

SOCIAL DESIRABILITY & OVERCLAIMING

Dr. Felipe Valentini

Dr. Ricardo Primi



G R A N A D A



Socially desirable responding

- An “enhancer” gives inflated/deflated responses to some items of a questionnaire when incentives are available to perform in a given manner.
 - Person × item × situation
 - *Unconscious or personal impression management*
 - *Very hard to control...*

Interaction of:

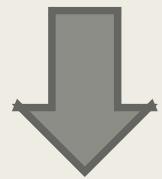
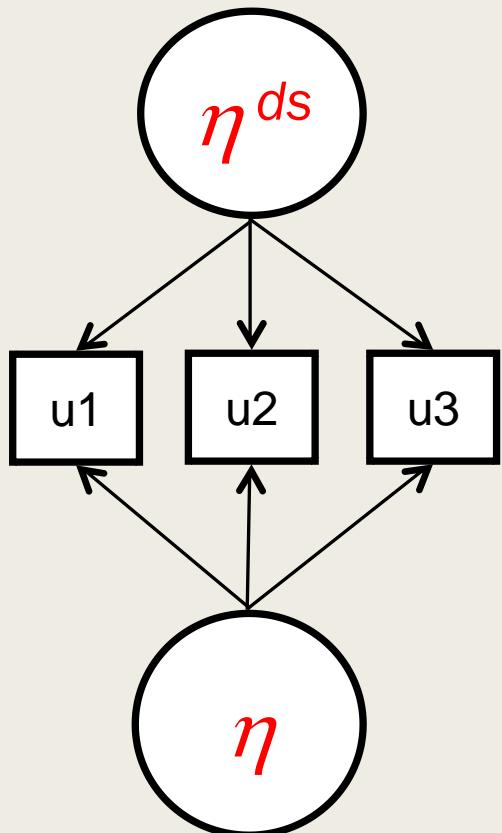
Person

Situation

related to the item content

✓ General model

- $U_{ij} = \mu_{ij} + \lambda_{ij} \eta_i$



$$U_{ij} = \mu_{ij} + \lambda_{ij} \eta_i$$

$$\mu_{ij} = \mu_0 + \lambda^{ds}_{ij} \eta^{ds}_i$$

Uniform personal DIF

Perhaps the Most Difficult Bias to Control...

- Difficult to separate from other types of response biases
 - Can be confused with extreme responses and acquiescence depending on the questionnaire item design
- Difficult to separate the bias from the true content of the fac
 - Example: I am kind.

What Social Desirability is?

- We all agree it means self-enhancement
- Is it Faking?
 - Bias in self-rating
- Is it a Genuine Personal Trait?
 - Intelligence in reading and comprehension of what is demanded in the environment

Two-component models of socially desirable responding.

[EXPORT](#)[★ Add To My List](#)[© Request Permissions](#)

Database: APA PsycArticles

Journal Article

[Paulhus, Delroy L.](#)

Citation

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598–609. <https://doi.org/10.1037/0022-3514.46.3.598>

Abstract

J. Millham and L. I. Jacobson's (1978) 2-factor model of socially desirable responding based on denial and attribution components is reviewed and disputed. A 2nd model distinguishing self-deception and impression management components is reviewed and shown to be related to early factor-analytic work on desirability scales. Two studies, with 511 undergraduates, were conducted to test the model. A factor analysis of commonly used desirability scales (e.g., Lie scale of the MMPI, Marlowe-Crowne Social Desirability Scale) revealed that the 2 major factors were best interpreted as Self-Deception and Impression Management. A 2nd study employed confirmatory factor analysis to show that the attribution/denial model does not fit the data as well as the self-deception/impression management model. A 3rd study, with 100 Ss, compared scores on desirability scales under anonymous and public conditions. Results show that those scales that had loaded highest on the Impression Management factor showed the greatest mean increase from anonymous to public conditions. It is recommended that impression management, but not self-deception, be controlled in self-reports of personality. (54 ref) (PsycINFO Database Record (c) 2016 APA, all rights reserved)

HATERS WILL SAY



IT'S PHOTOSHOP

Self-deception

Unconscious mechanism where individuals convince themselves of their positive qualities and downplay or ignore their negative traits to maintain a favorable self-image and gain social approval.

Impression Management

Attempt to control the perceptions others have of them. This concept is widely used in social psychology and involves presenting oneself in a way that aligns with social expectations and personal goals.



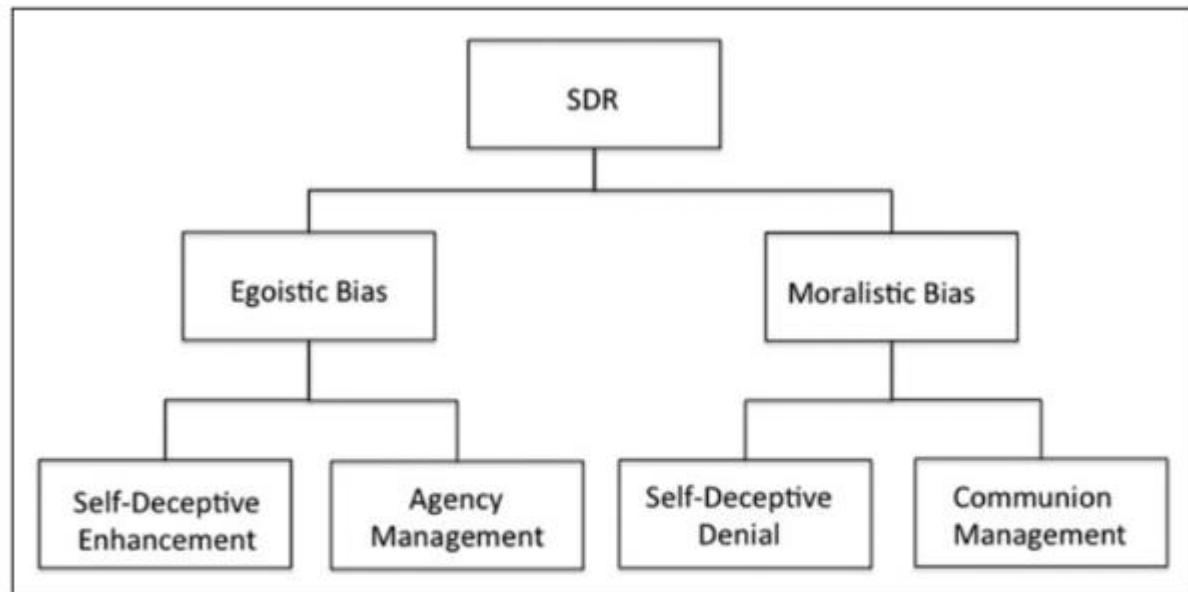


Figure 2. Paulhus's (2002) socially desirable responding (SDR) model.

solution (Wiggins, 1964). A more complex model from Paulhus (2002) is shown in Figure 2. SDR is positioned at the top of the hierarchy. Below, the two-tier model differentiates between the theme of the bias (egoistic vs. moralistic) and type of audience (self or others). Egoistic bias is defined as a “tendency to exaggerate one’s social and intellectual status” (Paulhus, 2002, p. 63), and moralistic bias as a “tendency to deny socially-deviant impulses and claim sanctimonious, ‘saint-like’ attributes” (p. 64). The outcomes of this combination are four types of SDR. Paulhus (2002) replicated the two-factor structure reported by Wiggins (1964), who had named the factors Alpha and Gamma. Alpha is associated with egoistic bias, and Paulhus showed that it was correlated with openness, extraversion, and intelligence. Gamma involves a moralistic bias and was correlated with conscientiousness and agreeableness.

The construct of social desirability: one or two dimensions?

Edward Helmes^a, Ronald R. Holden^{b,*}

^aJames Cook University, Townsville, Queensland, Australia

^bDepartment of Psychology, Queen's University, Kingston, Ontario, Canada K7L 3N6

Received 22 Aug

Table 3

One-, two-, and three-factor varimax rotated solutions for dimensions of social desirability^a

Abstract

The construct of social desirability has been studied for years ago. Recent work has focused on impression management. Other studies have tested the one- and two-dimensional models of social desirability. In this study, we test the one- and two-dimensional models of social desirability. We also examine the relationship between social desirability and other psychological constructs. Undergraduates were asked to complete a questionnaire that measured social desirability. There was even some support for the two-dimensional model. The results correspond to the expected findings. The results suggest that social desirability is a multidimensional construct. All rights reserved.

	One-factor	Two-factor		Three-factor		
		I	II	I	II	III
Marlowe-Crowne	0.79	0.86	0.05	0.83	0.13	-0.19
BIDR Self-deception	0.77	0.69	0.34	0.61	0.38	-0.28
BIDR Impression Management	0.71	0.76	0.07	0.73	0.17	-0.13
PRF Defendance	-0.52	-0.46	-0.25	-0.36	0.18	0.50
PRF Social Recognition	-0.24	0.02	-0.56	0.18	-0.22	0.71
Jackson Desirability	0.67	0.48	0.55	0.40	0.68	-0.08
JPI Self esteem	0.31	0.02	0.64	0.05	0.79	0.11
JPI Social Adroitness	-0.39	-0.33	-0.20	-0.22	0.12	0.73
JPI Value Orthodoxy	0.32	0.54	-0.34	0.61	-0.13	0.26
BPI Denial	0.61	0.68	0.01	0.66	0.10	-0.10
Edwards Desirability	0.59	0.27	0.77	0.14	0.78	0.28

Overall, the current results most strongly support a two-factor model of social desirability. The results are consistent with the self-deception and impression management model, but are more suggestive of the original need for approval and personal adjustment. Further research is needed to

- *Social Desirability Can Bias:*
- *Scores:*
 - *Meta-analyses have indicated an average inflation/deflation of $\frac{1}{2}$ standard deviation for scores on the five personality factors.*
 - *Higher for Neuroticism ($d = 0.44$) and Conscientiousness ($d = 0.45$).*
- *Ranking (in recruitment, for example)*
 - *People with higher Social Desirability (SD) tend to present scores closer to the desired profile in recruitment.*

Estimating and controlling

- ✓ Controlling using an external SD measure
- ✓ Marlowe-Crowe
- ✓ Balanced Inventory of Desirable Responding

BIDR Version 6—Form 40

Using the scale below as a guide, write a number beside each statement to indicate how much you agree with it.

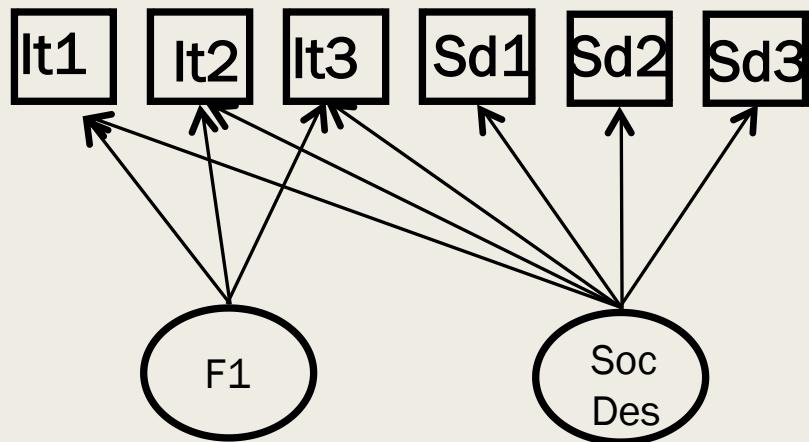
1	—	2	—	3	—	4	—	5	—	6	—	7
NOT TRUE		SOMEWHAT TRUE								VERY TRUE		

17. I am very confident of my judgments.
 *18. I have sometimes doubted my ability as a lover.
 19. It's all right with me if some people happen to dislike me.
 *20. I don't always know the reasons why I do the things I do.
 *21. I sometimes tell lies if I have to.
 22. I never cover up my mistakes.

Self-Deception

Impression
Management

- ✓ Controlling using an external SD measure



- ✓ Assume a ‘pure’ measure of SD
- ✓ Social desirability might be correlated to the content factor Agreeableness personality
- ✓ Social desirability factor might “steal” part of the true variance due to the content factor (overcontrol)

3. I don't care to know what other people really think of me.

23. There have been occasions when I have taken advantage of someone.

✓ Agreeableness

11. I never regret my decisions.

✓ Neuroticism

30. I always declare everything at customs.

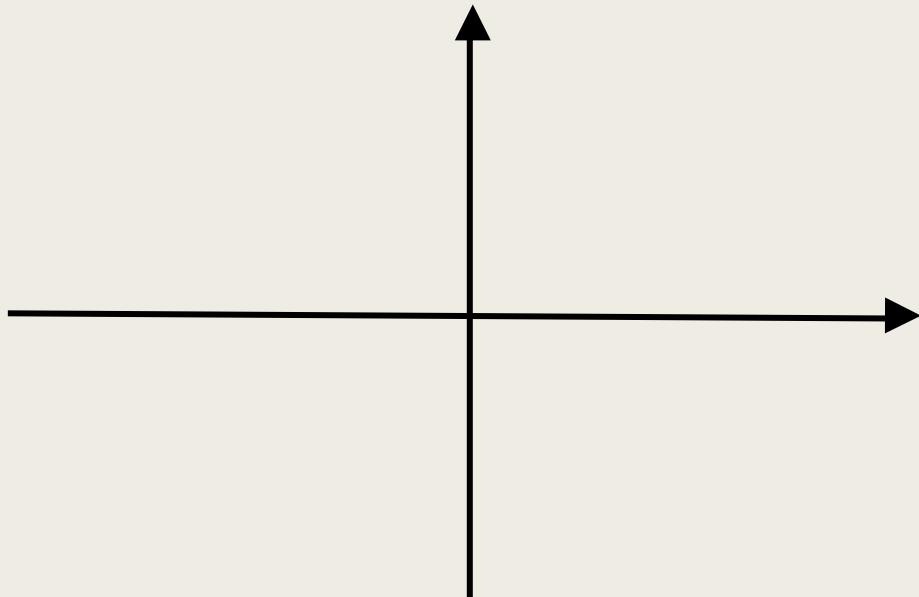
✓ Conscientiousness

Quadruplets: Descriptive versus evaluative content

UNDESIRABLE



TRAIT +



DESIRABLE



TRAIT -

Rethinking Social Desirability Scales: From Impression Management to Interpersonally Oriented Self-Control

Perspectives on Psychological Science
5(3) 243–262
© The Author(s) 2010
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691610369465
<http://pps.sagepub.com>



Liad Uziel

Psychology Department, Bar-Ilan University, Ramat-Gan, Israel

Abstract

Social desirability (specifically, impression management) scales are widely used by researchers and practitioners who bias self-reports in a self-favoring manner. These scales also serve to identify individuals at risk for social and health problems. The present review explores the evidence with regard to the ability of these scales to assess these problems. In the first part of the review, I present six criteria to evaluate impression management scales and conclude that they are unsatisfactory as measures of response style. Next, I explore what individual differences in impression management scales measure. I compare two approaches: a defensiveness approach, which argues that these scales measure a personality construct, and an adjustment approach, which suggests that impression management scales measure personal well-being and interpersonal adjustment. Data from a wide variety of fields including social behavior, health, and job performance tend to favor the adjustment approach. Finally, I argue that scales measuring impression management should be redefined as measures of interpersonally oriented self-control that identify individuals who differ in their levels of self-control, especially in social contexts.

Summary and Concluding Points

Over 2,500 published studies include scales of IM. The lion's share treats these scales as a measure of socially desirable response style and uses them to validate self-report tools. The present review questioned the utility of this practice. The ability of IM scales to detect deception has been tested over the years, and the scales have mostly failed to produce. Even if this were not the case, and IM scales had succeeded in detecting deception on some of these tests, their status as integral part of the questionnaire validation process should have been debated. The argument here is not that people do not overclaim or that there are no biases in self-reports. There is little reason to believe in such an ideal world (although some evidence implies that the situation is not as bad as typically assumed). However, the way to correct for socially desirable response bias in self-reports is probably not by statistically controlling for results on another self-report measure (cf. Paulhus & Vazire, 2007). In spite of its intuitive appeal, at present, such a procedure does more harm than good in removing valid variance of unknown magnitude.

Journal of Personality and Social Psychology

Volume 7, No. 4
Part 2 of 2 Parts

December, 1967
Whole No. 644

MONOGRAPH

TRAIT INFERENCES:

EVALUATIVE AND DESCRIPTIVE ASPECTS¹

DEAN PEABODY

Department of Psychology and Education, Swarthmore College

The common emphasis on the importance of evaluation in judgments has been critically examined in a study of inferences between traits. Sets of items were selected to remove the usual confounding between evaluative and descriptive aspects of judgment. Subjects made inferences from 90 traits defined by two opposing traits. Results revealed the secondary nature of evaluation. On 70 items where they were directly opposed, the descriptive aspects were always decisive over evaluation. Factor analysis showed that none of the factors was evaluative. General descriptive dimensions, however, could account for the evaluative consistency found in the results. Evidence suggested that evaluation is typically based on a subjective judgment of the degree of extremeness. A model was proposed to account for the systematic relation of evaluation to descriptive judgment.

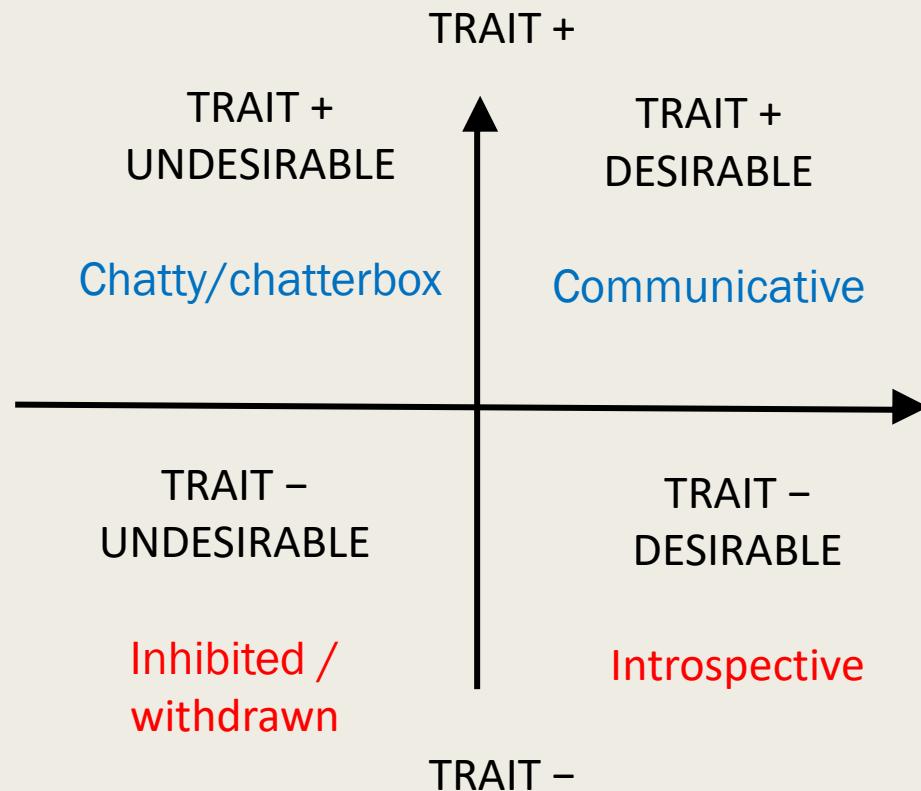
TABLE 1
SCHEME FOR UNCONFOUNDING EVALUATION AND A DESCRIPTIVE ATTRIBUTE

		Descriptive attribute	
		X	Un-X
Evaluation	+	Term 1	Term 2
	-	Term 3	Term 4

Quadruplets: Descriptive versus evaluative content

Extroversion

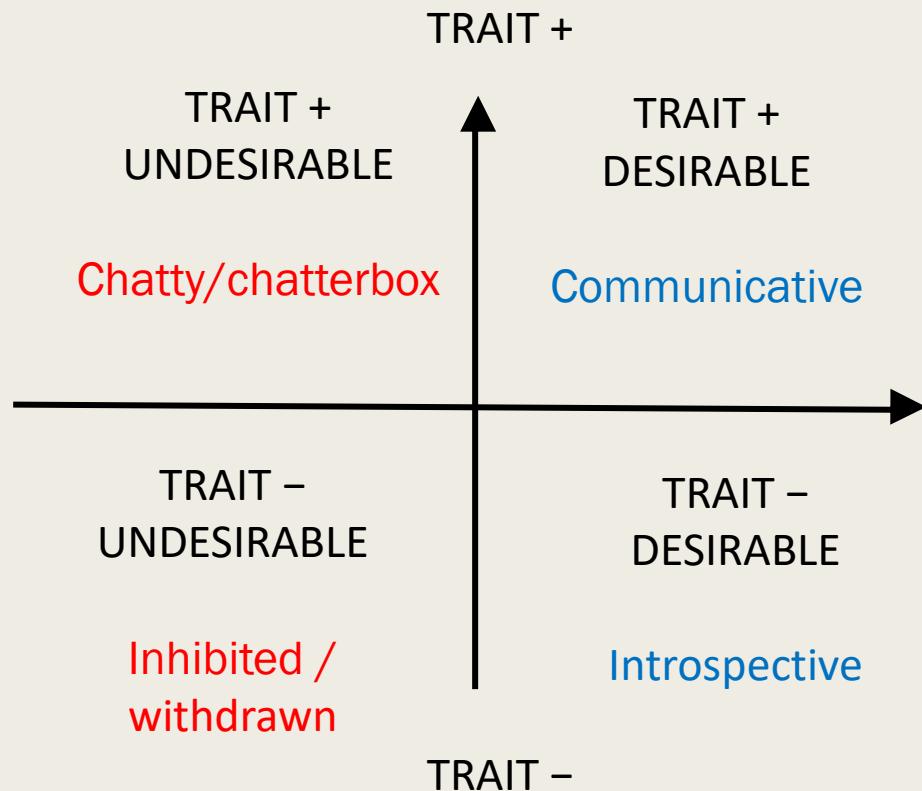
Chaty +
communicative +
Rev withdrawn +
Rev Introspective



Quadruplets: Descriptive versus evaluative content

Social
Desirability

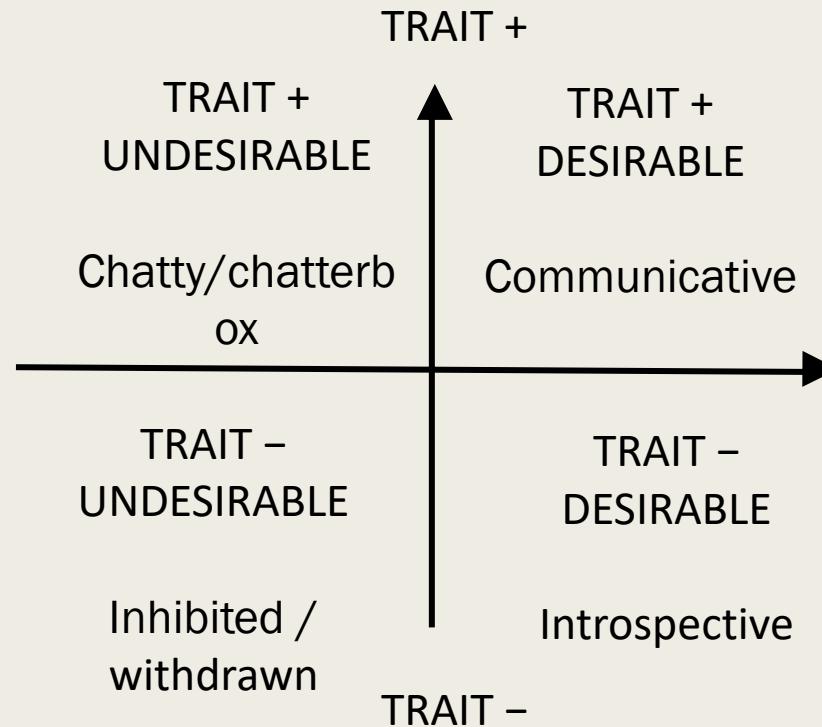
Communicative +
Introspective +
Rev Chaty +
Rev withdrawn +



Quadruplets: Descriptive versus evaluative content

Extroversion

Chaty +
communicative +
Rev withdrawn +
Rev Introspective



Social
Desirability

Rev Chaty +
Communicative +
Rev withdrawn +
Introspective +

- ✓ Why quadruplets partial DS out from TCT scores

$$U^{f+ ds-} = + \lambda \eta - \lambda^{ds} \eta^{ds}$$

$$U^{f- ds-} = - \lambda \eta - \lambda^{ds} \eta^{ds}$$

$$U^{f+ ds+} = + \lambda \eta + \lambda^{ds} \eta^{ds}$$

$$U^{f- ds+} = - \lambda \eta + \lambda^{ds} \eta^{ds}$$

- ✓ Why quadruplets partial DS out from TCT scores
 - ✓ Summing up the items
 - ✓ First... reverse back the negative keyed (descriptive content)

$$U^{f+ \text{ ds-}} = + \lambda \eta - \lambda^{ds} \eta^{ds}$$

$$U^{f- \text{ ds-}} = - \lambda \eta - \lambda^{ds} \eta^{ds}$$

$$U^{f+ \text{ ds+}} = + \lambda \eta + \lambda^{ds} \eta^{ds}$$

$$U^{f- \text{ ds+}} = - \lambda \eta + \lambda^{ds} \eta^{ds}$$



$$U^{f+ \text{ ds-}} = + \lambda \eta - \lambda^{ds} \eta^{ds}$$

$$\text{Rev}U^{f- \text{ ds-}} = + \lambda \eta + \lambda^{ds} \eta^{ds}$$

$$U^{f+ \text{ ds+}} = + \lambda \eta + \lambda^{ds} \eta^{ds}$$

$$\text{Rev}U^{f- \text{ ds+}} = + \lambda \eta - \lambda^{ds} \eta^{ds}$$

~~Sum = $4 \lambda \eta - 2\lambda^{ds} \eta^{ds} + 2\lambda^{ds} \eta^{ds}$~~

- ✓ Biggest issue of TCT for quadruplets:
 - ✓ Assume equal factor loadings within quadruplets

- ✓ Solution = use SEM or ESEM

European Journal of Personality, Eur. J. Pers. (2011)
 Published online (wileyonlinelibrary.com) DOI: 10.1002/per.839

The General Factor of Personality and Evaluation

ERIK PETTERSSON*, ERIC TURKHEIMER, ERIN E. HORN and ANDREW R. MENATTI
Department of Psychology, University of Virginia, Charlottesville, VA 22904-4400, USA

Abstract: According to the proposal of the general factor of personality (GFP), socially desirable personality traits have been selected for throughout evolution because they increase fitness. However, it remains unknown whether people high on this factor actually behave in socially desirable ways or whether they simply endorse traits of positive valence. We separated these two sources of variance by having 619 participants respond to 120 personality adjectives organised into 30 quadruples balanced for content and valence (e.g. unambitious, easy-going, driven and workaholic tapped the trait achievement-striving). An exploratory six-factor solution fit well, and the factors resembled the Big Five. We subsequently extracted a higher-order factor from this solution, which appeared similar to the GFP. A Schmid-Leiman transformation of the higher-order factor, however, revealed that it clustered items of similar valence but opposite content (e.g. at the negative pole, unambitious and workaholic), rendering it an implausible description of evolved adaptive behaviour. Isolating this evaluative factor using exploratory structural equation modelling generated factors consisting of items of similar descriptive content but different valence (e.g. driven and workaholic), and the correlations among these factors were of small magnitude, indicating that the putative GFP capitalises primarily on evaluative rather than descriptive variance. Implications are discussed. Copyright © 2011 John Wiley & Sons, Ltd.

Key words: general factor of personality; Big One; evaluation; valence; social desirability

Table 4 Standardised exploratory structural equation modelling loadings based on target rotation

Item	Agreeableness	
	Item	Loading
Happ	Gullible	0.35
Secu	Submissive	0.31
Chee	Agreeable	0.30
Wan	Trusting	0.28
Coop	Conservative	0.27
Socia	Weak	0.23
Dow	Gushy	0.22
Conf	Cheerful	0.22
Ratio	Timid	0.21
Eas	Inert	0.21

The first factor (Table 4), which by design was fixed to evaluation, included positive loadings on *happy*, *cooperative* and *rational*, and negative loadings on *depressed*, *selfish* and *weak*. This dimension clustered items of opposite descriptive content but equal valence. For example, the evaluative factor comprised loadings of the same sign for *sluggish* (loading = -0.59) and *manic* (loading = -0.50), *grim* (loading = -0.61) and *frivolous* (loading = -0.37), *suspicious* (loading = -0.39) and *gullible* (loading = -0.24), and *self-deprecating* (loading = -0.43) and *conceited* (loading = -0.47). The evaluation factor also included positive loadings for pairs with opposite descriptive content, including *orderly* (loading = 0.31) and *flexible* (loading = 0.29), *modest* (loading = 0.26) and *assertive* (loading = 0.23), *easy-going* (loading = 0.44) and *driven* (loading = 0.40), and *intellectual* (loading = 0.34) and *down-to-earth* (loading = 0.47). Although one can imagine circumstances in which any one of these pairs might describe the actual behaviour of a respondent, it is unlikely that a behavioural style could be characterised by so many paradoxical pairs of descriptors, so this dimension may be better interpreted as a response bias rather than as a description of a consistent behavioural propensity.

The remaining exploratory factors were forced to be inde-

pendent from the Agreeableness factor. The second factor had positive loadings on *modest*, *assertive*, *easy-going*, *driven*, *intellectual*, *down-to-earth*, *inert* and *weak*, and negative loadings on *conservative*, *disciplined*, *rigid* and *restricted*. Thus, after isolating participants' tendency to endorse positive or negative markers—a tendency that in and of itself seems unlikely to describe a consistent behavioural pattern—the items clustered by their descriptive content, regardless of their valence. It is interesting to note that these factors overlap quite well with

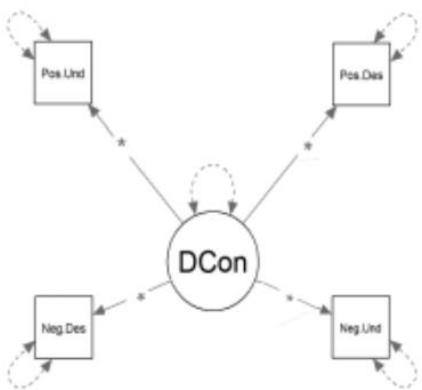
Social desirability, social-emotional competencies and intelligence: using quadruplets to estimate evaluative and descriptive content

Left running head: T. D. L. F. FRANCHI ET AL.

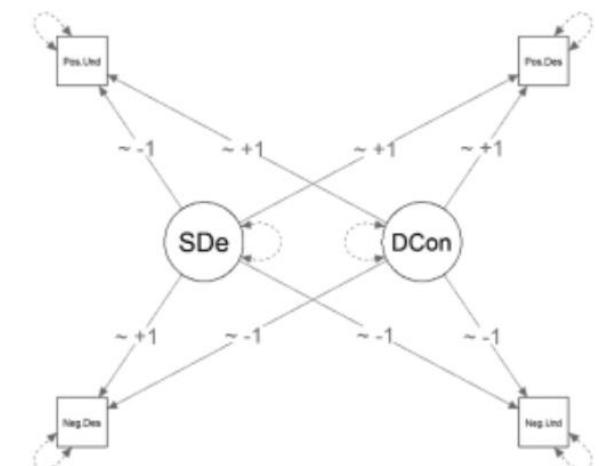
Short title : International Journal of Testing

AQ0

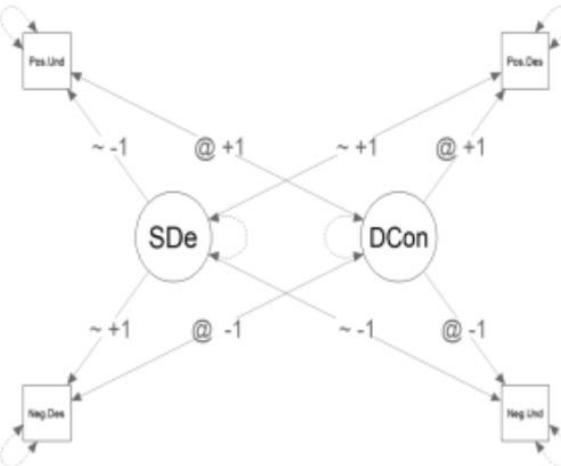
AQ6  Tatiana di Lucia Faion [Please change the affiliation of the first author to match that of the others: Graduate School of Psychology, University São Francisco, Campinas, São Paulo, Brazil] Franchi^a,  Felipe Valentini^b,  Leonardo Botinhon de Campos^b,  Letícia da Silva de Souza^b,  Pedro Vanni^b,  Leonardo de Barros Mose^b and  Ricardo Primi^b

A

χ^2 (gl) = 271.57 (100); CFI = .88;
TLI = .76; RMSEA = .08

B

χ^2 (gl) = 221.20 (85); CFI = .95;
TLI = .89; RMSEA = .06

C

χ^2 (gl) = 795.07 (146); CFI = .76;
TLI = .69; RMSEA = .10

**B. General Factor Control -
Target Rotation for all Factors**

Quad & Items	Descrip Factor	SDe	SDe	EO	AM	OM	SM	ER
1a	+	-	-.01	.45*				-.32*
1b	-	-	-.39*	-.54*			-.32*	-.33*
1c	+	+	.63*	.68*				
1d	-	+	.44*	-.45*			-.31*	
5a	+	+	.30*		.61*			
5b	+	-	.14*		.66*		-.27*	-.23*
5c	-	+	.21*		-.37*			
5d	-	-	-.01		-.6*			
10a	+	+	.41*		.73*		-.25*	
10b	-	+	.41*		-.43*	.21*		
10c	-	-	-.25*		-.48*	-.25*		
10d	+	-	-.16*		.47*	-.57*		
4a	-	-	-.36*			-.74*		
4b	-	+	.05		.32*	-.61*		
4c	+	-	.46*			.30*	-.32*	
4d	+	+	.66*			.36*		
8a	+	+	.62*				.37*	
8b	+	-	.35*	-.32*		-.37*	.33*	
8c	-	+	.23*	-.26*			-.46*	
8d	-	-	-.47*				-.72*	

item 10b "I like to do things objectively, without fantasizing" initially loaded on self-management; after including social desirability control, it negatively loaded on open-mindedness.

I have a lot of imagination

I like to do things objectively, without fantasizing

I'm a person without creativity

I fantasize too much and I get lost in my imagination

**B. General Factor Control -
Target Rotation for all Factors**

Quad & Items	Descrip Factor	SDe	SDe	EO	AM	OM	SM	ER
1a	+	-	-.01	.45*				-.32*
1b	-	-	-.39*	-.54*			-.32*	-.33*
1c	+	+	.63*	.68*				
1d	-	+	.44*	-.45*			-.31*	
5a	+	+	.30*		.61*			
5b	+	-	.14*		.66*		-.27*	-.23*
5c	-	+	.21*		-.37*			
5d	-	-	-.01		-.6*			
10a	+	+	.41*		.73*			-.25*
10b	-	+	.41*		-.43*	.21*		
10c	-	-	-.25*		-.48*	-.25*		
10d	+	-	-.16*		.47*	-.57*		
4a	-	-	-.36*			-.74*		
4b	-	+	.05		.32*	-.61*		
4c	+	-	.46*			.30*	-.32*	
4d	+	+	.66*			.36*		
8a	+	+	.62*				.37*	
8b	+	-	.35*	-.32*		-.37*	.33*	
8c	-	+	.23*	-.26*			-.46*	
8d	-	-	-.47*				-.72*	

The notable exception was in the positive content yet undesirable items.

For instance, item 4c ("I am obsessive - strict- about staying focused on what I am doing at the moment") was expected to align with the undesirable pole but instead loaded on the desirable pole.

My biggest flaw is that I am a workaholic

Our main conclusions are:

1. Uncontrolled social desirability levels can distort the internal structure of items;
2. Balanced quadruplets of items are effective for controlling social desirability.
3. Statistics alone may not solve the problem; we need to design items carefully

	OM	AM	SM	EO	TRAE Intelligence
Open-mindedness	-				0.14*
Amity	0.35*	-			0.08
Self-Management	-0.01	-0.04	-		-0.06
Engaging with Others	-0.07	-0.15*	0.26*	-	-0.26*
Emotional Resilience	0.07	0.05	0.16	0.10	-0.03
Social Desirability	-	-	-	-	0.24*

Note. OM = Open-mindedness; AM = Amity; SM = Self-Management; EO = Engaging with Others; ER = Emotional Resilience.

* = Statistically significant correlation ($p < 0.05$).

Other empirical study

In Brazil, every driver candidate must be submitted to a psychological assessment:

- first drive license – once
- professional drivers – once every 5 years

We applied the Big-Five Inventory (BFI) for three groups:

- **High stakes (n = 78)** – professional drivers, for which the personality assessment is a condition for keeping the professional licence.
- **Medium stakes (n=70)** – candidates for the first drive license. “Fail” on the assessment would stop them (temporarily) to get the license, however it would not interfere on their jobs.
- **Low stakes (n=148)** – College students (control group).

Empirical Illustration

We use only 3 factors (Neuroticism, Conscientiousness, and Agreeableness)

2 triplets for each factor

F	SD	Item
+	+	I think on Other people needs and I like to help him
-	-	I am selfish, think only on my needs
+	-	I think too much on the others needs, I am a fool
-	+	I can disconnect from the others problems

Structure Analysis

Agreeableness

I can disconnect from the others problems



I think too much on the others needs, I am a fool



Items	Fact Cont	SD	CFA		Bifactor	
			Cont	DS	Cont	DS
4	-	-	-0,41	-0,06	-0,36	
8	-	+	0,15	-0,32	0,18	
9	+	+	0,52	0,03	0,44	
10	-	+	-0,2	-0,21	-0,15	
13	+	+	0,35	0,63	0,25	
16	+	-	-0,01	0,68	-0,13	
28	-	-	-0,68	-0,42	-0,58	
31	+	-	0,07	0,01	0,03	

Structure Analysis

Neuroticism

I have too much
energy, that I get
nervous When
something go wrong

Items	Fact Cont	SD	CFA		Bifactor	
			Cont	DS	Cont	DS
1	+	-	0,77	0,19	-0,75	
14	-	+	-0,74	-0,35	0,65	
18	+	-	0,85	0,41	-0,73	
20	-	+	-0,79	-0,53	0,63	
30	+	+	0,51	0,08	0,51	
33	-	-	-0,4	-0,49	-0,24	
34	-	-	-0,09	-0,43	-0,08	
36	+	+	0,52	0,15	-0,51	

Latent Means

We set the item parameters as invariant across groups

Grupos	AFC	Bifactor
Low stakes (college)		
Agreeableness	0,00	0,00
Conscientiousness	0,00	0,00
Neuroticism	0,00	0,00
SD	0,00	
Medium stakes (1 licence)		
Agreeableness	1,16	0,14
Conscientiousness	0,81	0,02
Neuroticism	-1,05	-0,27
SD	1,29	
High stakes (Drivers)		
Agreeableness	1,33	0,09
Conscientiousness	1,07	0,32
Neuroticism	-2,16	-0,91
SD	1,46	

Issues... Nothing is perfect...

Quadruplets are hard to build

- Not all items can be manipulated

- Some manipulations might insert content multidimensionality

- Some manipulations ends up non-natural language

Future directions...

Submitted Paper

Peabody with MIMIC as a theoretical sound control for social desirability: simulations for Likert and Forced-Choice

Running Head: Social Desirability: Peabody with MIMIC

Authors:

Rafael Valdece Sousa Bastos

Felipe Valentini

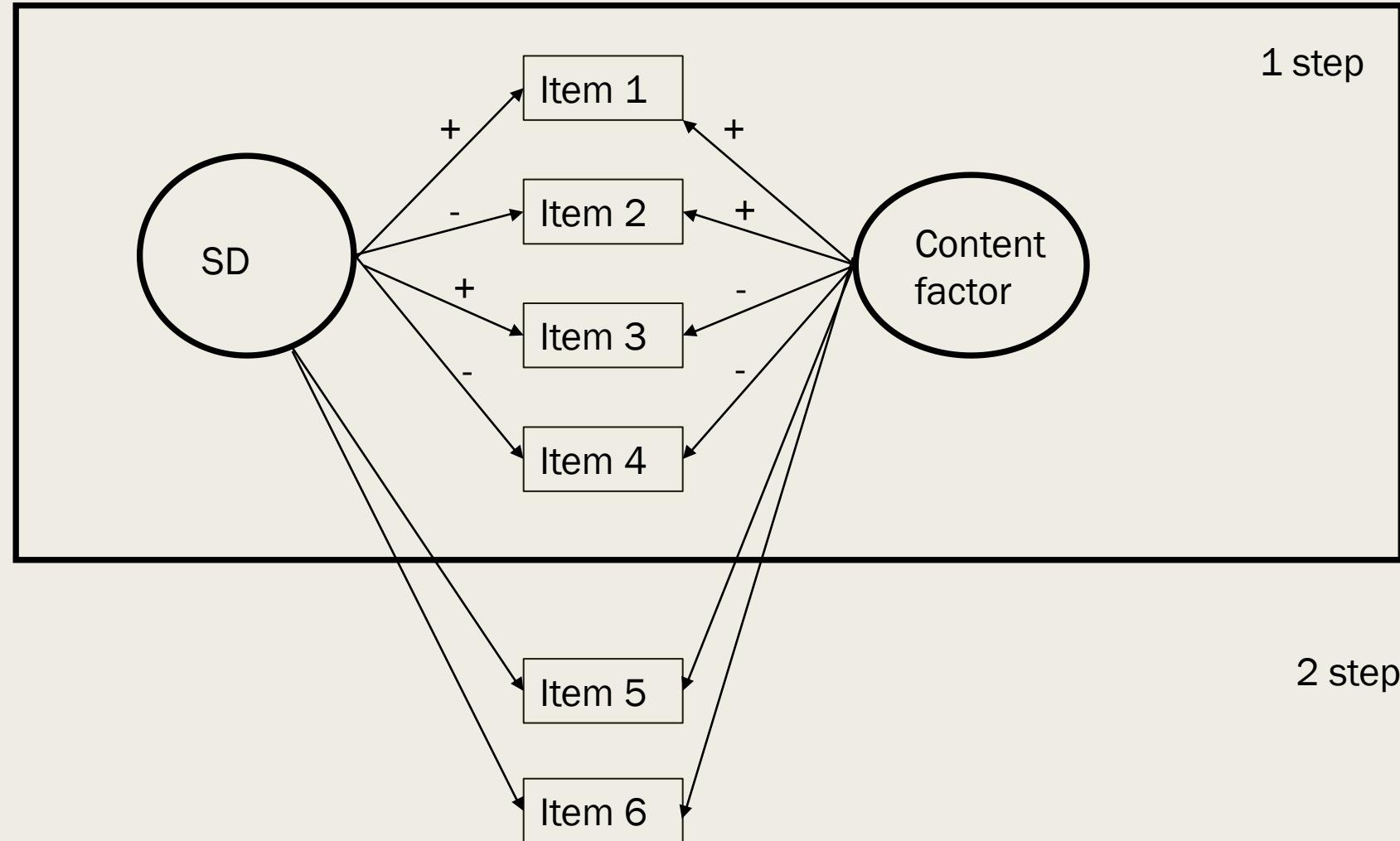
Future directions...

Submitted Paper

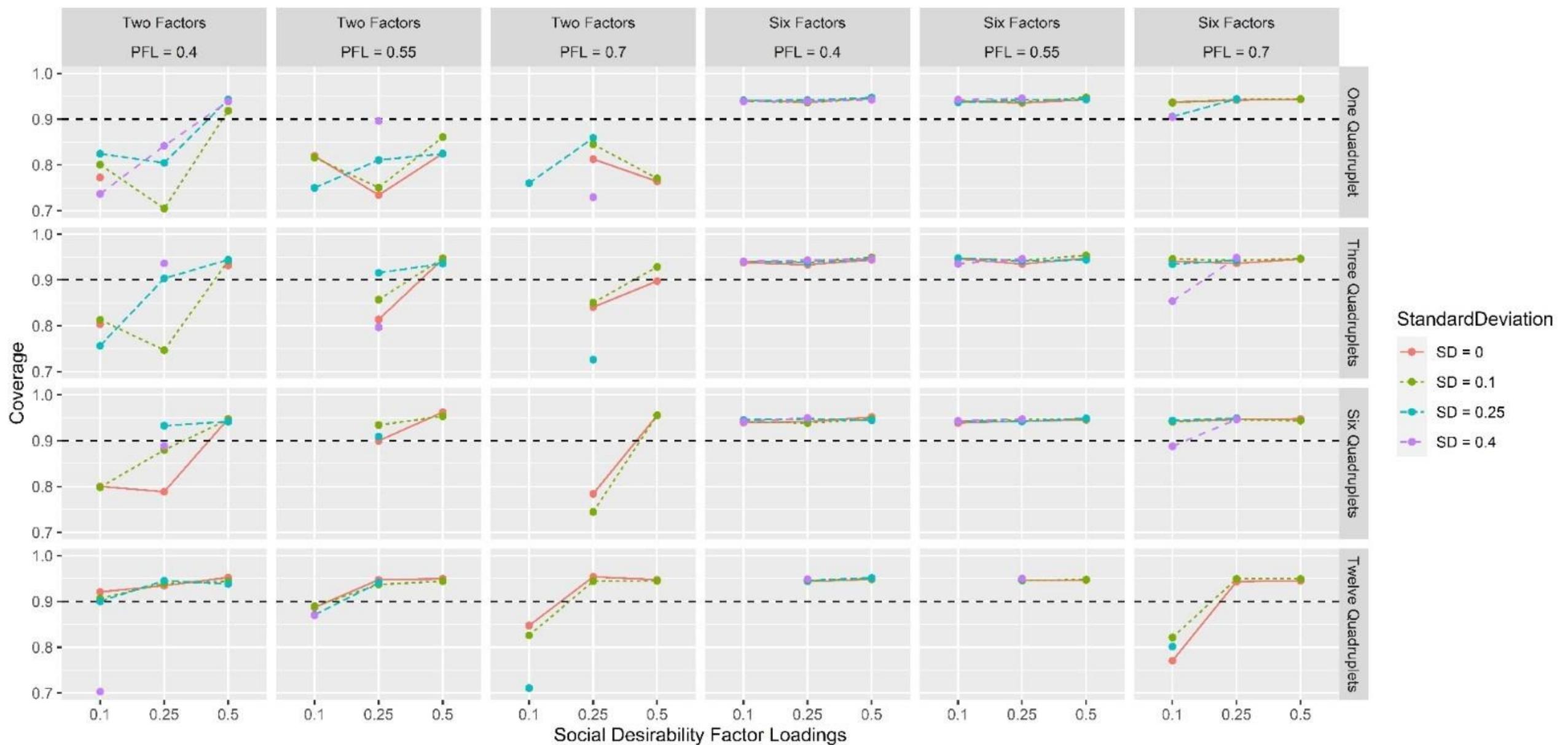
Items in and out of the quadruplet

- 1 step – estimate item parameters within the quadruplets;
- 2 step – fix the parameters within quadruplets (estimated on the 1 step),
and regress them to the remaining items (out from the quadruplet)

Only good manipulations get into the quadruplets...



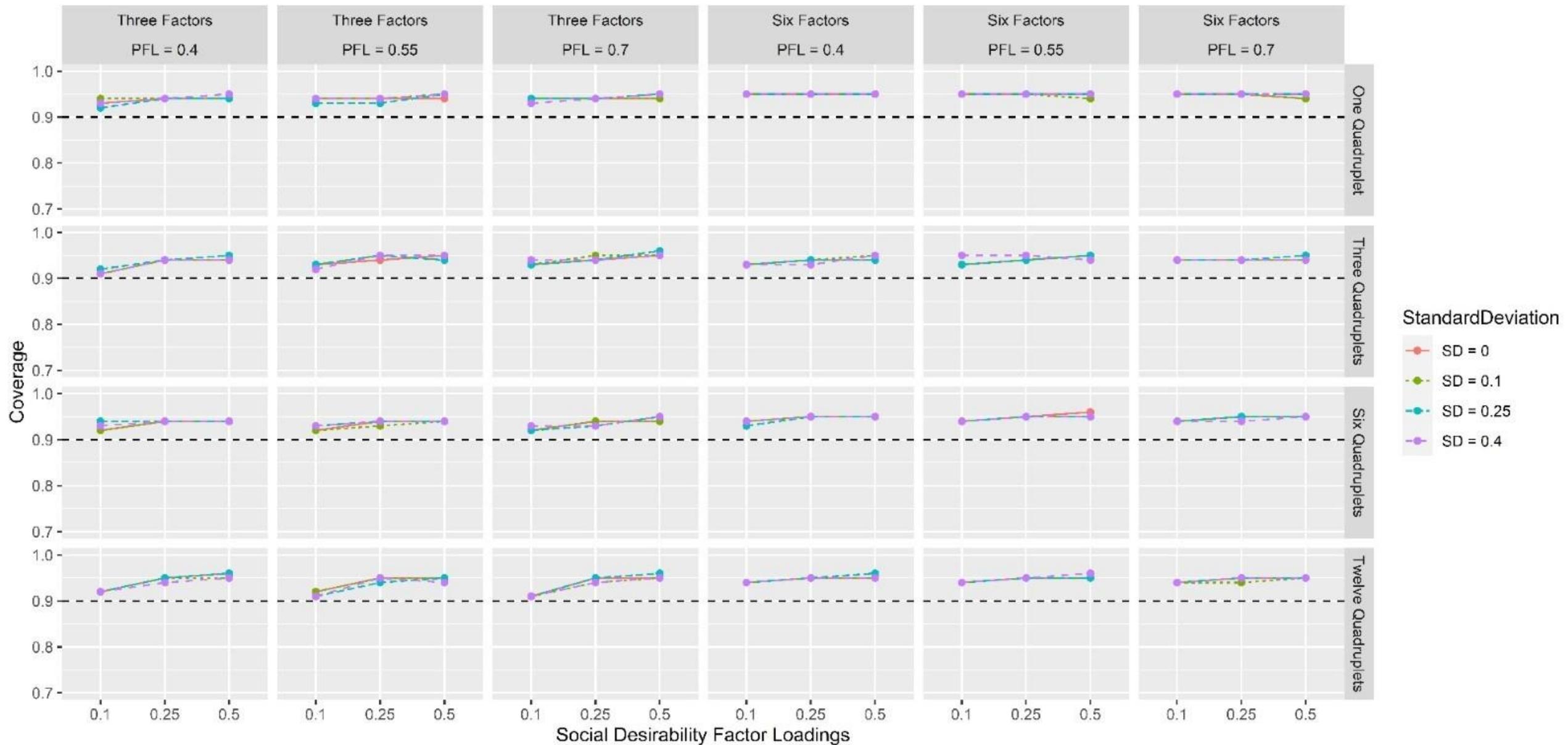
Coverage of Social Desirability Regressions



Note. PFL = Peabody Quadruplets Factor Loadings on Random Constructs; SD = Standard Deviation from the Mean of Peabody's Quadruplets Factor Loadings on Random Constructs. The dashed black line represents the cut-off value.

Using Forced-choice items

Coverage of Social Desirability Regressions



The neutralization approach

- “Simply put, the valence of an existing item is changed by rephrasing it in a way that makes the item more neutral. Positive items are framed less positively, whereas negative items are framed less negatively, making all items relatively more neutral.” (Bäkström, Björklund, & Larsson, 2012).

$$\blacksquare \quad U_{ij} = \mu_{ij} + \lambda_{ij} \eta_i + \lambda^{ds}_{ij} \eta^{ds}_i \quad (1)$$

Keep this parameter unchanged

Make this parameter as close to 0 as possible

Personality and Social Psychology

Social desirability in personality inventories: Symptoms, diagnosis and prescribed cure

MARTIN BÄCKSTRÖM and FREDRIK BJÖRKLUND

Department of Psychology, Lund University, Lund, Sweden

Bäckström, M. & Björklund, F. (2013). Social desirability in personality inventories: Symptoms, diagnosis and prescribed cure. *Scandinavian Journal of Psychology* 54, 152–159.

An analysis of social desirability in personality assessment is presented. Starting with the symptoms, Study 1 showed that mean ratings of graded personality items are moderately to strongly linearly related to social desirability (Self Deception, Impression formation, and the first Principal Component), suggesting that item popularity may be a useful heuristic tool for identifying items which elicit socially desirable responding. We diagnose the cause of socially desirable responding as an interaction between the evaluative content of the item and enhancement motivation in the rater. Study 2 introduced a possible cure; evaluative neutralization of items. To test the feasibility of the method lay psychometrists (undergraduates) reformulated existing personality test items according to written instructions. The new items were indeed lower in social desirability while essentially retaining the five factor structure and reliability of the inventory. We conclude that although neutralization is no miracle cure, it is simple and has beneficial effects.

Key words: Personality assessment, test items, self-ratings, social desirability.

Martin Bäckström, Department of Psychology, Lund University, Box 213, 221 00, Lund, Sweden. E-mail: martin.backstrom@psychology.lu.se

European Journal of Personality, Eur. J. Pers. 28: 620–633 (2014)

Published online 5 July 2014 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/per.1960

Criterion Validity is Maintained When Items are Evaluatively Neutralized: Evidence from a Full-scale Five-factor Model Inventory

MARTIN BÄCKSTRÖM*, FREDRIK BJÖRKLUND and MAGNUS R. LARSSON

Department of Psychology, Lund University, Sweden

Abstract: The original version and an evaluatively neutralized version (with items rephrased to reduce popularity) of a personality inventory were compared. The results revealed (i) similar criterion validity across three different sets of self-rated behaviours, (ii) stronger relations to the rated social desirability of criteria for the original version and (iii) less correlation between factors for the neutralized version. We take the results to indicate that evaluative neutralization is a viable technique for reducing social desirability in self-ratings. Implications for test construction are discussed.
Copyright © 2014 European Association of Personality Psychology

Key words: social desirability; personality assessment; five-factor model; criterion validity



Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally

Martin Bäckström *, Fredrik Björklund, Magnus R. Larsson

Department of Psychology, Lund University, Box 213, 22100 Lund, Sweden

ARTICLE INFO

Article history:
Available online 31 December 2008

Keywords:
Personality
Social desirability
Factor analysis
Big Five

ABSTRACT

The factors in self-report inventories measuring the five-factor model (FFM) correlate with one another although they theoretically should not. Study 1 showed, across three different FFM-questionnaires, that almost all of the common variance between factors can be attributed to a single general factor related to social desirability. In Study 2, simple rephrasing of items from a FFM-questionnaire made them substantially less socially desirable, while the inventory's empirical (five factor) structure remained the same. Participants low in social desirability showed little difference between how they responded to the original items vs. the neutral items. For participants high in social desirability the difference was considerably larger. The simplicity of reducing social desirability in self-rating inventories of the FFM, and the usefulness of this endeavor, is discussed.

© 2008 Elsevier Inc. All rights reserved.



Is the general factor of personality based on evaluative responding? Experimental manipulation of item-popularity in personality inventories

Martin Bäckström *, Fredrik Björklund

Department of Psychology, Lund University, Box 213, SE-221 00 Lund, Sweden

ARTICLE INFO

Article history:
Received 21 November 2015
Received in revised form 12 February 2016
Accepted 22 February 2016
Available online xxxx

Keywords:
Self-ratings
Personality
Big Five
General factor of personality

ABSTRACT

The general factor of personality (GFP) is understood as a hierarchically superordinate factor, which suggests that it and the subordinate personality traits are mutually dependent on one another. If a personality inventory captures the subordinate traits the GFP should appear too. Likewise, manipulating the GFP should affect the subordinate traits and vice versa. The current study was an attempt to uniquely affect the size of the GFP by manipulating the evaluativeness of the inventory. First we estimated a general factor in a standard (evaluative) personality inventory, and found it to be robust. Then we estimated it in an inventory with evaluatively neutralized items, and found it to be unreliable. Finally, the neutralized inventory was made evaluative again. As expected, the GFP reappeared, suggesting the increased evaluative content to be the cause. Results are discussed in relation to personality assessment and to higher order factors in personality theory. It is suggested that for determining whether the GFP exists or not researchers should turn to other measures than personality inventories.

© 2016 Elsevier Ltd. All rights reserved.



- Start with items with good face validity.
- Have them rated on their social desirability on a scale from 1 to 9 by a sample of naïves
- Compute mean scores for each item.
- Items with a mean around 5 (+-) are acceptable; otherwise, they should be rewritten
- Repeat until
 - *The overall items mean is not significantly distinct from 5*
 - *The standard deviation is < 1*
 - *Social desirability scores are uncorrelated to theoretical item keying (each item coded 1 or -1)*

1 = Totally undesirable



9 = Totally desirable



(Bäckström & Björklund, 2013; Bäckström, Björklund, & Larsson, 2009)

Neutralization examples

I often cause the impression of being someone snobish and arrogant.

Mean score = **2.0**.

I have few worries about what if other people dislike me.

Mean score = **5.0**.

I am a peaceful person that makes everything to avoid using violence. (reversed)

Mean score = **7.9**.

Even if someone teases me, I avoid behaving aggressively. (reversed)

Mean score = **6.125**.

Even when aggressiveness is needed, I feel unable to express it. (reversed)

Mean score = **5.2**.

Other approaches

Use Mixture Factor Models

factors to capture descriptive
class to capture faking

Machine Learning



Volume 9, Issue 3
June 2021

JOURNAL ARTICLE

An Evaluation of Mixture Confirmatory Factor Analysis for Detecting Social Desirability Bias

[Get access >](#)

Alexandru Cernat , Caroline Vandenplas

Journal of Survey Statistics and Methodology, Volume 9, Issue 3, June 2021, Pages 496–522, <https://doi.org/10.1093/jssam/smaa032>

Published: 22 December 2020

Detecting Social Desirability Bias Using Factor Mixture Models

Walter L. Leite

*Research and Evaluation Methodology Program, College of Education
University of Florida*

Lou Ann Cooper

*Office of Medical Education, College of Medicine
University of Florida*

Based on the conceptualization that social desirable bias (SDB) is a discrete event resulting from an interaction between a scale's items, the testing situation, and the respondent's latent trait on a social desirability factor, we present a method that makes use of factor mixture models to identify which examinees are most likely to provide biased responses, which items elicit the most socially desirable responses, and which external variables predict SDB. Problems associated with the common use of correlation coefficients based on scales' total scores to diagnose SDB and partial correlations to correct for SDB are discussed. The method is demonstrated with an analysis of SDB in the Attitude toward Interprofessional Service-Learning scale with a sample of students from health-related fields.

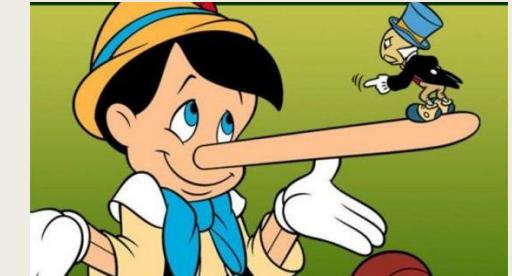
Enhancing the Detection of Social Desirability Bias Using Machine Learning: A Novel Application of Person-Fit Indices

[Sanaz Nazari](#)  , [Walter L. Leite](#) , and [A. Corinne Huggins-Manley](#) [View all authors and affiliations](#)

[OnlineFirst](#) | <https://doi.org/10.1177/00131644241255109>

Overclaiming

Tendency to exaggerate in self-reporting of skills and competencies and overestimated self-efficacy.



Part of faking bias

The over-claiming technique: Measuring self-enhancement independent of ability.

EXPORT ★ Add To My List ⌂ Request Permissions

Database: APA PsycArticles Journal Article

Paulhus, Delroy L., Harms, P. D., Bruce, M. Nadine, Lysy, Daria C.

Citation

Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology*, 84(4), 890–904. <https://doi.org/10.1037/0022-3514.84.4.890>

Abstract

Over-claiming is a concrete operationalization of self-enhancement based on respondents' ratings of their knowledge of various persons, events, products, and so on. Because 20% of the items are nonexistent, responses can be analyzed with signal detection formulas to index both response bias (over-claiming) and accuracy (knowledge). Study 1 demonstrated convergence of over-claiming with alternative measures of self-enhancement but independence from cognitive ability. In Studies 2-3, the validity of the over-claiming index held even when respondents were (a) warned about the foils or (b) asked to fake good. Study 3 also showed the utility of the over-claiming index for diagnosing faking. In Study 4, the over-claiming technique was applied to the debate over the adaptive value of positive illusions. (PsycInfo Database Record (c) 2023 APA, all rights reserved)

Journal of Personality and Social Psychology
Journal TOC
Get
Cited by
Citation stats
401 18

Related Content
The impact of ignoring foils on the self-reporting of knowledge

Overclaiming on personality questionnaires.

EXPORT ★ Add To My List ⌂ Request Permissions

Database: APA PsycInfo Chapter

Paulhus, Delroy L.

Citation

Paulhus, D. L. (2012). Overclaiming on personality questionnaires. In M. Ziegler, C. MacCann, & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 151–164). Oxford University Press.

Abstract

In this chapter, the term *faking* will be interpreted in the broader sense of self-presentation, that is, motivated distortion of self-reports. At the private level, self-presentation is typically labeled *self-deception* (Paulhus, 1984) or self-enhancement (Baumeister, 1982). At the public level, self-presentation is most commonly labeled impression management (Paulhus, 1984). I will treat them together because both forms of positive self-presentation constitute a threat to the validity of personality scales. Moreover both forms of positive self-presentation can be measured with the overclaiming technique (OCT). The OCT was designed to measure knowledge exaggeration and knowledge accuracy simultaneously and independently (Paulhus, Harms, Bruce, & Lysy, 2003; Paulhus & Harms, 2004). Respondents are asked to rate their familiarity with a set of topics relevant to a content domain (e.g., academic facts, workplace items, consumer products). Critical to the technique is the inclusion of some items that do not actually exist (i.e., *foils*). A respondent's knowledge exaggeration and accuracy are calculated from two values: (a) the proportion of real items rated as familiar and (b) the proportion of foils rated as familiar. Exaggeration is indexed by the respondent's tendency to claim familiarity with items (especially foils) whereas accuracy is indexed by the respondent's ability to distinguish real items from foils. To the extent that an audience is salient, exaggeration can be interpreted as impression management; otherwise, it is best interpreted as self-deceptive enhancement. Details about the history, psychometrics, and applications of the OCT are fleshed out in the following section. (PsycInfo Database Record (c) 2022 APA, all rights reserved)

Overclaiming

How confident you are about your knowledge about?

9th Symphony by Beethoven

Book White nights by Dostoyevsky

Book The exile of the foreigner by Camus

Paint “The Dream of the Eternal Summer” by van Gogh

Overclaiming

How confident you are about your knowledge about?

9th Symphony by Beethoven

Book White nights by Dostoyevsky

Book The exile of the foreigner by Camus

Paint “The Dream of the Eternal Summer” by van Gogh

foils

Overclaiming

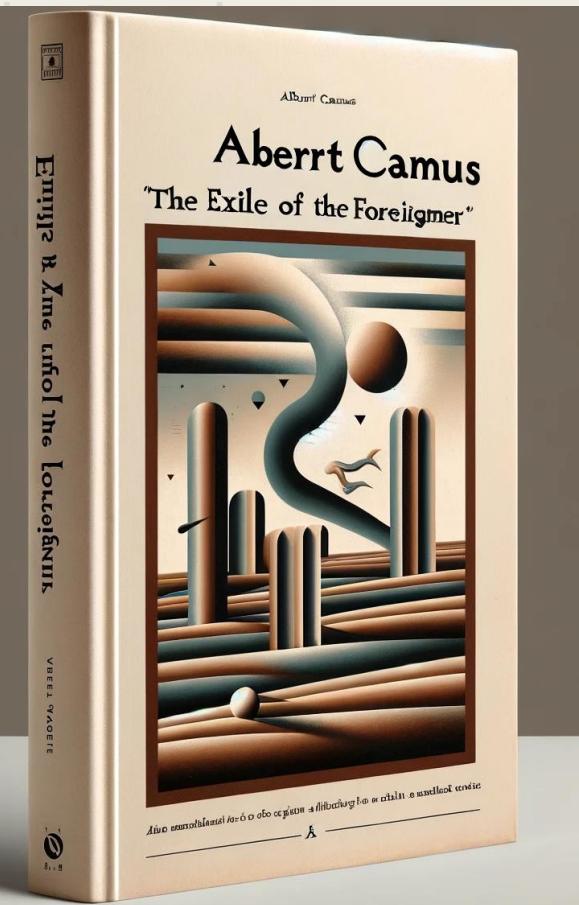
How confident you are about your knowl

9th Symphony

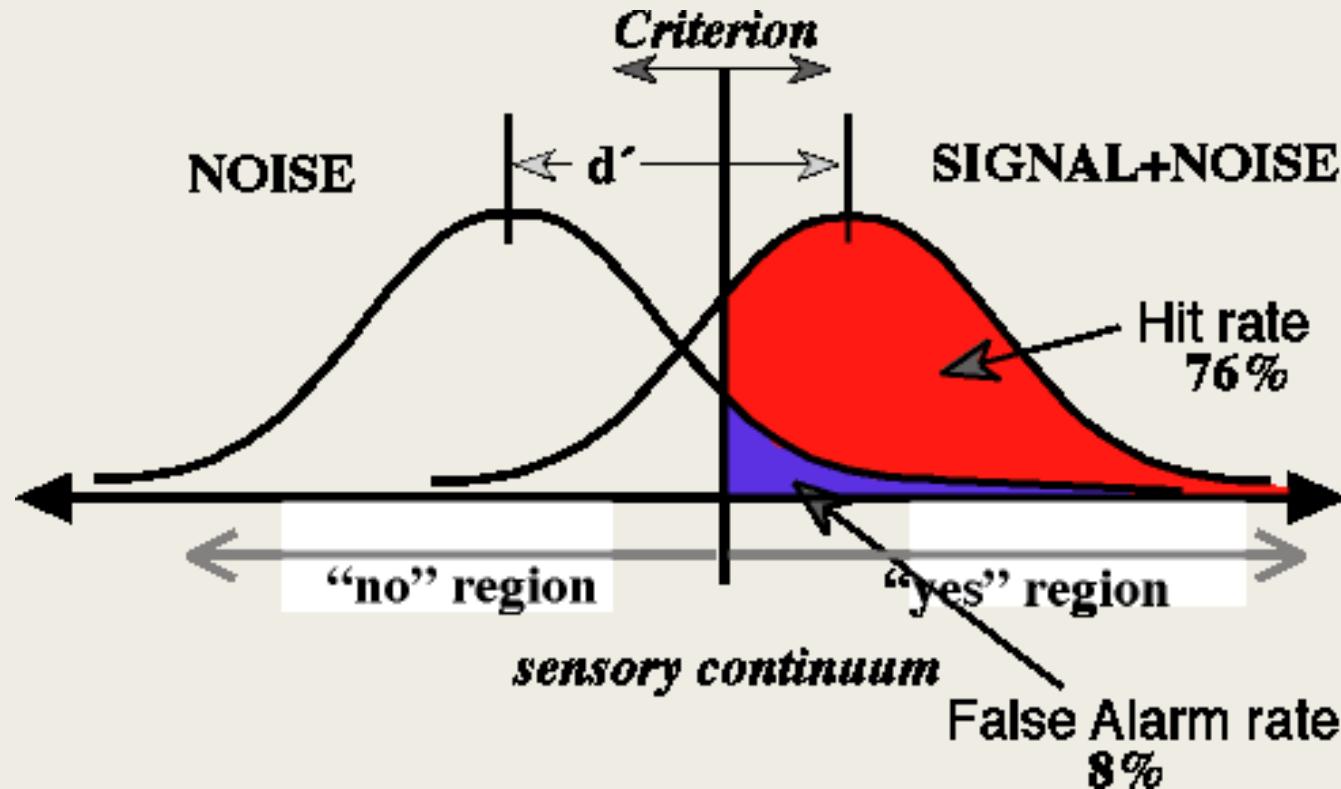
Book White ni

Book The exile

Paint "The Dream



	Test taker 1	Test taker 2	Test taker 3
9 th Symphony by Beethoven	Yes	Yes	Yes
Book White nights by Dostoyevsky	Yes	No	Yes
Book The exile of the foreigner by Camus	Yes	No	No
Paint “The Dream of the Eternal Starry Summer” by van Gogh	Yes	Yes	No
Hits (H)	1 (2/2)	.5 (1/2)	1 (2/2)
False Alarm (FA)	1 (2/2)	.5 (1/2)	0 (0/2)
Bias/overstatement (H + FA)/2	1	1	0.5
Accuracy (H - FA)	0	0	1



d - sensibility for distinguish between signal and noise
 c - Criterion – threshold for reporting a detected stimulus

We can use as a framework for selecting items

The Vocabulary and Overclaiming Test (VOC-T)

Matthias Ziegler¹, Christoph Kemper², and Beatrice Rammstedt²

Table 2. Correlations of VOC-T measures with demographic characteristics, the Big 5, risk-taking, and knowledge

	Overclaiming	(VOC-T-bias)	Accuracy	(VOC-T-accuracy)
Study 2				
Sex	-.11*	(-.13)	-.06	(-.09)
Age	-.02	(-.02)	.20**	(.30)
Neuroticism	-.13**	(-.15, -.17)	-.02	(-.02, -.02)
Extraversion	.12**	(.13, .15)	-.05	(-.05, -.06)
Openness	.30**	(.34, .38)	.06	(.06, .07)
Agreeableness	-.05	(-.06, -.06)	-.07	(-.07, -.08)
Conscientiousness	.08	(.09, .10)	.04	(.04, .05)
Risk-taking	.19**	(.22)	.02	(.03)
Study 3				
Sex	.02	(.02)	-.15**	(-.21)
Age	.02	(.02)	.19**	(.26)
IST Knowledge	.02	(.02, .03)	.15**	(.21, .23)
IST Knowledge verbal	-.01	(-.01, -.01)	.18**	(.25, .29)
IST Knowledge numerical	.03	(.04, .05)	.09**	(.12, .16)
IST Knowledge figural	.04	(.05, .06)	.10**	(.14, .18)

Notes. IST = Intelligence Structure Test 2000-R. Values in parentheses corrected for unreliability within VOC-T scores and both, within VOC-T scores and criterion. * $p < .05$, ** $p < .01$.

Abstract. The present study examined the construct validity of the Vocabulary and Overclaiming Test (VOC-T-accuracy), using a large sample of German adults. The resulting questionnaire was compared with the VOC-T-bias measure and other measures of self-enhancement. The results showed that the VOC-T-bias measure had similar psychometric properties in two further samples. The VOC-T-accuracy measure had good internal consistency and construct validity. Overclaiming did not contribute to the variance of VOC-T-accuracy scores. The VOC-T-accuracy measure has good construct validity and reliability. In all, the psychometric properties of the VOC-T-accuracy measure are comparable with those of the VOC-T-bias measure.

Keywords: overclaiming, self-enhancement, vocabulary, overclaiming test, bias, accuracy, psychometrics

Data derived from large-scale surveys such as the European Social Survey (ESS) or the Socio-Economic Panel (SOEP), have enriched psychological research (e.g., Headey, 2008). Extraversion and agreeableness are the two Big 5 personality traits that have been most frequently studied. Extraversion and agreeableness have been shown to contribute substantially to life satisfaction and well-being. This in turn led to the development of the VOC-T, which is a population-representative measure of self-enhancement.



Overclaiming. An international investigation using PISA data

John Jerrim, Philip D. Parker & Nikki Shure

To cite this article: John Jerrim, Philip D. Parker & Nikki Shure (14 Sep 2023): Overclaiming. An international investigation using PISA data, *Assessment in Education: Principles, Policy & Practice*, DOI: [10.1080/0969594X.2023.2238248](https://doi.org/10.1080/0969594X.2023.2238248)

To link to this article: <https://doi.org/10.1080/0969594X.2023.2238248>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 14 Sep 2023.

Table 1. The association betw**(a) Gender**

England
Scotland
New Zealand
Canada
Australia
Wales
Ireland
USA
Northern Ireland

(b) Socio-economic status

New Zealand
Scotland
Wales
Australia
USA
England
Canada
Northern Ireland
Ireland

(c) Immigrant group

Northern Ireland
New Zealand
Ireland
England
Scotland
Canada
Australia
Wales
USA

3. Results**Who overclaims?**

Table 1 considers how average scores on the overclaiming scale differ between demographic groups. There is an important difference between genders; boys are much more likely to overclaim than girls. This holds true across all nine countries, with all differences statistically significant (even after a Bonferroni correct is made) and equivalent to an effect size of between 0.2 and 0.3 standard deviations in most countries. Consequently, Table 1 provides strong and consistent evidence that teenage boys are more likely to overclaim than teenage girls.

A similar difference is found with respect to socio-economic status; young people from more advantaged socio-economic backgrounds have higher average overclaiming scores than their less advantaged peers. The magnitude of the difference is again not trivial and varies somewhat across countries. For instance, the difference in

The final difference considered in Table 1 is between immigrant and native-born citizen groups. In most countries, immigrants having significantly higher scores than young people who are native-born. This is particularly pronounced in New Zealand and Northern Ireland, where immigrants score around 0.3–0.4 standard deviations higher on

	Native-born	Immigrants	Gap (effect size)	SE
(a) Gender				
England	-0.14	0.27	.40*	0.17
Scotland	-0.03	0.25	.28**	0.06
New Zealand	-0.23	0.03	.25**	0.07
Canada	0.05	0.27	.23**	0.06
Australia	-0.23	-0.05	.18*	0.09
Wales	0.13	0.26	.14**	0.03
USA	0.01	0.13	.12**	0.03
Northern Ireland	-0.01	0.21	.25	0.13
Ireland	0.21	0.22	.01	0.05
(b) Socio-economic status				
New Zealand	18***			0.03
Scotland	17**			0.05
Wales	16**			0.05
Australia	16**			0.03
USA	16**			0.03
England	16**			0.03
Canada	16**			0.03
Northern Ireland	16*			0.06
Ireland	14**			0.05
(c) Immigrant group				
Northern Ireland				
New Zealand				
Ireland				
England				
Scotland				
Canada				
Australia				
Wales				
USA				

The overclaiming scale has been standardised within each country to mean zero and standard deviation one. The gap refers to the difference between groups in terms of an effect size. SE refers to the standard error of the gap. Northern Ireland excluded from socio-economic status results due to factor scores not able to be calculated. * indicates statistical significance at $p = 0.05$; ** indicates statistical significance at $p = 0.006$ after a Bonferroni correction has been made. Data source: OECD (2014).

Liar! Liar! (When Stakes Are Higher): Understanding How the Overclaiming Tech

Patrick D. Dunlop
Curtin University

Reinout E. de Vries
Vrije Universiteit Amsterdam and Universit

Karina Jorritsma
Curtin University

Overclaiming questionnaire knowledge about a given topic can plague self-report assessments. OCQs in this respect is inconsistent and identify faking. We propose that the content of the OCQ must overclaiming knowledge of propositions through three demonstrate that overclaiming

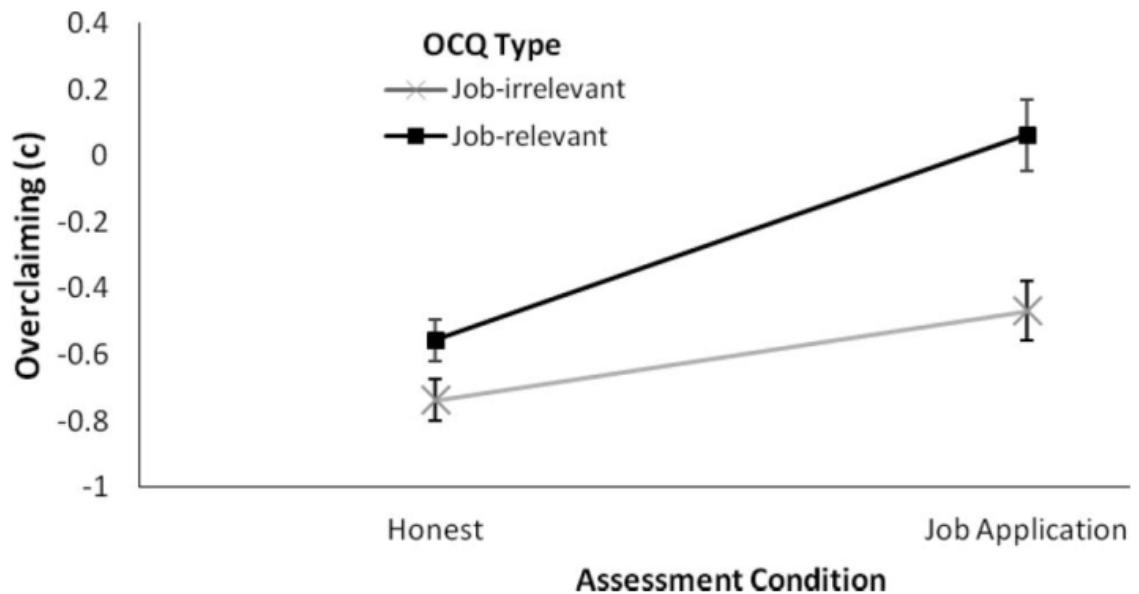


Figure 1. Interaction between assessment condition and OCQ type on the amount of overclaiming observed. Error bars represent 95% confidence intervals. OCQ = overclaiming questionnaires.



The utility of overclaiming questionnaires depends on the fit between test content and application context

Adrian Hoffmann¹ · Birk Diedenhofen¹ · Sascha Müller^{2,3}

Accepted: 25 October 2022
© The Author(s) 2022

Abstract

Overclaiming questionnaires (OCQs) were proposed as a means to counteract social desirability bias by capturing individual differences in participants' self-enhancement tendencies in self-report assessments. Previous studies that evaluated OCQs reported mixed results. However, fit between the content of an OCQ in terms of its items and the context in which the measure is presented has not been tested systematically. In a mock application study ($N=432$), we compared different levels of content-context fit between conditions. Results show that the utility of a general knowledge OCQ varied as a function of its content fit to different application contexts. Expectedly, overclaiming was most pronounced in an application context with optimal content fit to the OCQ, followed by a context with lower fit and an honest control condition without application context. Furthermore, participants in the application conditions were shown to successfully fake on conventional personality scales while incorporating specific requirements of the application context into their faking behavior. Our results thus corroborate previous findings suggesting a high susceptibility of personality scales to deliberate faking. In contrast, when content-context fit is taken into account, OCQs may be a promising method for assessing applicant faking.

Keywords Overclaiming · Faking · Personnel selection · Social desirability · Validity

Introduction

incentive to present themselves favorably, for instance, to increase their chances of being offered a job because they fit

Using IRT

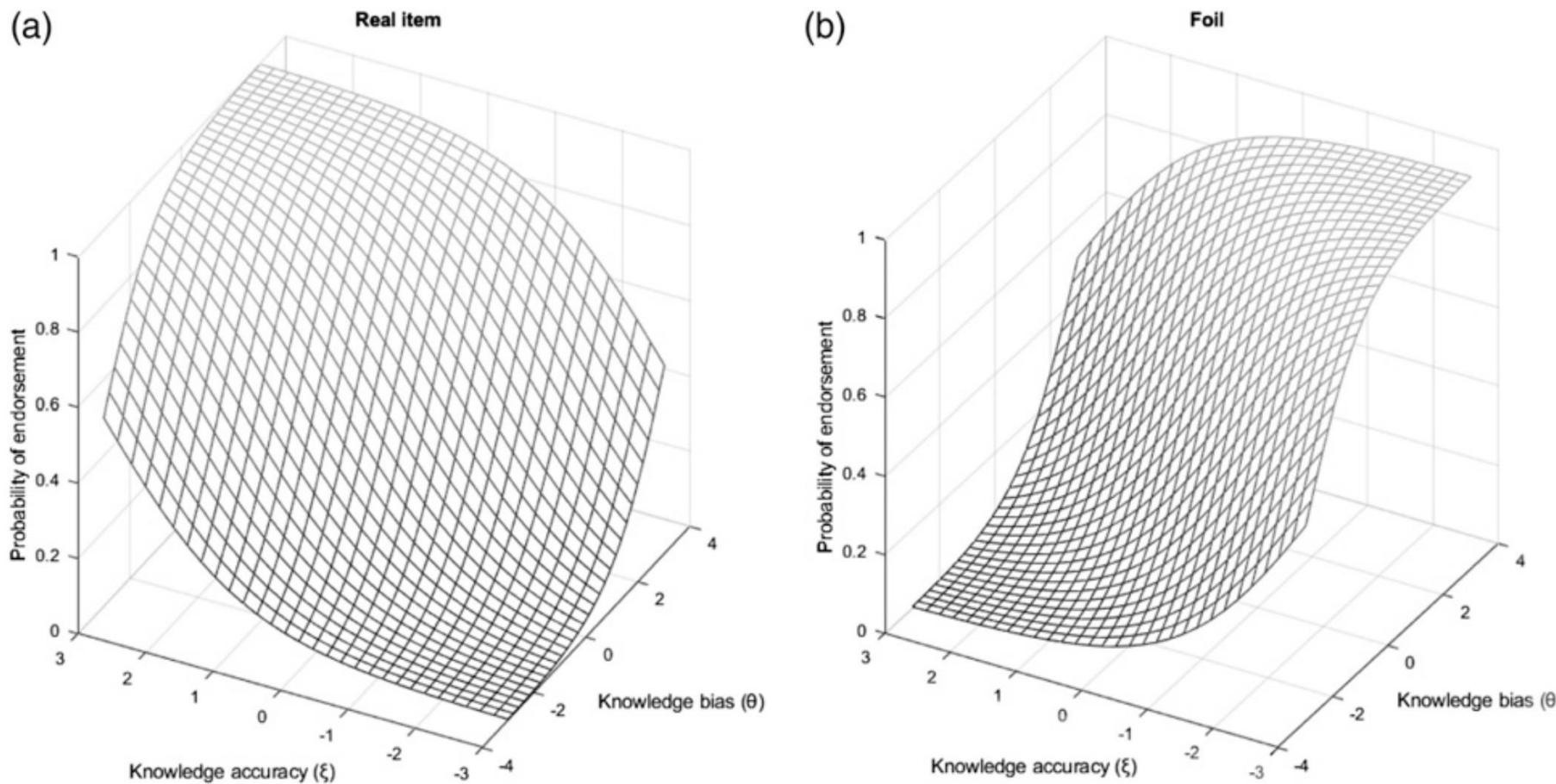
Article

A New Approach to Desirable Responding: Multidimensional Item Response Model of Overclaiming Data

Applied Psychological Measurement
2023, Vol. 47(3) 221–236
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/01466216231151704
journals.sagepub.com/home/apm


g-Lin Shih^{3,2} 

desirable responding in self-report. We extend previous research by applying signal detection formulas to the analysis of knowledge accuracy and personality. Here, we develop a multidimensional item response theory (MIRT) model to analyze overclaiming data. First, a new approach is proposed to yield comparable indices of knowledge accuracy. Two empirical studies based on Chinese idioms—are then conducted to validate the proposed approach for group comparisons and its practical applications. The results are illustrated and discussed.



Leonardo Botinhon de Campos



**OVERCLAIMING, DESEJABILIDADE SOCIAL E NARCISISMO
EM AVALIAÇÕES DE PERSONALIDADE**

Letícia da Silva de Souza



**EU SEI A MENTIRA QUE VOCÊ ESTÁ CONTANDO:
OVERCLAIMING EM CENÁRIOS *HIGH STAKES* E *LOW STAKES***

	Overclaiming			
	Falso Alarme	Hits	Inflation (overclaiming)	Accuracy
BFI2 - EXT	0,07 / 0,24*	0,05 / 0,22*	0,06 / 0,25*	-0,01 / 0,01
BFI2 - AMA	0,06 / 0,10*	0,07 / 0,10	0,07 / 0,11*	0,03 / 0,02
BFI2 - CON	0,07 / 0,14*	0,01 / 0,13*	0,04 / 0,14*	-0,08 / 0,02
BFI2 - NEU	0 / -0,15**	-0,01 / -0,17*	-0,01 / -0,17*	-0,01 / -0,06
BFI2 - OPE	0,27* / 0,20*	0,34* / 0,26*	0,33* / 0,25*	0,17* / 0,14*

Teas Over Desir

Doreen
and Ma

Abstract
Contamin
impact of j
range from
literature
network—
detection
and cognit
not loaded
factors. Al

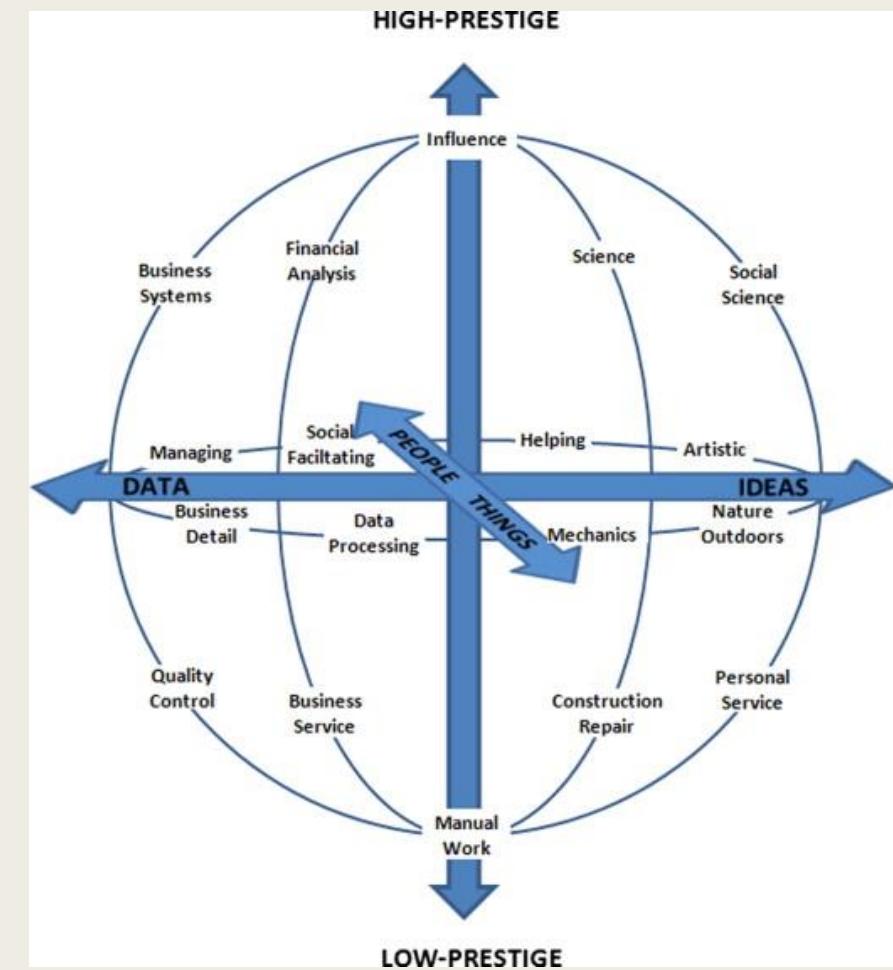
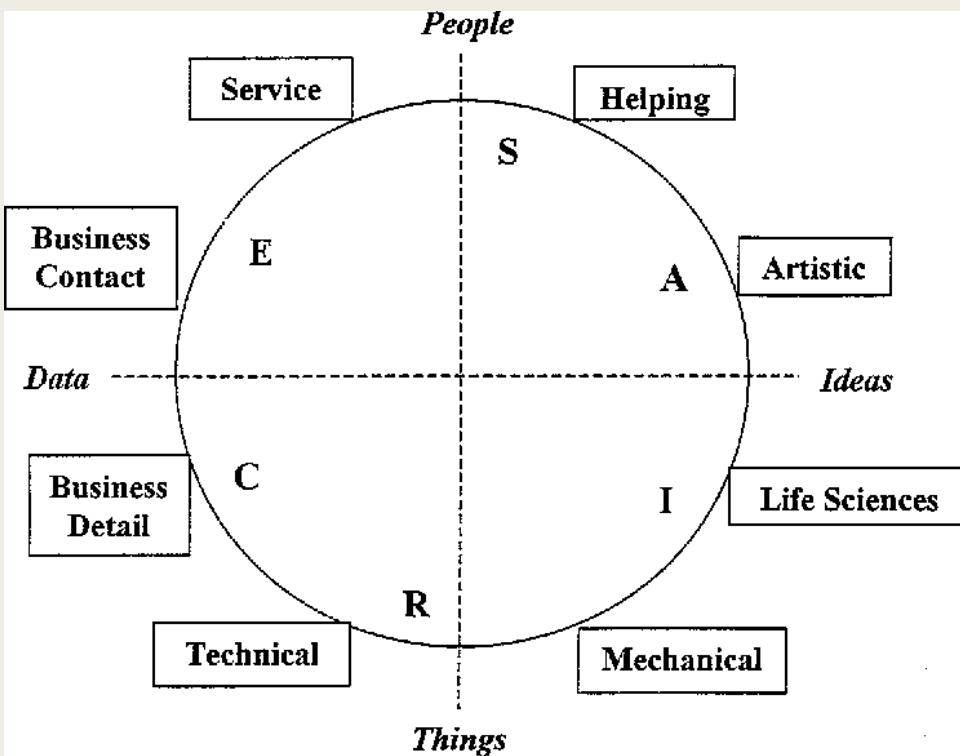
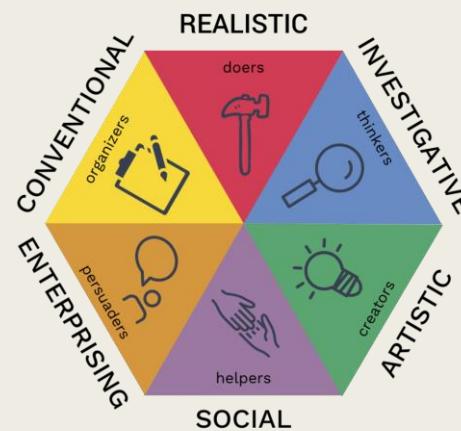
1.	Extraversion (P)													
2.	Agreeableness (P)	.16*	—											
3.	Conscientiousness (P)	.14*	.04	—										
4.	Neuroticism (P)	-.24*	-.17*	-.11*	—									
5.	Openness (P)	.22*	.05	.09*	.03	—								
6.	Grandiose narcissism (P)	-.07	-.05	-.06	.11*	<.01	—							
7.	Edwards SDR scale (SDR)	-.36*	-.25*	-.34*	.68*	-.04	.15*	—						
8.	First-factor Marlowe–Crowne scale (attribution of desirable behavior; SDR)	.16*	.29*	.33*	-.08*	.21*	-.06	-.22*	—					
9.	Second-factor Marlowe–Crowne scale (denial of undesirable behavior; SDR)	-.02	-.29*	-.22*	.35*	-.04	.03	.47*	-.20*	—				
10.	Self-deceptive enhancement (SDR)	-.12*	-.22*	-.14*	.22*	.03	.01	.36*	<.01	.34*	—			
11.	Impression management (SDR)	-.04	-.11*	.06	.10*	-.01	-.01	.11*	.03	.21*	.26*	—		
12.	Lie scale (SDR)	.07	-.22*	-.29*	.22*	-.03	.09*	.32*	-.29*	.52*	.25*	.11*	—	
13.	Confidence bias (OCO)	.07*	-.07	.01	-.11*	.05	-.08*	-.06	.03	.02	-.03	.11*	-.02	—
14.	Crystallized intelligence (I)	-.02	.05	-.02	.04	.02	-.05	.32*	.19*	.21*	.38*	.48*	.43*	.46*
15.	Overclaiming (OCL)	-.11*	-.02	-.06	-.05	.02	-.01	.07*	-.03	-.04	.03	.35*	.15*	.26*
														.38*

is the most distinct potential indicator of positivity bias and independent of known personality measures.

including
nn

Prestige

Bias or real content interest?



Prestige

Bias or real content interest?

How interested you would be in:

Driving a bus 1 2 3 4 5

Driving a bus in a dirty city 1 2 3 4 5

Driving a touristic bus 1 2 3 4 5

Driving a high-speed train 1 2 3 4 5

Pilot an airplane 1 2 3 4 5

Prestige

Bias or real content interest?

How interested you would be in:

Driving a bus 1 2 3 4 5

Driving a bus in a dirty city 1 2 3 4 5

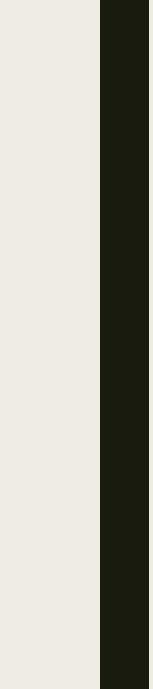
Driving a touristic bus 1 2 3 4 5

Driving a high-speed train 1 2 3 4 5

Pilot an airplane 1 2 3 4 5

Take-home message

- Response bias and set occur in assessments
- Biases are difficult to control, but it is important to measure them
- interaction between
 - item content
 - Descriptive
 - Evaluative
 - personal characteristics
 - testing situation



ITC CONFERENCE

02 · 05 JULY 2024



G R A N A D A

