dipaksa untuk memilih pernyataan yang sesuai diantara sepasang pernyataan yang disajikan. Dengan demikian, skor yang digunakan adalah frekuensi dari

pilihan responden terhadap pernyataan pada keenam dimensi. Hanya terdapat 13 item yang dapat dianalisis pada bentuk ini, karena 17 item lainnya bersifat negatif dan ketika dianalisis bersamaan tidak dapat diperoleh model yang bersifat unidimensional dalam mengukur PT. Tiga belas item yang dianalisis menghasilkan model fit (*chi-square*=9.24, *df*=5, *p-value*=.099). Untuk lebih jelasnya dapat dilihat pada tabel 3.5

Tabel 3.5a.

Hasil analisis CFA first order bentuk forced-choice

Dimensi	<i>Gamma</i>	<i>Std. Error</i>	<i>t-value</i>	<i>r-square</i>
<i>Altruism</i>	0.41	0.06	7.04	0.16
<i>Compliant</i>	0.34	0.05	7.01	0.10
<i>Emotional</i>	0.19	0.04	4.80	0.03
<i>Public</i>	0.36	0.06	5.76	0.12
<i>Anonymous</i>	-1.68 ^b	0.14	-12.18	2.58
<i>Dire</i>	0.15	0.04*	4.14	0.02

^a*factor loading* item ditetapkan bernilai 1

^bmuatan faktor negatif

*memiliki korelasi kesalahan pengukuran dengan banyak dimensi lain

Hasil analisis CFA *first order* bentuk *forced-choice* menunjukkan bahwa dimensi *Anonymous* pada bentuk ini memiliki koefisien muatan faktor yang negatif. Hal ini dikarenakan jumlah pernyataan positif yang mengukur dimensi *Anonymous* lebih banyak jika dibandingkan pada dimensi lainnya. Dengan demikian,

jika semua dimensi bermuatan positif maka dimensi *Anonymous* –lah yang bermuatan negatif. Sebaliknya, ketika dimensi *Anonymous* ditetapkan bernilai 1, maka semua dimensi bermuatan negatif hal ini dapat dilihat pada tabel 3.5b.. Pada bentuk *forced-choice* dimensi yang paling mempengaruhi dalam hal mengukur *Prosocial Ten-*

dencies adalah *Anonymous*. Dimensi *Dire* memiliki korelasi kesalahan pengukuran dengan dimensi *Emotional*, *Compliant*, dan *Altruism*. Hal ini mengindikasikan bahwa dimensi *Dire* selain mengukur *Prosocial Tendencies*, item-item dalam dimensi ini memiliki kemiripan dengan dimensi

Emotional, *Compliant*, dan *Altruism*. Sedangkan deimsni *Anony-mous* dan dimensi *Public* memiliki perbedaan, artinya item-item dalam kedua dimensi tersebut tidak memiliki kemiripan dengan item-item dalam dimensi lainnya.

Tabel 3.5b

Hasil analisis CFA first order bentuk forced-choice

Dimensi	Gamma	Std. Error	t-value	r-square
<i>Altruism</i>	-0.25 ^b	0.05	-4.81	0.16
<i>Compliant</i>	-0.20 ^b	0.04	-5.02	0.10
<i>Emotional</i>	-0.11 ^b	0.03	-3.78	0.03
<i>Public</i>	-0.21 ^b	0.05	-4.06	0.12
<i>Anonymous</i>	1.00 ^a	-	-	2.58
<i>Dire</i>	-0.09 ^b	0.03	-3.43	0.02

^afactor loading item ditetapkan bernilai 1

^bmuatan faktor negatif

*memiliki korelasi kesalahan pengukuran dengan banyak dimensi lain

Setelah pengujian model fit pada masing-masing bentuk PTM. Pengujian model fit juga dilakukan pada *Marlowe-Crowne Social Desirability Scale* (MCSDS). Pengujian ini diharapkan menghasilkan item-item yang valid dalam hal mengukur *social dsirability*. Item-item yang valid tersebut yang kemudian digunakan untuk mendapatkan *true score* dari variabel *social desirability* dan dikorelasikan dengan *true score* dimensi-dimensi PTM pada masing-masing bentuk soal. CFA untuk MCSDS menghasilkan model fit

dengan *chi-square*=472.86, *df*= 430, *p-value*=.07, dan selanjutnya item-item MCSDS yang valid dapat dilihat di tabel 3.6. Berdasarkan hasil analisis CFA MCSDS, ditemukan hanya terdapat 19 item yang valid mengukur *social desirability*. Sembilan belas item ini digunakan untuk mendapatkan *true score* MCSDS.

Tabel 3.6

Hasil analisis CFA first order MCSDS

Item	<i>Factor Loading</i>	<i>Std. Error</i>	<i>t-value</i>	<i>r-square</i>
1	0.39	0.09	4.20	0.07
2	0.36	0.09	4.03	0.06
4	0.30	0.09	3.37	0.05
7	0.25	0.09	2.77	0.03
9	0.50	0.09	5.74	0.13
12	0.54	0.09*	6.24	0.15
13	0.45	0.09	5.16	0.10
14	0.25	0.09*	2.72	0.03
15	0.29	0.09	3.31	0.04
16	0.24	0.09	2.65	0.03
17	0.60	0.09	7.01	0.18
18	0.20	0.09	2.15	0.02
19	0.40	0.09	4.54	0.08
20	0.23	0.09	2.60	0.03
21	0.47	0.09	5.35	0.11
24	0.29	0.09	3.21	0.04
25	0.86	0.08*	10.11	0.37
26	0.66	0.08	7.75	0.22
28	0.38	0.09	4.19	0.07
29	0.47	0.09	5.35	0.11
30	0.81	0.08	9.72	0.33
32	0.62	0.09	6.93	0.19

*item yang memiliki korelasi kesalahan 3 atau lebih dengan item lain

Korelasi PTM dengan *Social Deisrability*

Setelah melakukan pengujian model fit pada masing-masing bentuk PTM baik *first order* maupun *second order* dan pengujian model fit MCSDS, maka dapat dilakukan analisis selanjutnya. Analisis selanjutnya, yaitu mengkorelasikan masing-masing *true score* dimensi-dimensi PTM ketiga bentuk soal dengan *true score* MCSDS. Hal ini dilakukan

untuk memperoleh informasi mengenai mana dari ketiga bentuk PTM yang rentan terhadap *social desirability*.

Berdasarkan tabel 3.7 diperoleh informasi mengenai bentuk dari PTM mana yang rentan terhadap *social desirability*. Hasil menunjukkan bahwa pada bentuk soal *semantic differential*, semua dimensi dari PTM berkorelasi signifikan dengan MCSDS. Hal ini mengindikasikan

bahwa seseorang yang mendapatkan nilai tinggi pada dimensi-dimensi PTM bentuk *semantic differential* maka semakin tinggi skor *social desirability* orang tersebut. Pada bentuk *semantic differential*, dimensi *Public* ($R=-.177, p<0.01$) memiliki korelasi yang signifikan dan berlawanan arah (negatif). Hal ini berarti bahwa mereka yang menjawab tidak setuju pada pernyataan dimensi *Public* adalah mereka yang cenderung menjawab secara *social desirability*. Hal ini dikarenakan responden mengetahui bahwa item-item yang

terdapat dalam dimensi *Public* bersifat *socially undesirable*. Item-item pada dimensi *Public* mengukur sikap menolong seseorang yang terjadi dihadapan orang banyak. Jika mereka yang cenderung menolong seseorang dihadapan orang banyak, maka mereka cenderung menyetujui item-item tersebut dan respon mereka tidak mengandung *social desirability*. Berbeda halnya dengan mereka yang memiliki skor rendah (merespon tidak setuju), maka mereka yang menjawab secara *social desirability* yang tinggi.

Tabel 3.7.

Korelasi antara true score PTM dengan MCSDS pada masing-masing dimensi berdasarkan bentuk soal.

Bentuk	Dimensi	R	Sig
Likert Scale	<i>Altruism</i>	.198**	.000
	<i>Compliant</i>	.175**	.002
	<i>Emotional</i>	.086	.132
	<i>Public</i>	-.154**	.006
	<i>Anonymous</i>	.085	.136
	<i>Dire</i>	.057	.313
<i>Semantic Differential</i>	<i>Altruism</i>	.163**	.004
	<i>Compliant</i>	.253**	.000
	<i>Emotional</i>	.238**	.000
	<i>Public</i>	-.177**	.002
	<i>Anonymous</i>	.204**	.000
	<i>Dire</i>	.245**	.000
<i>Forced-Choice</i>	<i>Altruism</i>	.014	.775
	<i>Compliant</i>	.071	.193
	<i>Emotional</i>	.045	.450
	<i>Public</i>	.206**	.001
	<i>Anonymous</i>	-.115*	.044
	<i>Dire</i>	-.080	.187

** $p<0.01$

* $p<0.05$

Pada bentuk Likert, terdapat tiga dimensi yang memiliki korelasi yang signifikan dengan MCSDS (rentan terhadap *social desirability*) yaitu

dimensi *Altruism* ($R=.198, p<0.01$), *Compliant* ($R=.175, p<0.01$) dan *Public* ($R=-.154, p<0.01$). Sama halnya dengan bentuk *semantic*

differential, dimensi *Public* pada bentuk ini memiliki korelasi yang signifikan dan berlawanan arah (negatif). Hal ini mengindikasikan bahwa mereka yang menjawab tidak setuju pada item-item dimensi ini adalah mereka yang menjawab secara *social desirability*. Sedangkan mereka yang menjawab setuju pada item-item ini adalah mereka yang tidak menjawab secara *social desirable*. Hal ini dikarenakan responden mengetahui bahwa item-item pada dimensi *Public* mengandung *socially undesirable*.

Bentuk terakhir adalah *forced-choice*. Pada bentuk soal ini, hanya terdapat dua dimensi saja yang memiliki korelasi dengan MCSDS (rentan terhadap *social desirability*). Dimensi tersebut adalah dimensi *Public* ($R=.206, p<0.01$) dan *Anonymous* ($R=-.115, p<0.05$). Berbeda dengan kedua bentuk sebelumnya, dimensi *Public* pada bentuk ini memiliki korelasi signifikan dan positif. Hal ini terjadi karena dalam bentuk soal *forced-choice* responden dipaksa untuk memilih salah satu pernyataan diantara sepasang pernyataan yang berbeda dimensi. Dengan demikian, dimensi *Public* tidak memiliki korelasi yang negatif karena semakin sering seseorang memilih pernyataan pada dimensi *Public* maka semakin *social desirable*. Berbeda halnya dengan dimensi *Anonymous*, dimensi ini memiliki korelasi signifikan dan negatif. Hal ini berarti bahwa semakin sering pernyataan pada dimensi *Anonymous* dipilih maka semakin

tidak mengandung *social desirability*. Sedangkan semakin pernyataan pada dimensi *Anonymous* tidak dipilih, maka semakin mengandung *social desirability*. Hal ini dikarenakan pernyataan pada *Anonymous* dinilai sebagai *socially undesirable*. Responden dengan sadar mengetahui bahwa jika mereka tidak memilih pernyataan pada dimensi *Public* maka mereka dianggap merespon secara jujur. Sedangkan jika mereka memilih pernyataan pada dimensi *Anonymous* maka mereka dianggap merespon secara tidak jujur. Dengan demikian, bentuk *forced-choice* –lah yang dianggap mampu menekan respon *social desirability*. Adapun pernyataan-pernyataan pada dimensi *Public* dan *Anonymous* harus diperbaiki agar responden tidak mudah memberikan respon palsu mereka.

Analisis ketiga adalah analisis CFA *second order* berbagai bentuk dari PTM. Pada analisis berikut ini data yang digunakan adalah data berupa faktor skor dari masing-masing dimensi PTM dari ketiga bentuk yang berbeda. Analisis ini dimaksudkan untuk melihat apakah terdapat perbedaan *factor loading* masing-masing dimensi *Prosocial Tendencies Measures* masing-masing bentuk soal. Berdasarkan dari ketiga bentuk soal yang disajikan, manakah bentuk soal yang paling baik dalam hal mengukur *Prosocial Tendencies*. Setelah pengujian model fit pada ketiga bentuk soal PTM ($\chi^2=6219.14, df=1592, p\text{-value}=1.000$) selanjutnya dapat dilakuk-

analisis. Tabel 3.8 menjelaskan informasi mengenai *factor loading* masing-masing dimensi terhadap masing-masing bentuk PTM. Selain itu diperoleh juga informasi mengenai *factor loading* masing-masing bentuk PTM (*Gamma*) terhadap PTM itu sendiri.

Hasil analisis ini menunjukkan bahwa pada bentuk Likert dan *semantic differential* memang mengukur *Prosocial Tendencies* akan tetapi mengandung *social desirability*. Sedangkan bentuk *forced-choice* selain mengukur *Prosocial Tendencies*, bentuk ini juga mengandung *social*

undesirability sehingga arah *loading*-nya berlawanan. Hal ini mengindikasikan bahwa semakin tinggi nilai PTM yang diukur dengan bentuk Likert dan *semantic differential*, maka semakin rendah nilai PTM yang diukur dengan menggunakan *forced-choice*. Dengan demikian, bentuk *forced-choice* mengukur *Pro-Anti Social Tendencies*. Skor pada bentuk *forced-choice* didominasi oleh item-item yang bersifat *socially undesirable*. Sedangkan pada bentuk Likert dan *semantic differential*, didominasi oleh item-item yang bersifat *socially desirable*.

Tabel 3.8
Second order CFA masing-masing bentuk soal

Bentuk	Dimensi	Factor Loading	Std. Error	T-value	Sig
Likert Scale	<i>Altruism</i>	0.251	0.118	2.123	0.034
	<i>Compliant</i>	0.563	0.062	9.096	0.000
	<i>Emotional</i>	0.585	0.076	7.655	0.000
	<i>Public</i>	-0.287	0.120	-2.389	0.017
	<i>Anonymous</i>	0.708	0.064	11.015	0.000
	<i>Dire</i>	0.386	0.082	4.735	0.000
	<i>Gamma</i>	-0.999	0.098	-10.172	0.000
Semantic Differential	<i>Altruism</i>	0.499	0.057	8.773	0.000
	<i>Compliant</i>	0.675	0.042	15.985	0.000
	<i>Emotional</i>	0.811	0.025	32.500	0.000
	<i>Public</i>	-0.413	0.063	-6.569	0.000
	<i>Anonymous</i>	0.500	0.056	8.956	0.000
	<i>Dire</i>	0.889	0.023	38.501	0.000
<i>Gamma</i>	-0.545	0.084	-6.485	0.000	
Forced-choice	<i>Altruism</i>	0.799	0.012	66.128	0.000
	<i>Compliant</i>	0.567	0.041	13.722	0.000
	<i>Emotional</i>	0.287	0.052	5.477	0.000
	<i>Public</i>	0.582	0.052	11.295	0.000
	<i>Anonymous</i>	-0.997	0.001	-859.226	0.000
	<i>Dire</i>	0.266	0.059	4.510	0.000
	<i>Gamma</i>	0.415	0.017	23.901	0.000

Analisis selanjutnya adalah mengkorelasikan bentuk-bentuk PTM (Likert, *semantic differential* dan *forced-choice*). Analisis ini dilakukan untuk mengetahui apakah hasil PTM menurut Likert *scale* (LK) berkorelasi tinggi dengan hasil PTM bentuk

semantic differential (SD) dan *forced-choice* (FC). Jika ketiga bentuk skala tersebut memang mengukur satu konstruk yaitu *Prosocial Tendencies* maka seharusnya ketiga bentuk tersebut saling berkorelasi tinggi (Validitas Konvergen).

Tabel 3.9
Matriks Korelasi berbagai bentuk PTM

Variabel	LK	SD	FC
LK	1		
SD	0.84*	1	
FC	-0.02	0.01	1

* $p < 0.05$

Berdasarkan tabel 3.9 diperoleh informasi mengenai korelasi antara ketiga bentuk soal (Likert, *semantic differential* dan *forced-choice*). Faktor skor pada *forced-choice* tidak berkorelasi signifikan dengan faktor skor bentuk *semantic differential* ($R = .01$, *ns*) dan bentuk Likert ($R = .02$, *ns*). Sedangkan antara bentuk Likert dan *semantic differential* memiliki korelasi yang signifikan ($R = .84$, $p < 0.05$). Hal ini menunjukkan bahwa bentuk Likert dan *semantic differential* mengukur hal yang sama, akan tetapi memiliki perbedaan dengan bentuk *forced-choice*. Berdasarkan temuan sebelumnya, bentuk *semantic differential* dan Likert sangat rentan terhadap *social desirability*. Sedangkan bentuk *forced-choice* adalah bentuk soal yang dimensinya paling sedikit berkorelasi dengan *social desirability*. Adanya korelasi yang

signifikan antara bentuk soal *semantic differential* dan Likert *scale* disebabkan karena kedua bentuk soal tersebut memiliki korelasi yang signifikan dengan variabel *social desirability*. Dengan kata lain, bentuk soal *semantic differential* dan Likert *scale* selain mengukur PTM, keduanya juga mengukur *social desirability*.

Diskusi

Berikut ini akan dipaparkan mengenai beberapa temuan yang didapat setelah melakukan analisis terhadap data yang dikumpulkan. Berdasarkan hasil analisis validitas yang telah dilakukan, maka dapat ditarik kesimpulan bahwa bentuk *semantic differential* –lah yang dilaporkan memiliki validitas tertinggi (dilihat dari besaran *gamma*). Hasil analisis korelasi antara ketiga bentuk soal dengan *social desirability*

menunjukkan bahwa bentuk *forced-choice* lah yang kurang rentan terhadap *social desirability*. Hasil analisis CFA *second order* dari ketiga bentuk PTM mengindikasikan bahwa bentuk Likert lah yang dilaporkan paling baik dalam hal mengukur *Prosocial Tendencies*. Selain itu juga terdapat koefisien *gamma* masing-masing bentuk PTM yang bersifat negatif. Hal ini menunjukkan bahwa kedua bentuk PTM (Likert dan *semantic differential*) selain valid mengukur *Prosocial Tendencies* tetapi juga mengandung *social desirability*. Sedangkan pada bentuk *forced-choice* mengukur hal yang sebaliknya yaitu *Pro Anti Social Tendencies*, dimana item-itemnya mengandung bias *social undesirability*. Analisis korelasi tambahan memberikan informasi tambahan mengapa kedua bentuk PTM (Likert dan *semantic differential*) memiliki muatan yang negatif pada analisis CFA *second order* yang telah dilakukan. Perbedaan arah ini mengindikasikan bias yang terjadi pada ketiga bentuk soal. Pada item-item Likert dan *semantic differential*, bias yang terjadi adalah *social desirability*. Sedangkan pada item-item bentuk *forced-choice* bias yang terjadi adalah *social undesirability*.

Berdasarkan hasil analisis yang telah dilakukan, maka sesuai hipotesis yang menyatakan bahwa bentuk *forced-choice* lah yang kurang rentan terhadap *social desirability* jika dibandingkan dengan bentuk Likert dan *semantic differential*. Hasil penelitian ini sejalan dengan peneli-

tian yang telah dilakukan oleh Jackson, Wroblewski dan Ashton (2000) yang menemukan bahwa respon *faking* lebih rendah terdapat pada format pernyataan *forced-choice* dari pada *single-stimulus*. Penelitian lain yang dilakukan Christiansen et al (Jackson et al, 2000) dengan menggunakan skala *Personality Adjective* yang dirubah kedalam bentuk *single-stimulus* (skala Likert) dan format *forced-choice* pada 350 sampel. Setengah dari sampel merespon dibawah instruksi normal dan setengahnya lagi diinstruksikan untuk merespon seakan-akan melamar pekerjaan menjadi seorang *sales*. Mereka menemukan bahwa perbedaan *mean* antara kelompok instruksi normal dengan kelompok pelamar adalah rendah pada format *forced-choice* daripada *single-stimulus* (skala Likert). Mereka juga menguji korelasi skala dengan menggunakan kuesioner kepribadian dibawah instruksi normal. Pada kelompok pelamar korelasi dari *single-stimulus* dengan skala kepribadian jauh lebih rendah dari pada korelasi dengan menggunakan sampel kelompok normal. Akan tetapi, korelasi untuk skala *forced-choice* hampir sama tinggi untuk kelompok pelamar daripada kelompok normal, meskipun terdapat perbedaan format item antara *forced-choice* dan perbandingan skala kepribadian.

Dalam penelitian kali ini, penulis menggunakan 3 strategi untuk menekan atau mengurangi dampak dari *social desirability* (Paulhus,

1991). Pertama, membuat *self-report* yang efektif dalam menghindari respon palsu (*faking response*), yang dalam hal ini menggunakan *self-report* bentuk *forced-choice* (*Rational Technique*). Kedua, menggunakan analisis faktor yang dapat berguna untuk membuang item-item yang tidak valid, hal ini dikarenakan item mengukur hal lain (*Factor Analytic Technique*). Terakhir, menyertakan skala *social desirability* yang kemudian dikorelasikan dengan *self-report* yang digunakan (*Covariate Technique*).

Meskipun *self-report* bentuk *forced-choice* dapat digunakan untuk mengurangi atau mencegah responden berbohong, akan tetapi dalam pelaksanaannya sangat sulit untuk membuat *self-report* bentuk *forced-choice*. Pertama, memasang dua pernyataan yang memiliki nilai *social desirability* yang sama merupakan hal yang sulit dan hampir tidak mungkin dilakukan. Selain itu, ketika kedua item tersebut dipasangkan tidak menutup kemungkinan bahwa pasangan pernyataan ini tidak memiliki *social desirability* yang seimbang (Messick, 1960; Edward, 1970). *Missing* -nya data juga umum terjadi dalam format *forced-choice* ini, karena responden bisa sangat frustrasi ketika mereka harus memilih satu diantara pernyataan yang sangat sesuai dan sangat tidak sesuai dengan dirinya. Yang terakhir adalah sangat sulit menggunakan format *forced-choice* ini dalam prosedur analisis faktor, karena sulit untuk mengetahui

mana dari dua pernyataan yang mungkin di pilih oleh responden (Paulhus, 1981).

Beberapa masalah lain juga diungkapkan dalam hal penggunaan format *forced-choice*. Pertama, penelitian terdahulu menunjukkan bahwa format *forced-choice* masih rentan terhadap respon *social desirability*, meskipun dalam penggunaannya menyamakan pilihan pernyataan yang memiliki tingkat *social desirability* yang sama. Kedua, masalah yang terkait dengan psikometrik yaitu adalah sistem skoring yang dilakukan dalam format *forced-choice*. Kritik lain juga melibatkan masalah mengenai relativitas dalam hal pilihan mana yang relatif paling baik dari yang lain. Akibatnya nilai yang dihasilkan tidak mengungkapkan dengan benar mengenai *trait*. Dan demikian tidak pantas jika dibandingkan dengan keseluruhan individu. Kerena pilihan pernyataan terdiri dari *trait* yang berbeda, dengan memilih satu pernyataan berarti salah satu pernyataan dari *trait* lain tidak dipilih (Christiansen et al., 2005).

Untuk penelitian selanjutnya, penulis menyarankan untuk menggunakan metode MTMM (*multitrait-multimethod*). Diantaranya adalah validitas diskriminan, metode yang dilakukan adalah mengkorelasikan metode-metode yang sama pada konstruk yang berbeda. Selain itu juga, dapat dikorelasikan dengan konstruk lain yang telah diketahui berkorelasi positif atau negatif dengan konstruk yang diukur. Selanjutnya

adalah validitas konvergen, metode yang dilakukan pada validitas ini adalah mengkorelasikan metode-metode yang berbeda tetapi mengukur hal yang sama (konstruk yang sama). Metode ini berasumsi bahwa interkorelasi hasil pengukuran dari *trait* yang sama dengan metode yang berbeda harus tinggi daripada korelasi anantara *trait* yang berbeda.

Saran lain ketika hendak ingin menggunakan format *forced-choice* perlu memperhatikan pernyataan yang akan digunakan. Pernyataan tersebut harus bersifat positif dan memiliki tingkat *social desirability* yang seimbang. Lebih dianjurkan untuk memasang pernyataan yang memiliki tingkat *social desirability* yang sama dibandingkan memasang pernyataan dengan tingkat *social undesirability* yang sama. Hal ini dilakukan untuk menghindari bias *social undesirability*. Selanjutnya, penggunaan bentuk *semantic differential* dalam *self report* sebaiknya dihindari, karena meskipun melaporkan validitas yang tinggi akan tetapi mengandung *social desirability* yang paling tinggi. Dan jika ingin menggunakan bentuk *semantic differential* maka perlu dihilangkan respon tengah. Ketika peneliti lain merasa kesulitan dalam hal menyusun *self report* dalam bentuk *forced-choice*, maka dianjurkan untuk menggunakan skala Likert. Namun, sebaiknya dengan menghilangkan pilihan respon tengah dan menyertakan skala *social desirability* yang kemudian dikore-

lasikan dengan konstruk yang ingin diukur.

Daftar Pustaka

- Anastasi, Anne., & Urbina, Susana. (1997). *Psychological testing*. 7th edition. Pearson Prentice Hall.
- Barrick, R. Murray., & Mount, K. Michael. (1996). Effects of impression management and a self-deception on the predictive validity of personality construct. *Journal of Applied Psychology*, 81, 261,272.
- Carlo, Gustavo., & Randall, A. Brandy. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of Youth and Adolescence*, 33, 31-34.
- Christiansen, Niel. D., Burns, Garry. N., & Montgomery, George. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance*, 18 (3), 267-307.
- Edwards, AL., Walsh, JA., & Diers, CJ. (1970). The relationship between social desirability and internal consistency of personality scales. *Journal of Applied*, 47 (4), 255-259.
- Ellingson, E. Jill., Smith, D. Brent., & Sackett, R. Paul. (2001) Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology*, 86, 122-133.

- Jackson, Douglas., Wroblewski, R. Victor., & Ashton, C. Michael. (2000). The impact of faking on employment tests: does forced-choice offer a solution?. *Human Performance*, 13, 371-388.
- Joinson, Adam. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instrument, & Computers*, 31, 433-438.
- Joreskog, K. G., & Sorbom, D. (2004). *LISREL 8.70 for Windows (computer software)*. Lincoln Wood, IL: Scientific Software International, Inc.
- Messick, Samuel. (1960). Dimensions of social desirability. *Journal of Consulting Psychology*, 24 (4), 279-287.
- Mills, F. Jeremy & Kroner, G. Daryl. (2005). An investigation into the relationship between socially desirable responding and offender self-report. *Psychological Services*, 2, 70-80.
- Paulhus, DL. (1981). Control of social desirability in personality inventories: principal-factor deletion. *Journal of Research in Personality*, 15, 383-388.
- Paulhus, DL. Measurement and control of response bias. (1991). In J. P. Robinson, P.R. Shaver & L.S. Wrightsman (Eds). *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press, Inc.
- Rothstein, g. Mitchell & Goffin, D. Richard. (2006). The use of personality measures in personnel selection: What does current research support?. *Human Resource Management Review*, 16, 155-180.
- Zickar, MJ & Drasgow, F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied Psychological Measurement*, 20, 71-87.

