Methodol Comput Appl Probab (2015) 17:1037–1055
DOI 10.1007/s11009-014-9427-2

SGR Modeling of Correlational Effects in Fake Good Self-report Measures

Luigi Lombardi · Massimiliano Pastore · Massimo Nucci · Andrea Bobbio

Received: 16 January 2014 / Revised: 23 June 2014 /
Accepted: 1 October 2014 / Published online: 28 October 2014

Abstract In many self-report measures (i.e., personality survey items and diagnostic test items) the collected samples often include fake records. A case of particular interest in self-report measures is the presence of caricature effects in participants' responses under faking good motivation conditions. We say that a pattern of fake responses is a caricature pattern if it shows higher structural intercorrelations among faked items relative to the expected intercorrelations under the corresponding uncorrupted responses. In this paper we generalized a recent probabilistic perturbation procedure, called SGR - Sample Generation by Replacements - (Lombardi and Pastore (2012) Multivar Behav Res 47:519–546), to simulate caricature effects in fake good responses. To represent this particular faking behavior we proposed a novel extension of the SGR conditional replacement distribution which is based on a discrete version of the truncated multivariate normal distribution. We also applied the new procedure to real behavioral data on the role of perceived affective self-efficacy in social contexts and on self-report behaviors in reckless driving.

Keywords Sample generation by replacement · Fake-good data · Truncated multivariate normal distribution · Correlational structures · Caricature effect

L. Lombardi (✉)
Department of Psychology and Cognitive Science, University of Trento, corso Bettini, 31, 38068 Rovereto (TN), Italy
e-mail: luigi.lombardi@unitn.it

M. Pastore
Department of Developmental and Social Psychology, University of Padova, via Venezia, 8, 35131 Padova, Italy
e-mail: massimiliano.pastore@unipd.it

M. Nucci
Department of General Psychology, University of Padova, via Venezia, 8, 35131 Padova, Italy
e-mail: massimo.nucci@unipd.it

A. Bobbio
Department of Philosophy, Sociology, Education and Applied Psychology, University of Padova, via Venezia, 8, 35131 Padova, Italy
e-mail: andrea.bobbio@unipd.it

Mathematics Subject Classification (2010) C15 · C34 · C46 · C63

1 Introduction

Intentional response distortion on recruitment and selection surveys and personality questionnaires has been one of the most relevant concern in using self-report measures in socio-behavioral studies. Because many self-report scales of attitudes, beliefs, personality, and pathology include transparent items that can be easily manipulated by respondents, some authors have noted that responded faking may be commonplace (e.g., Griffith and Converse 2011; Levin and Zickar 2002; Rosse et al. 1998). Faking good can be defined as a conscious attempt to present false information to create a favorable impression with the goal of influencing others (e.g., Furnham 1986; McFarland and Ryan 2000; Zickar and Robie 1999). More in general, there is a broad consensus that faking is an intentional response distortion aimed at achieving a personal gain (e.g., MacCann et al. 2011). For example, in personnel selection some job applicants may misrepresent themselves on a personality test hoping to increase the likelihood of being offered a job (e.g., Paulhus 1984; Zickar and Robie1999; Donovan et al. 2013).

Past research has established that respondents who have been instructed to fake good are able to substantially modify their scale scores by providing more extreme response values (e.g., Furnham 1986; Hesketh et al. 2004; McFarland and Ryan 2000; Viswesvaran and Ones 1999). Moreover, faking good can also affect the covariance structure of distorted scales. Generally, evidence suggests that scores under faking-good motivating conditions tend to have smaller variances and lower reliability estimates (Ellingson et al. 2001; Eysenck et al. 1974; Hesketh et al. 2004; Topping and O'Gorman 1997). However, opposite results have also been observed where simple fake good instructions tend to increase the intercorrelations between the manipulated or faked items (Ellingson et al. 1999; Galić et al. 2012; Pauls and Crost 2005; Zickar and Robie 1999; Ziegler and Buehner 2009).

Our study focuses on a particular aspect of faking good behavior in self-report measures that we term the *caricature effect* of faking. In data modeling, a caricature pattern can be understood as a transformed data pattern which exaggerates specific characteristics of an original data pattern. Similarly, a fake data pattern is extreme, relative to its true data pattern as it magnifies some of its relevant features. So, for example, in a personality test a caricature pattern can reveal a stronger association between emotion stability and conscientiousness as compared with the corresponding real correlation between the two dimensions. Similarly, in a post-traumatic stress disorder (PTSD) scenario, a caricature pattern can be identified by more emphasized associations between intensity-level symptoms relative to patterns of individuals who really suffer from PTSD. In general, we distinguish caricature effects from spurious (or inflated) correlations that may be elicited from extreme responding in high-stakes tests (e.g., Cronbach 1946; Landers et al. 2011). Unlike spurious correlations, caricature effects are not necessarily related to extreme response styles (defined as the tendency to prefer the highest responses when confronted with a Likert-type item). By contrast, they are characterized by (fake) self report measures which correspond to moderate shifts in the values of the original responses (Pastore and Lombardi 2014).

In this contribution we adopt a recent modeling approach, called Sample Generation by Replacement (SGR; Lombardi and Pastore 2012), to investigate caricature effects in faking good responses. SGR is a general probabilistic procedure that allows a detailed exploration of what outcomes are produced by particular sets of faking assumptions and provides a kind of *what-if-analysis* of hypothetical faking scenarios. This kind of prospective analysis can be used to quantify uncertainty in inferences based on possible fake data as well as to evaluate the implications of fake data for statistical results. For example, SGR has been successfully applied to evaluate the impact of hypothetical faking good manipulations on therapy-compliance indicators in a sample of liver transplant patients (Lombardi and Pastore 2012). It has also been used to study the sensitivity of reliability indices to fake perturbations in dichotomous and ordered data under the tau-equivalent condition (Pastore and Lombardi 2014) or to test simple inferential hypotheses about faking manipulations (Lombardi and Pastore 2014). Unfortunately, the standard SGR approach is limited to the simulation of conditionally independent fake data which do not allow to represent caricature effects in the covariance structure. To fill this gap, in this contribution, we introduce a new generalization of the SGR approach that accounts for caricature effects in the covariance structure using a direct representation for correlated patterns in the simulated fake data.

The next section provides a little summary about the most relevant statistical approaches dealing with fake data analysis. The third section briefly recapitulates the main aspects of the SGR procedure to simulate fake data. In the fourth section the new model of faking to mimic caricature effects is introduced. The fifth section illustrates our method with two applications to real data sets about the role of perceived affective self-efficacy in social contexts and the effect of some environmental determinants on self-report behaviors in reckless driving, respectively. Finally, the last section presents conclusions and some relevant comments about limitations and potential new applications of the SGR approach.

2 Some Currently Used Approaches for Fake Data

Statistical approaches for dealing with faking in self-report measures are not new and many methods have been used to minimize the impact of possible fake data in sample surveys. For example, ethnographic methods (i.e., nominative techniques and snowball sampling) have been constructed to estimate characteristics of stigmatized behaviors which often result in underreporting or fake data (Tracy and Fox 1981; Miller 1981). Similarly, psychometric methods have been developed to identify and evaluate subjects' responses for feigning (fake-bad, malingering) or defensiveness (fake-good, self-deception, social desirability) using factor analytic approaches (e.g., Ferrando 2005, Ferrando and Anguiano-Carrasco 2011; Fox and Meijer 2008; Holden and Book 2009; Leite and Cooper 2010; McFarland and Ryan 2000; Paulhus 1991; Zickar and Robie 1999; Ziegler and Buehner 2009), factor mixture models (Leite and Cooper 2010), and case-diagnostic procedures (Pek and MacCallum 2011). Another well known method is represented by randomized response (RR; Chaudhuri and Mukerjee 1988; Fox and Tracy 1986; Warner 1965). RR is a general approach that was developed in the statistics community for the purpose of protecting surveyees' privacy and has been used especially in self-administered questionnaires for large scale sample surveys (e.g., Campbell 1987; Cohen 1987; Kolata 1987). Generally, RR is characterized by complex and not always transparent sampling procedures as well as by the need of a large number of cases which is usually necessary to produce estimates with a sufficient level of reliability (Campbell 1987). In particular, RR and its derivatives are often criticized not only because of their exacting demands on the skills of responders in handling the required devices, but also, and mainly, because these techniques ask respondents to provide information that seems useless or even tricky to them (Campbell 1987). In these circumstances the interviewee may feel that s/he is being tricked by the interviewer or eventually s/he may simply doubt about the method itself.

Unlike the previous methods, SGR takes an interpretation perspective which incorporates in a global model all the available information (empirical or hypothetical) about the process of faking and the underlying true model representation. In particular, SGR is not a method for detecting faking at the individual level but a rational approach to evaluate statistical results under potential faking corrupted data. Moreover, SGR has a statistical descriptive nature and does not hinge on a specific psychological theory of faking. It simply tries to capture the phenomenological effect of faking according to an informational, data-oriented perspective based on a data replacement (information replacement) scheme. This makes SGR more related in spirit to other statistical approaches such as, for example, uncertainty and sensitivity analysis (Helton et al. 2006) and prospective power analysis (Cohen 1988) which are characterized by an attempt to directly quantify uncertainty of general statistics computed on the data by means of specific hypothesis.

3 Sample Generation by Replacement

SGR is a probabilistic resampling procedure that can be used to simulate fake discrete or ordinal data with a restricted number of values (Lombardi and Pastore 2012). SGR is characterized by a two-stage sampling procedure based on two distinct generative models: the model defining the process that generates the data prior to any fake perturbation (*data generation process*) and the model representing the faking process to perturb the data (*data replacement process*). In SGR the data generation process is modeled by means of standard Monte Carlo procedures for ordinal data whereas the data replacement process is implemented using ad hoc probabilistic faking models. In sum, the overall generative process is split into two conceptually independent and possibly simpler components (divide and conquer strategy): data generation + data replacement.

In a more formal way, we may think of the original (fake-uncorrupted) data as being represented by an $I \times J$ matrix $\mathbf{D}$, that is to say, I i.i.d. observations (hypothetical participants) each containing J elements (hypothetical participant's responses). We assume that entry d_{ij} of $\mathbf{D}$ ($i = 1, \ldots, I; j = 1, \ldots, J$) takes values on a small ordinal range, $1, 2, \ldots, Q$, (e.g., $Q = 5$ for 5-point Likert items). In particular, let $\mathbf{d}_i$ be the $(1 \times J)$ array of $\mathbf{D}$ denoting the pattern of responses of participant i. The response pattern $\mathbf{d}_i$ is a multidimensional ordinal random variable with probability distribution $p(\mathbf{d}_i|\theta)$, where θ indicates the vector of parameters of the probabilistic model for the data generation process. Therefore, the data matrix $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_I]^T$ is drawn from the joint probability distribution

$$p(\mathbf{D}|\theta) = \prod_{i=1}^{I} p(\mathbf{d}_i|\theta). \tag{1}$$

representing the original data generation process. The main intuition of our replacement approach is to construct a new $I \times J$ ordinal data matrix $\mathbf{F}$, called the *fake data matrix* of $\mathbf{D}$, by manipulating each element d_{ij} in $\mathbf{D}$ according to a replacement probability distribution (data replacement process). Let $\mathbf{f}_i$ be the $(1 \times J)$ array of $\mathbf{F}$ denoting the replaced pattern of fake responses of participant i. The fake response pattern $\mathbf{f}_i$ is a multidimensional ordinal random variable with conditional replacement probability distribution $p(\mathbf{f}_i|\mathbf{d}_i, \theta_F)$.

It is important to note that in the standard SGR framework the replacement distribution $p(\mathbf{f}_i|\mathbf{d}_i, \theta_F)$ is restricted to satisfy the *conditional independence assumption* (see Lombardi and Pastore 2012; Pastore and Lombardi 2014). More precisely, in the replacement distribution each fake response f_{ij} only depends on the corresponding data observation d_{ij} and the model parameter θ_F. Therefore, because the patterns of fake responses are also i.i.d. observations, the simulated data array $(\mathbf{D}, \mathbf{F})$ is drawn from the joint probability distribution

$$p(\mathbf{D}, \mathbf{F}|\theta, \theta_F) = \prod_{i=1}^{I} p(\mathbf{d}_i|\theta) p(\mathbf{f}_i|\mathbf{d}_i, \theta_F) \quad (2)$$

$$= \prod_{i=1}^{I} p(\mathbf{d}_i|\theta) \prod_{j=1}^{J} p(f_{ij}|d_{ij}, \theta_F) \quad (3)$$

By repeatedly sampling data from the two generative models we can simulate the so called *fake data sample* (FDS). We can then study the distribution of some relevant statistics computed on this FDS.

Unfortunately, the conditional independence assumption limits the domain of applicability of the SGR approach. In particular, Eq. 3 does not allow to directly represent modulations in the covariance structure of the the faked responses that are typical of caricature effects in faking good scenarios. To overcome this limitation, in the next section we present a generalization of the SGR modeling that does not hinge on the conditional independence assumption.

4 Representing the SGR Components

4.1 Data Generation Process

In the multivariate latent variable framework there are many possible approaches to modeling ordinal variables according to Eq. 1. In this contribution we focus on the Underlying Variable Approach (UVA; Muthén 1984; Lee et al. 1990; Jöreskog 1990). Following the UVA framework we assume that there exists a continuous data matrix $\mathbf{D}^*$ underlying the original ordinal data matrix $\mathbf{D}$. Let $\mathbf{d}_i^*$ be the $(1 \times J)$ array of $\mathbf{D}^*$ denoting the pattern of underlying continuous responses of participant i. Without loss of generality, it is convenient to let $\mathbf{d}_i^*$ have the multivariate normal distribution with density function $\phi(\mathbf{0}, \mathbf{R})$ where $\mathbf{0}$ and $\mathbf{R}$ denote the $(1 \times J)$ array of zeros representing the location vector of ϕ and the $(J \times J)$ correlation matrix $\mathbf{R}$ of the multivariate normal distribution, respectively. The connection between the ordinal variable d_{ij} and the underlying variable d_{ij}^* in $\mathbf{D}^*$ is given by the following rule:

$$d_{ij} = h \qquad \text{iff} \qquad \alpha_{h-1}^j < d_{ij}^* < \alpha_h^j; \quad h = 1, \ldots, Q; i = 1, \ldots, I; j = 1, \ldots, J,$$

where

$$-\infty = \alpha_0 < \ldots < \alpha_{h-1}^j < \alpha_h^j < \ldots < \alpha_Q = +\infty,$$

are threshold parameters for the continuous data d_{ij}^*. Note that, for each variable d_{ij} with Q categories, there are $Q - 1$ strictly increasing threshold parameters. Therefore, the probability distribution for the multidimensional ordinal random variable $\mathbf{d}_i = (h_1, \ldots, h_J)$ is given by

$$p(\mathbf{d}_i|\theta_M) = \int_{\alpha_{h_1-1}^1}^{\alpha_{h_1}^1} \cdots \int_{\alpha_{h_J-1}^J}^{\alpha_{h_J}^J} \phi(\mathbf{z}_i|\mathbf{0}, \mathbf{R}) d\mathbf{z}_i \quad (4)$$

with $\theta_M = (\alpha, \mathbf{R})$ and $\mathbf{z}_i = (z_{i1}, \ldots, z_{iJ})$ being the parameter vector of the original data generation model and the values for the continuous variables $\mathbf{d}_i^*$, respectively. A simple way to simulate data according to Eq. 4 can be obtained by first generating the continuous data $\mathbf{D}^*$ (according to the target model) and subsequently transform it into its discrete counterpart $\mathbf{D}$ by using appropriate fixed threshold values α.

4.2 Data Replacement Process

The SGR approach offers an elegant way to simulate faking good scenarios. Notice that, the faking good (as well as the faking bad) scenario always entails a conditional replacement model in which the conditioning is a function of response polarity. We assume a perturbation context in which responses are exclusively subject to positive feigning:

$$f_{ij} \geq d_{ij}; \qquad i = 1, \ldots, I; \; j = 1, \ldots, J.$$

More precisely, a pure faking good scenario requires that in the replacement distribution the following condition holds:

$$p(\mathbf{f}_i|\mathbf{d}_i, \theta_F) = 0, \qquad \exists j : f_{ij} < d_{ij}.$$

In other words, this model does not allow to substitute the original observed value with lower ones. However, to represent caricature effects in faking good scenarios we must also capture the magnified correlations among the faked items, that is to say, we need to directly control the correlational patterns in the conditional replacement distribution. To this aim we introduce a novel representation of the faking model which accounts for modulations in the covariance structure of the fake responses.

4.2.1 The Truncated Multivariate Replacement Distribution

As a kernel for the conditional replacement distribution we consider the truncated multivariate normal distribution $TN(\mu, \Sigma, \mathbf{a}, \mathbf{b})$ (e.g., Horrace 2005). This distribution can be expressed as

$$f(\mathbf{x}|\mu, \Sigma, \mathbf{a}, \mathbf{b}) = \frac{\exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)\right\}}{\int_{\mathbf{a}}^{\mathbf{b}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)\right\} d\mathbf{x}} \tag{5}$$

for $\mathbf{a} \leq \mathbf{x} \leq \mathbf{b}$ and 0 otherwise. The $(1 \times J)$ vectors $\mathbf{a}$ and $\mathbf{b}$ are the lower and upper truncation points $(a_j < b_j; \; j = 1, \ldots, J)$ for the multivariate normal distribution with J dimensions. Finally, μ and Σ are the location parameter vector and the covariance matrix of the (not truncated) multivariate normal distribution.

Now, let $\mathbf{f}_i = (k_1, \ldots, k_J)$ and $\mathbf{d}_i = (h_1, \ldots, h_J)$ be the replaced values and the original values for the i^{th} simulated observation, respectively. According to the UVA paradigm we can set

$$p(\mathbf{f}_i|\mathbf{d}_i, \theta_F) = \int_{\beta^1_{k_1-1}}^{\beta^1_{k_1}} \cdots \int_{\beta^J_{k_J-1}}^{\beta^J_{k_J}} f(\mathbf{x}|\mathbf{0}, \Sigma, \mathbf{a}^i, \mathbf{b}^i) d\mathbf{x}, \tag{6}$$

for all items j such that $1 \leq h_j \leq k_j \leq Q$. By contrast, the replacement distribution simply takes value 0 whenever it exists at least one item j such that $k_j < h_j$. In the replacement distribution we adopted a truncated multivariate standard distribution with location parameter vector $\mu = \mathbf{0}$ and correlation matrix Σ. Moreover, in the replacement distribution the pair $(\beta^j_{k_j-1}, \beta^j_{k_j})$ denotes the thresholds corresponding to the discrete value k_j for item j. Finally, the bounds $\mathbf{a}^i = (a^i_1, \ldots, a^i_J)$ and $\mathbf{b}^i = (b^i_1, \ldots, b^i_J)$ are set to

$$a^i_j = \beta_{h^j_{j-1}} \qquad b^i_j = +\infty, \qquad j = 1, \ldots, J$$

where $(\beta^j_{h_j-1}, \beta^j_{h_j})$ is the pair of thresholds corresponding to the value h_j for the original response d_{ij} in $\mathbf{d}_i$. In sum, we can describe the information characterizing the faking model by means of the parameter array

$$\theta_F = (\beta, \Sigma).$$

with β being the $J \times (Q-1)$ threshold matrix for the replacement distribution.

4.2.2 *Standard Replacement Distribution Assumptions*

We recall that in the standard SGR simulation procedure (Lombardi and Pastore 2012; Pastore and Lombardi 2014) the fake perturbations are restricted to satisfy the conditional independence assumption. More precisely, under the faking good condition ($\mathbf{d}_i \leq \mathbf{f}_i$) the conditional replacement distribution reduces to the following multiplicative conditional distribution:

$$p(\mathbf{f}_i|\mathbf{d}_i, \theta_F) = \int_{\beta^1_{k_1-1}}^{\beta^1_{k_1}} \cdots \int_{\beta^J_{k_J-1}}^{\beta^J_{k_J}} f(\mathbf{x}|\mathbf{0}, \mathbf{I}, \mathbf{a}^i, \mathbf{b}^i) d\mathbf{x}, \tag{7}$$

$$= \prod_{j=1}^{J} \int_{\beta^j_{k_j-1}}^{\beta^j_{k_j}} f(x|0, 1, a^i, b^i) dx, \tag{8}$$

$$= \prod_{j=1}^{J} p(f_{ij}|d_{ij}, \theta_F) \tag{9}$$

with $\mathbf{I}$ and $f(x|0, 1, a^i, b^i)$ being the $J \times J$ identity matrix and the one-dimensional truncated standard distribution, respectively. Unfortunately, this restriction clearly limits the range of empirical faking processes that can be mimicked by the SGR simulation procedure. In particular, because the replacement distribution acts as a perturbation process for the original data, the resulting fake data sets will generally yield covariance structures that are (on average) weaker than the ones observed for the original uncorrupted data, thus showing a sort of residual correlation effect (or anticaricature effect). By contrast, the general form of the new replacement distribution described in Eq. 6 does allow to represent different levels of correlational patterns in the simulated fake data. In particular, the correlation matrix $\mathbf{R}_f$ of the fake data matrix $\mathbf{F}$ can be modulated by the covariance matrix Σ in the replacement model. In the next section we will show by means two empirical applications how the resulting correlation matrix $\mathbf{R}_f$ can be affected from the interaction between different modulations of faking (represented by different configurations of threshold values α) and different structures for the covariance matrix Σ in the replacement model.

5 Applicative Examples

The new replacement distribution is illustrated using two empirical applications. The first application deals with data from a questionnaire about the role of perceived affective self-efficacy in personality evaluation (Bandura et al. 2003). The second application considers data about self-report behaviors in reckless driving (Taubman-Ben-Ari et al. 2004). These two examples show alternative ways of performing fake-data analysis using the SGR approach.

5.1 Perceived Affective Self-efficacy

The current section is divided into two subsections: the first introduces the empirical data sets and the generative model for the data generation process; the second discusses how we can use SGR to compare the performances of four distinct faking models with respect to the observed data.

5.1.1 Original Data Set and Generative Model

Participants were 498 undergraduate students (404 females) at the University of Padua (Italy). Ages ranged from 18 to 56, with a mean of 20.96 years and a standard deviation of 3.86. The 498 participants were randomly assigned to two groups. The first group ($n_1 = 247$) received a control set of instructions, whereas the second group ($n_2 = 251$) received faking motivating instructions. The two groups resulted also matched for gender ($p = 0.26$), age ($p = 0.80$), and education ($p = 0.38$).

In particular, in the faking group the participants were induced to believe that a renowned Italian selection and recruitment company was interested in hiring some candidates for a very appealing and challenging job position. Key competences in order to succeed were told to be both the ability to manage affective and empathic social relationships (e.g., prosocial behavior), and a positive attitude towards teamwork. Answers to the items of a specific questionnaire would then be considered by the company as the first criteria to meet in order to have access to the subsequent steps of the selection process. By contrast, in the control group participants were instructed to join a relevant scientific project, whose aim was to translate and adapt to the Italian context a new and innovative psychological instrument. However, unlike the faking group, the controls had to complete the same questionnaire with the request to answer all items as honestly and accurately as possible, so that both robustness, reliability and validity of the instrument would not be threatened, and results could be of benefit both for scholars and practitioners.

Data consisted of the participants' responses to four of the 12 items of the Perceived Empathic Self-Efficacy Scale, Adult version (AEP/A; Caprara 2001) scored on a 5-point scale where 1 denotes that she/he "Cannot do at all" the behavior described in the item, while 5 denotes that she/he "Certain can do" it. AEP/A was designed to assess individuals' perceived capability to recognize emotions, feelings, preferences and needs of other people.

Table 1 Items selected for the study. Items were introduced by the following statement "How well can you"

Item	Description
aep1	When you meet new friends, find out quickly the things they like and those they do not like?
aep4	Recognize if a person is seriously annoyed with you?
aep7	Understand the state of mind of others when you are very involved in a discussion?
aep8	Understand when a friend needs your help, even if he/she doesn't overtly ask for it?

Table 2 Frequency tables for the ordinal responses as a function of item number and type of group

	Value					
Item	1	2	3	4	5	Mean response value
aep1(N)	2	21	149	72	3	3.215
aep1(F)	0	8	121	111	11	3.498
aep4(N)	0	23	68	123	33	3.672
aep4(F)	0	9	56	135	50	3.904
aep7(N)	4	62	86	82	13	3.154
aep7(F)	3	32	91	96	29	3.462
aep8(N)	0	6	65	135	41	3.854
aep8(F)	0	3	51	131	66	4.036

N and F denote the control group and the faking group, respectively

A description of the four selected items is reported in Table 1. The four items were chosen in order to guarantee representativeness of the complete item pool, a good factorial structure (NNFI = .985, RMSEA = .022) as well as a clear difference between the two groups in response frequencies (Table 2) and correlation patterns (Table 3).

The resulting responses were collected into two empirical data matrices $\mathbf{D}_e$ (247×4) and $\mathbf{F}_e$ (251×4) for the control group and the faking group, respectively. Having developed these two scenarios, we argued that in the fake condition, as frequently happens in personnel selection situations (e.g., Donovan et al. 2013), participants would be motivated to enhance or overestimate their scores on the Perceived Empathic Self-Efficacy scale, in order to increase the likelihood of being appreciated by the recruitment company and, consequently, of being offered the job. As expected the second group showed a sort of fake-good effect for the observed responses (see Table 2). More precisely, the participants in the faking group seemed to deliberately manipulate their responses using larger values of the scale to create better impressions. Similarly, the responses in the same group revealed also stronger associations among the four items as compared with the observed correlations in the control group (see Table 3).

The main idea of our SGR analysis was to use the data matrix $\mathbf{D}_e$ (control group) to set the values of the parameters $\theta_M = (\alpha, \mathbf{R})$ in the original generative model. In this context, $\mathbf{D}_e$ would represent a sort of (empirically based) a priori knowledge about fake uncorrupted responses. In particular, the parameters of the generative model were derived according to

Table 3 Polychoric correlations among the four items for the control and faking groups

	aep1	aep4	aep7	aep8
aep1		0.16	0.30	0.33
aep4	0.04		0.20	0.25
aep7	0.11	0.07		0.21
aep8	0.11	0.18	0.14	

Note: Values to the left of the diagonal are correlations for the control group, and values to the right of the diagonal are correlations for the faking group

Table 4 Estimated maximum likelihood thresholds for the generative model

	h			
	1	2	3	4
aep1	-2.40	-1.32	0.51	2.25
aep4	-2.27	-1.32	-0.34	1.11
aep7	-2.14	-0.62	0.29	1.62
aep8	-2.27	-1.97	-0.56	0.97

the following procedure. First, the thresholds α were set equal to the maximum likelihood estimates

$$\widehat{\alpha}_h^j = \Phi^{-1}\left(\sum_{q=1}^{h} \frac{n_q^j}{N_D}\right), \qquad h = 1, \ldots, Q-1;\ j = 1, \ldots, 4$$

with N_D, n_q^j, and Φ^{-1} being the total number of participants in control group, the total number of responses for item j falling in the ordinal category q, and the inverse of the cumulative density function (CDF) for the standardized distribution $N(0, 1)$, respectively (see Table 4).

Next, the correlation matrix $\mathbf{R}$ in the original generative model was set equal to the polychoric correlation matrix computed from $\mathbf{D}_e$ (for more details about the estimation procedures the reader may refer to, for example, Yang-Wallentin et al. 2010). Therefore, on the basis of the parameter values of the generative model we were able to simulate samples according to Eq. 4.

5.1.2 Comparing Faking Models

We performed an SGR analysis on the basis of different hypothetical scenarios of faking. By using a simulation design, we evaluated the mimicking ability of four different faking models with respect to the empirical fake data set $\mathbf{F}_e$ (faking group condition). To that end, we defined four perturbation models derived by the combination of two factors with two levels each. The first factor in the simulation design defined two structures for the covariance matrix Σ in the truncated replacement distribution: a) an identity matrix representing the standard SGR independence model b) a correlation matrix reflecting the patterns of

Table 5 Correlation matrix Σ in the truncated replacement distribution

	aep1	aep4	aep7	aep8
aep1		0.41	0.55	0.53
aep4	0		0.50	0.45
aep7	0	0		0.41
aep8	0	0	0	

Values to the left of the diagonal are correlations for the independence model, and values to the right of the diagonal are correlations for the correlational model

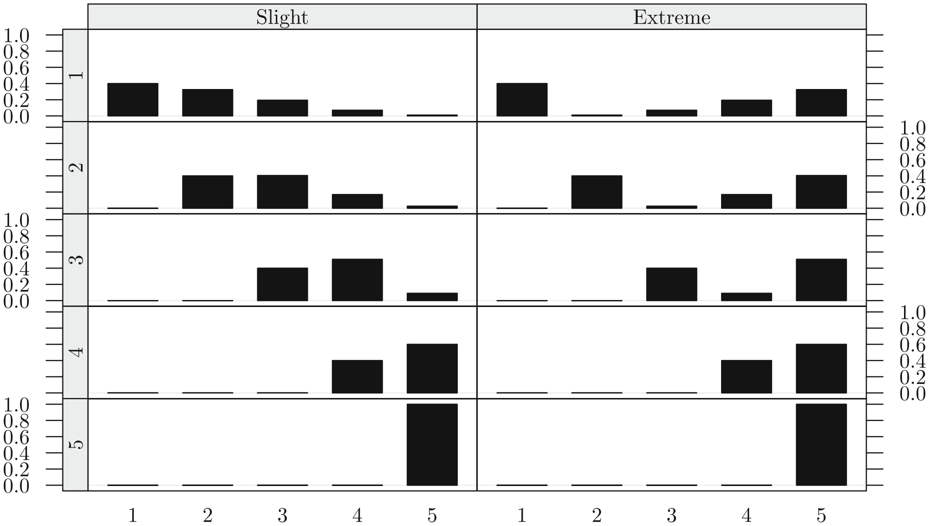

Fig. 1 Two models of conditional replacement distributions for a 5-point discrete r.v. Each column in the graphical representation corresponds to a different conditional replacement distribution with one of the two different assignments for the shape parameters ($\gamma = 1.5, \delta = 4$ and $\gamma = 4, \delta = 1.5$). Each row in the graphical representation corresponds to a different original 5-point discrete value. Note that in the replacement distributions, the probability of a replaced value that is lower than the original discrete value is always 0 (fake good condition)

associations among the items in the faking group. In particular, in this latter condition, the model correlation matrix Σ was obtained by transforming the polychoric correlation matrix computed from $\mathbf{F}_e$ to correct for reduced covariances among the simulated values in the truncated distribution.[1] This resulted in a model correlation matrix Σ with cells having larger values than those in the observed polychoric correlation matrix of $\mathbf{F}_e$ (see Table 5).

The second factor in the simulation design defined two different theoretical response styles for faking: a) *slight faking* b) *extreme faking* (Zickar and Robie 1999). Slight faking describes a response style where the observed self report measure corresponds to a moderate positive shift in the value of the original response. In particular, in this representation the chance to replace an original value h with another greater value k decreases as a function of the distance between k and h (Fig. 1, first column). By contrast, extreme faking describes a response style where the observed self report measure corresponds to an exaggerated positive shift in the value of the original response. More specifically, unlike the slight configuration, in the extreme response style the chance to replace an original value h with another greater value k increases as a function of the distance between k and h (Fig. 1, second column). In order to set the thresholds for the two faking style conditions, we

[1] Σ is the covariance matrix of the original (not truncated) multivariate normal distribution. In particular, it can be seen that truncation can significantly reduce the variance and change the covariance between variables. Therefore, if we wish to simulate correlated fake patterns with associations that are of the same magnitude of the empirical covariance matrix we need to choose a particular Σ which boosts the final simulated correlations.

Table 6 Thresholds corresponding to the two theoretical models of faking

	k			
	1	2	3	4
Slight	-0.14	0.73	1.51	2.41
Extreme	-2.41	-1.51	-0.73	0.14

adopted the generalized beta distribution for discrete variables, DG, originally introduced by Pastore and Lombardi (2014):

$$\beta_l^j = \Phi^{-1}\left(\sum_{q=1}^{l} DG(q; q+1, Q, \gamma, \delta)\right), \qquad l = 1, \ldots, Q-1; j = 1, \ldots, 4.$$

In the DG distribution the values $q + 1$ and $Q(= 5)$ represent the lower and upper bounds of the function DG, whereas γ and δ denote the shaping parameters for the distribution. In this characterization, slight faking and extreme faking are represented by different values in the shaping parameters: ($\gamma = 1.5, \delta = 4$) for the slight faking model and ($\gamma = 4, \delta = 1.5$) for the extreme faking model (Fig. 1). Here the main assumption is that the threshold values are considered invariant across the four items.[2] However, the type of thresholds can change according to the specific response style considered in the model (slight faking against extreme faking, Table 6).

5.1.3 Data Simulation and Results

To test the four faking models we first simulated 2000 original data matrices $\mathbf{D}$ (with size 251×4) using the generative model defined in the previous section. Next, for each simulated data $\mathbf{D}$ the two factors were systematically varied in a complete two-factorial design to generate new fake data sets and test the faking models against the empirical data $\mathbf{F}_e$.

The crucial question now becomes: if the data contained fake observations, would a model based on the caricature effect assumption be able to correctly reconstruct the empirical relations in the observed data matrix $\mathbf{F}_e$? To reach this objective, we studied the difference between the empirical marginal means for the four items in the faking conditions ($\mathbf{F}_e$) and the reconstructed marginal means derived from the simulated data under the four faking models. Moreover, we also evaluated the difference between the empirical polychoric correlation matrix computed on $\mathbf{F}_e$ and the reconstructed correlation matrices derived from the simulation study conditions. We used the ARB index (Average Relative Bias) to evaluate the four faking models:

$$ARB = 100(1/B)\sum_{b=1}^{B}(1/V)\sum_{v=1}^{V}\left(\frac{\hat{\theta}_{bv} - \theta_v}{\theta_v}\right)$$

with $\hat{\theta}_{bv}$ and θ_v being the v-element of the reconstructed statistic (either reconstructed marginal means or polychoric correlations) in the b-sample replicate ($b = 1, 2, \ldots, B$), and

[2]This reduces the complexity of the parameter array β from 4×4 to 1×4.

the υ-element of the observed statistic (either empirical marginal means or polychoric correlations), respectively. A large absolute value of ARB indicates a large discrepancy between the empirical and the reconstructed statistics. Because of the sufficiently large number of replicates ($B = 2000$) in the simulated samples, we were confident to achieve reasonable estimation stability even in the tail regions of the ARB index.

The results of the SGR analysis are shown in Figs. 2 and 3. Figure 2 represents the simulated marginal means of the fake-good data as a function of the two simulation study factors. The results showed that the slight faking model yielded a better performance (ARB = 0.85) as compared with the extreme faking model (ARB = 14.47). Figure 3 shows the simulated correlations of the fake-good data as a function of the two factors. The results showed that the truncated replacement distribution with a correlational structure provided a better performance (ARB = -9.26) as compared with the independence model (ARB = -85.67). In sum, taken together, the two results confirm that a slight faking model with correlated patterns is more consistent with the empirical data $\mathbf{F}_e$. Therefore, according to our definition of caricature effect in faking contexts, we can conclude that the observed data in the faking group condition were more consistent with moderate shifts in the values of the uncorrupted true responses.

5.2 Self-report Behaviors in Reckless Driving

The former application compared responses from participants who were given different instructions for self-representation on a personality questionnaire in a laboratory-type situation (e.g., honest motivating condition vs faking motivating condition). However, laboratory studies comparing situations with different types of instructions for self-representation may suffer from the lack of ecological validity and provide only a limited view of the faking process. There is some evidence that experimental manipulations of faking do not induce

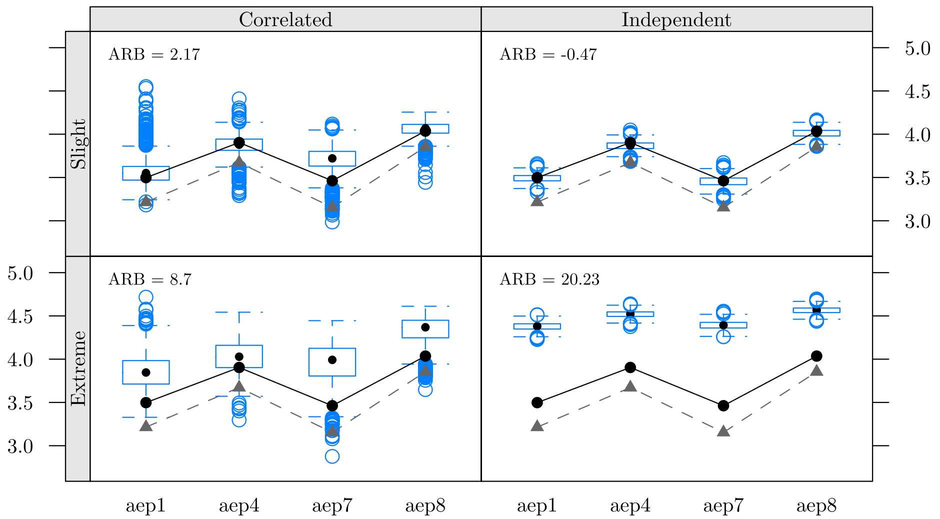

Fig. 2 Boxplots for the simulated marginal means of the fake-good data for the four models. The solid line denotes the observed pattern for the marginal means in $\mathbf{F}_e$. The dashed line indicates the observed pattern for the marginal means in $\mathbf{D}_e$. aep1, aep4, aep7, and aep8 denote the four selected items of the AEP/A scale. The data represented in each boxplot were derived from 2000 fake data samples

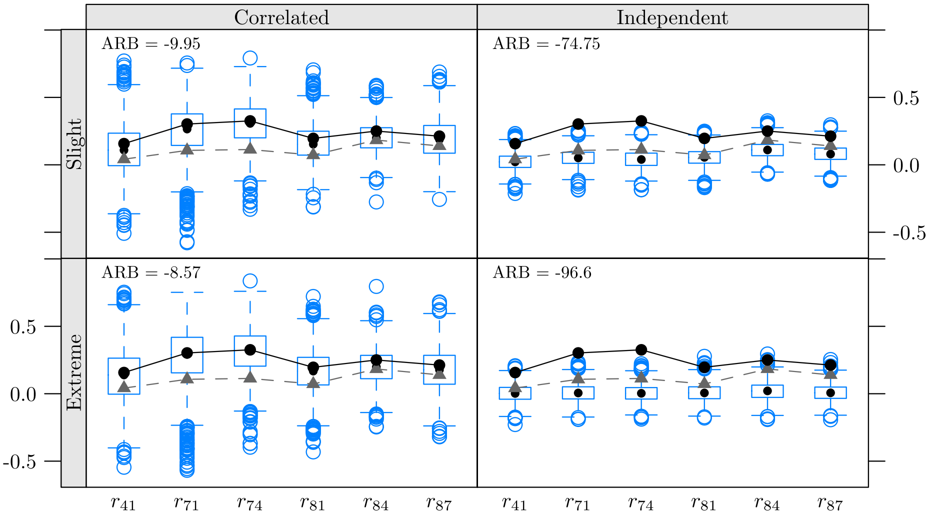

Fig. 3 Boxplots for the simulated correlations of the fake-good data for the four models. The solid line denotes the observed correlational pattern in $\mathbf{F}_e$. The dashed line indicates the observed correlational pattern in $\mathbf{D}_e$. The label $r_{jj'}$ denotes the correlation between item j and item j' of the AEP/A scale. The data shown in each boxplot were derived from 2000 fake data samples

homogeneous patterns of faking (e.g., Zickar et al. 2004). In particular, we are not sure what set of instructions describing hypothetical conditions tell us about faking in real situations (e.g., Galić et al. 2012). So, for example, individuals' profiles under faking motivating instructions may not match those of actual applicants in personnel selection. For this reason, in this second application we illustrated how an analyst can test inferential hypothesis about observed statistical results on data collected in real sensitive contexts. To this aim we applied the SGR procedure to self-report driving experiences about reckless driving in a group of young males.

5.2.1 Original Data Set

A four-item questionnaire was adapted from a previous reckless driving scale (Taubman-Ben-Ari et al. 2004) and administered to a group of 76 young male drivers from the Trentino region (North-East Italy). The only criteria for inclusion in the study were possession of a driving license and at least six months of driving experience. Table 7 reports the item descriptions. Participants were asked to read each item carefully and report how often they used to drive according to the described way when they had friends in the car as passengers.

Table 7 Items selected for the study. Items were introduced by the following statement "How often you used to drive according to the described way"

Item	Description
item 3	Driving at a higher speed than allowed.
item 7	Overtaking another vehicle on a continuous white line (no pass zone)
item 8	Not keeping the right distance from a vehicle in front of me
item 14	Turning high speed

Table 8 Frequency tables for the ordinal responses as a function of item number

Item	value 1	2	3	4	5	Mean response value
item 3	1	2	11	31	31	4.171
item 7	5	16	22	21	12	3.250
item 8	16	20	26	8	6	2.579
item 14	2	20	15	38	1	3.211

Data were collected using a Likert scale with 5 anchor points, ranging from 1 (*never*) to 5 (*very often*). Table 8 shows the response frequencies. By a quick inspection of the counts shown in Table 8 we can easily recognize that a relevant portion of the participants answered using high rating scale values. Some authors (e.g., Taubman-Ben-Ari et al. 2004) claim that reckless driving is perceived as related to both personal and environmental factors. In particular, the driving literature reveals that motivations like competitiveness or sense of power may influence on the way one drives recklessly, especially when peers encourage risk behaviors (e.g., Horvath et al. 1993). In particular, in comparison to more experienced drivers, younger drivers were found to be highly motivated to comply with the perceived wishes of their friends and eventually imitate risky driving (Parker et al. 1992). We hypothesized that young male drivers might tend to overemphasize self-report behaviors of driving transgressions because negatively influenced by risky-driving atmosphere influenced by the peers (Taubman-Ben-Ari et al. 2004).

5.2.2 *Hypothesis Testing*

The four items showed a good internal consistency (observed Cronbach's alpha $\text{Alpha}_C = 0.77$). Because of the relative high frequency of risky driving behaviors (see Table 8), we suspected that the raters might have artificially boosted their responses in order to comply with the perceived wishes of their friends. To test this hypothesis we performed a new SGR analysis on the observed data set by assuming a) a generative model implementing a factorial model reproducing the same observed internal consistency value but with thresholds α_h^j representing less extreme rating responses (compared with the observed ones)[3] b) an independent replacement model with slight faking configuration. We can easily reformulate this setting using a Fisher significance testing (Lehmann 1993; Lombardi and Pastore 2014). More precisely, we can construct the composite hypothesis:

$$H_I : \begin{cases} \textit{Generative model}: \\ \quad 1. \quad \text{Alpha}_C = 0.77, \ \text{(true alpha)} \\ \quad 2. \quad \alpha_1^j = -0.92, \quad \alpha_2^j = 0.16, \quad \alpha_3^j = 1.14, \quad \alpha_4^j = 2.17, \ j = 1, \ldots, 4 \ \text{(thresholds)} \\ \textit{Replacement model}: \\ \quad 1. \quad \gamma^j = 1.5, \ \delta^j = 4, \quad j = 1, \ldots, 4, \ \text{(slight faking)} \\ \quad 2. \quad \Sigma = \mathbf{I} \ \text{(independent model)} \end{cases}$$

[3]To simulate the threshold values for the generative model we used the inverse of the binomial cumulative distribution function with $n = 4$ and $p = 0.35$ (for further details about the inverse strategy see Jöreskog and Sörbom (1996)).

and examine whether or not the observed Cronbach's alpha value is consistent with H_I. In our example, the test procedure (Cronbach's alpha) was replicated 1000 times under the condition of the hypothesis. Next, an approximate p-value was computed as the proportion of the simulated test values which were larger than the observed Cronbach's alpha (0.77). Figure 4 (left panel) shows the distribution of the simulated Cronbach's alpha under the H_I hypothesis. The observed statistic (0.77) seemed not consistent with H_I (approximate p-value $<$ .001). In substantive terms, the observed reliability cannot be explained by an independent generative model with slight faking good manipulations to mimic the responses of the young male drivers.

It might be possible that the SGR modeling failed to represent the observed internal consistency value because of the independency assumption in the replacement model. Therefore, we rerun the SGR analysis this time including a replacement model under a correlational structure (caricature effect model). In particular, the new composite hypothesis was reformulated as follows:

$$H_C : \begin{cases} \textit{Generative model}: \\ \quad 1. \quad \text{Alpha}_C = 0.77, \ \text{(true alpha)} \\ \quad 2. \quad \alpha_1^j = -0.92, \quad \alpha_2^j = 0.16, \quad \alpha_3^j = 1.14, \quad \alpha_4^j = 2.17, \ j = 1, \ldots, 4 \ \text{(thresholds)} \\ \textit{Replacement model}: \\ \quad 1. \quad \gamma^j = 1.5, \ \delta^j = 4, \quad j = 1, \ldots, 4, \ \text{(slight faking)} \\ \quad 2. \quad \Sigma = \mathbf{R} \quad \text{with} \quad r_{jj'} = 0.4, \quad j \neq j' \ \text{(caricature model)} \end{cases}$$

Figure 4 (right panel) shows the distribution of the simulated Cronbach's alpha under the new hypothesis. This time the observed statistic (0.77) seemed more consistent with the hypothetical model. In sum, the observed reliability index is more consistent with a moderate caricature model ($r = 0.4$) mimicking slight faking good manipulations in the rating responses.

6 Limitations and Directions for Future Study

As with other Monte Carlo studies, our investigation involves simplifying decisions that result in lower external validity such as, for example, the assumption that the threshold

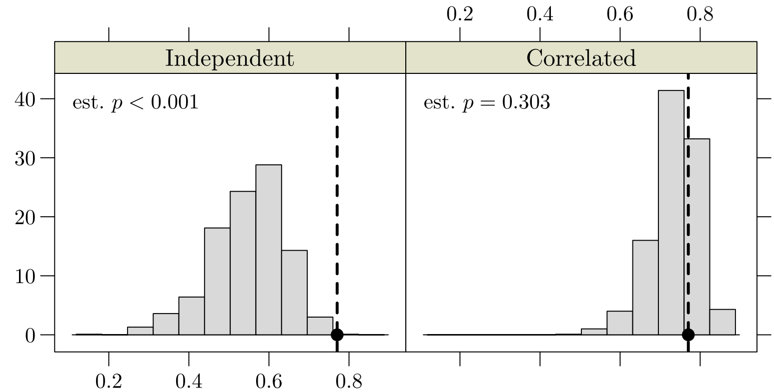

Fig. 4 Reproduced distribution for the test statistic Alpha_C under H_I (left panel) and H_C (right panel). Dashed lines represent the original sample value of Alpha_C (0.77)

values in the replacement distribution are considered invariant across the items. Unfortunately, this restriction clearly limits the range of empirical faking processes that can be mimicked by the current SGR simulation procedure. Moreover, faking is a complex phenomenon which is certainly influenced by the fakers' personality as well as by its interaction with the specific situation. A natural extension of the SGR approach would consider also differential aspects for the responders. More formally, we could model different values for the faking parameters as a function of the respondents' characteristics (and also of the specific items considered). So for example, in an extended version of the SGR approach the parameter vector θ_F would be replaced by specific parameter vectors θ_F^{ij} which depend on specific individuals i and items j. In this way, we could use additional information about the respondents' characteristics (e.g., desirability measures) to set the faking parameters of the replacement model as a function of these additional information.

Another limitation of the current version of the SGR approach is related to its pure descriptive nature. So, for example, in its basic form SGR can be useful for describing what the informational structures of fake data are but not for how they actually operate according to specific psychological processes. A possible way out would be to use an appropriate reparameterization of the replacement distribution on the basis of, for example, the optimal IRT approach. In this particular reparameterization, faking could be modeled as a change in the trait level of the individual that gives rise to the fake responses via the theta-shift parameterization (Zickar and Drasgow 1996). Alternatively, we might assume that while the trait levels of the individuals remain invariant, the item parameters can vary according to the differential effect of faking (Ferrando and Anguiano-Carrasco 2013). Therefore, although encouraging, the promise of the SGR approach should be examined across more varied conditions. We acknowledge that more work still needs to be done.

Nevertheless, one benefit of the SGR analysis is that it allows detailed exploration of what outcomes are produced by particular sets of faking assumptions. By changing the input in the model parameters and showing the effect on the outcome of a model, SGR provides a *what-if-analysis* of the faking scenarios. Therefore, the essential characteristic of SGR is its explicit use of mathematical models and appropriate probability distributions for quantifying uncertainty in inferences based on possible fake data. Moreover, SGR involves the derivation of new statistical results as well as the evaluation of the implications of such new results: Are the substantive conclusions reasonable? How sensitive are the results to the modeling assumptions about the process of faking? In sum, SGR takes an interpretation perspective by incorporating in a global model all the available information about the process of faking.

Clearly, SGR is different from other statistical approaches that, instead, are more oriented in solving the fake identification problem by using *ad hoc* empirical paradigms such as, for example, *coached faking* or *ad-lib faking* (e.g., Ferrando 2005; Ferrando and Anguiano-Carrasco 2011; Fox and Meijer 2008; Holden and Book 2009; Leite and Cooper 2010; McFarland and Ryan 2000; Paulhus 1991; Zickar and Robie 1999; Ziegler and Buehner 2009). In addition, SGR is also different from RR, which, instead, tries to estimate the true responses by using randomization to encourage honest reports. Finally, we think that SGR may complement or even integrate techniques like RR and new relevant SGR developments may indeed lie in applying it to diverse problems beyond those considered here (i.e., for different types of data and/or with different probabilities of faking for statistical units).

References

Bandura A, Caprara G, Barbaranelli C, Gerbino M, Pastorelli C (2003) Role of affective self-regulatory efficacy in diverse spheres of psychosocial functioning. Child Dev 74(3):769–782

Campbell AA (1987) Randomized-response technique. Science 236(4805):1049

Caprara GV (2001) La valutazione dell'autoefficacia. Costrutti e strumenti. Erikson, Trento, IT

Chaudhuri A, Mukerjee R (1988) Randomised response theory and technique. Marcel Dekker Inc, New York

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Erlbaum Associates, Hillsdale, NJ

Cohen JE (1987) Sexual-behavior and randomized responses. Science 236(4808):1503

Cronbach LJ (1946) A case study of the split-half reliability coefficient. J Educ Psychol 37:473–480

Donovan JJ, Dwight SA, Schneider D (2013) The impact of applicant faking on selection measures, hiring decisions, and employee performance. Journal of Business and Psychology. Online First Article

Ellingson JE, Sackett PR, Hough LM (1999) Social desirability corrections in personality measurement: issues of applicant comparison and construct validity. J Appl Psychol 84(2):155–166

Ellingson JE, Smith DB, Sackett PR (2001) Investigating the influence of social desirability on personality factor structure. J Appl Psychol 86:122–133

Eysenck SB, Eysenck HJ, Shaw L (1974) The modification of personality and lie scale scores by special 'honesty' instructions. Br J Soc Clin Psychol 13:41–50

Ferrando PJ (2005) Factor analytic procedures for assessing social desirability in binary items. Multivar Behav Res 40:331–349

Ferrando PJ, Anguiano-Carrasco C (2011) A structural equation model at the individual and group level for assessing faking-related change. Struct Equ Model A Multidiscip J 18:91–109

Ferrando PJ, Anguiano-Carrasco C (2013) A structural model–based optimal person-fit procedure for identifying faking. Educ Psychol Meas 73(2):173–190

Fox JA, Tracy PE (1986) Randomized response: a method for sensitive surveys. In: Quantitative applications in the social sciences. Sage Publications, Inc., California

Fox J-P, Meijer RR (2008) Using item response theory to obtain individual information from randomized response data: an application using cheating data. Appl Psychol Meas 32(8):595–610

Furnham A (1986) Response bias, social desirability and dissimulation. Personal Individ Differ 7:385–400

Galić Z, Jerneić v, Kovačić M (2012) Do applicants fake their personality questionnaire responses and how successful are their attempts? A case of military pilot cadet selection. Int J Sel Assess 20(2):229–241

Griffith R, Converse P (2011) The rules of evidence and the prevalence of applicant faking. In: Ziegler M, MacCann C, Roberts R (eds) Faking in personality assessment: reflections and recommendations. Oxford University Press, pp 34–52

Helton J, Johnson J, Salaberry C, Storlie C (2006) Survey of sampling based methods for uncertainty and sensitivity analysis. Reliab Eng Syst Saf 91:1175–1209

Hesketh B, Griffin B, Grayson D (2004) Applicants faking good: evidence of item bias in the NEO PI-R. Personal Individ Differ 36:1545–1558

Holden RR, Book AS (2009) Using hybrid Rasch-latent class modeling to improve the detection of fakers on a personality inventory. Personal Individ Differ 47(3):185–190

Horrace W (2005) Some results on the multivariate truncated normal distribution. J Multivar Anal 94(1):209–221

Horvath P, Zuckerman M (1993) Sensation seeking, risk appraisal and risky behavior. Personal Individ Differ 14(1):41–52

Jöreskog K (1990) New developments in LISREL: analysis of ordinal variables using polychoric correlations and weighted least squares. Qual Quant 24:387–404

Jöreskog K, Sörbom D (1996) PRELIS 2: user's reference guide. Scientific Software International, Inc., Lincolnwood, IL

Kolata G (1987) How to ask about sex and get honest answers. Science 236(4800):382

Landers RN, Sackett PR, Tuzinski KA (2011) Retesting after initial failure, coaching rumors, and warnings against faking in online personality measures for selection. J Appl Psychol 96(1):202–210

Lee SY, Poon WY, Bentler P (1990) A 3-stage estimation procedure for structural equation models with polytomous variables. Psychometrika 55(1):45–51

Lehmann EL (1993) The fisher, Neyman-Pearson theories of testing hypotheses: one theory or two? J Am Stat Assoc 424:1242–1249

Leite WL, Cooper LA (2010) Detecting social desirability bias using factor mixture models. Multivar Behav Res 45:271–293

Levin RA, Zickar MJ (2002) Investigating self-presentation, lies, and bullshit: understanding faking and its effects on selection decisions using theory, field research, and simulation. In: Brett JM, Drasgow F (eds) The psychology of work. Lawrence Erlbaum Associates, Mahwah, pp 253–275

Lombardi L, Pastore M (2012) Sensitivity of fit indices to fake perturbation of ordinal data: a sample by replacement approach. Multivar Behav Res 47:519–546

Lombardi L, Pastore M (2014) sgr: a package for simulating conditional fake ordinal data. The R Journal 6(1):164–177

MacCann C, Ziegler M, Roberts R (2011) Faking in personality assessment: reflections and recommendations. In: Ziegler M, MacCann C, Roberts R (eds) Faking in personality assessment: reflections and recommendations. Oxford University Press, pp 309–329

McFarland LA, Ryan AM (2000) Variance in faking across noncognitive measures. J Appl Psychol 85:812–821

Miller JD (1981) Complexities of the randomized response solution. Am Sociol Rev 46(6):928–930

Muthén B (1984) A general structural equation model with dichotomous, ordered categorical and continuous latent variables indicators. Psychometrika 49:115–132

Parker D, Manstead A, Stradling S, Reason J (1992) Detyerminants of intention to commit driving violations. Accid Anal Prev 24(1):117–134

Pastore M, Lombardi L (2014) The impact of faking on Cronbach's Alpha for dichotomous and ordered rating scores. Qual Quant 48:1191–1211

Paulhus DL (1984) Two-component models of socially desirable responding. J Personal Soc Psychol 46:598–609

Paulhus DL (1991) Measurement and control of response bias. In: Robinson JP, Shaver PR, Wrightsman LS (eds) Measures of personality and social psychological attitudes. Academic Press, New York, pp 17–59

Pauls CA, Crost NW (2005) Effects of different instructional sets on the construct validity of the NEO-PI-R. Personal Individ Differ 39(2):297–308

Pek J, MacCallum RC (2011) Sensitivity analysis in structural equation models: cases and their influence. Multivar Behav Res 46:202–228

Rosse JG, Stecher MD, Miller JL, Levin RA (1998) The impact of response distortion on preemployment personality testing and hiring decisions. J Appl Psychol 83:634–644

Taubman-Ben-Ari O, Mikulincer M, Iram A (2004) A multi-factorial framework for understanding reckless driving–appraisal indicators and perceived environmental determinants. Transp Res Part F Traffic Psychol Behav 7(6):333–349

Topping GD, O'Gorman J (1997) Effects of faking set on validity of the NEO-FFI. Personal Individ Differ 23(1):117–124

Tracy PE, Fox JA (1981) The validity of randomized response for sensitive measurements. Am Sociol Rev 46(2):187–200

Viswesvaran C, Ones DS (1999) Meta-analyses of fakability estimates: implications for personality measurement. Educ Psychol Meas 59:197–210

Warner SL (1965) Randomized response: a survey technique for eliminating evasive answer bias. J Am Stat Assoc 60(309):63–69

Yang-Wallentin F, Joreskog KG, Luo H (2010) Confirmatory factor analysis of ordinal variables with misspecified models. Struct Equ Model-A Multidiscip J 17(3):392–423

Zickar MJ, Drasgow F (1996) Detecting faking on a personality instrument using appropriateness measurement. Appl Psychol Meas 20:71–87

Zickar MJ, Gibby RE, Robie C (2004) Uncovering faking samples in applicant, incumbent, and experimental data sets: an application of mixed-model item response theory. Organ Res Methods 7:168–190

Zickar MJ, Robie C (1999) Modeling faking good on personality items: an item-level analysis. J Appl Psychol 84:551–563

Ziegler M, Buehner M (2009) Modeling socially desirable responding and its effects. Educ Psychol Meas 69(4):548–565