

## MEASURING SOCIAL DESIRABILITY RESPONDING. A SHORT VERSION OF PAULHUS' BIDR 6

ANDREA BOBBIO  
ANNA MARIA MANGANELLI  
UNIVERSITY OF PADOVA

---

The aim of this research was to devise a short version of the BIDR 6 scale, a well-known measure of the two main dimensions of socially desirable responding: self-deceptive enhancement and impression management. Three correlational studies are described, all conducted with Italian respondents. In the first, a sample of non-student adults and a sample of university students were involved. Exploratory and confirmatory factor analyses and reliability analysis were applied. Invariance of the factor structure was tested via the multi-sample procedure. A 16-item reliable version was achieved and the factor structure, comparing both non-student adults vs. university students and male vs. female samples, turned out to be invariant. Differences for age, gender, and level of education were addressed. In the second study, data were collected by means of a Web-based questionnaire. Factorial structure and reliability of the BIDR 6 short version scale found support. Gender differences were addressed and discussed. In the third study the scale was administered in an organizational context along with internal Locus of Control, Self-Efficacy, Alienation, Hopefulness scales, and a short version of the Crowne and Marlowe's Social Desirability scale. Again, the 16-item BIDR 6 turned out to be reliable, the two-latent factor structure was endorsed by data, and the correlations with all the other measures supported its validity.

Key words: BIDR; Impression management; Self-deceptive enhancement; Short-version; Social desirability.

*Correspondence concerning this article should be addressed to Andrea Bobbio, Dipartimento di Psicologia Applicata, Università di Padova, Via Venezia 8, 35131 Padova (PD), Italy. E-mail: andrea.bobbio@unipd.it*

---

### INTRODUCTION

The response style defined as Socially Desirable Responding (SDR) refers to the systematic tendency to give overly positive answers that make the respondent look good (Paulhus, 2002). SDR is a widely studied topic in methodology, personality, and social psychology and continues to be a serious concern in many fields of psychological assessment because of its potential to introduce response bias (Roccatò, 2003). In fact, a remarkable amount of evidence supported by cross-cultural research suggests that traditional self-report measures including surveys and questionnaires administered through the World Wide Web (Hancock & Flowers, 2001) distort the report of racist attitudes, nonnormative sexual attitudes, values, cultural orientation, deviant or socially undesirable behaviours (such as substance abuse or declaring bankruptcy), socially desirable behaviours (such as attending religious services or voting), and personality traits (Holtgraves, 2004). However, it should be noticed that the existing positive correlation between SDR and cultural values has recently been considered a support for values' relevance in the social context, and not only as an error phenomenon (Ferrari, Bristow, & Cowman, 2005; Lalwani, Shavitt, & Johnson, 2006; Li & Bagger, 2006; Schwartz, Verkasalo, Antonovsky, & Sagiv, 1997). Social

desirability also appeared to be positively correlated with age and negatively correlated with level of education (Dijkstra, Smit, & Comijs, 2001; Heerwig & McCabe, 2009; Nuevo et al., 2009; Ray & Lovejoy, 2003). Other major issues discussed in the literature are (Holtgraves, 2004; Dijkstra et al., 2001): i) the extent to which social desirability is actually a problem for self-report measures, meaning that its importance may have been exaggerated; ii) whether individual differences in social desirability reflect true substantive variation in personality trait, rather than being a response bias only, iii) how social desirability should be conceptualized and measured, because most well-know measures of social desirability do not correlate highly with each other, and iv) how and when social desirability operates, for example, “does social desirability operate as an editing process that occurs after a respondent has recalled the requested information or does it affect the manner in which people retrieve information in the first place? Is it a conscious and deliberate process or one that is largely automatic?” (Holtgraves, 2004, p. 162). In sum, a broad and complex debate exists on pervasiveness and impact of social desirability, involving scholars from different fields. The present research deals with a small portion of this extensive debate and focuses only on measurement of social desirability. The aim is to devise a short version of Paulhus’ (1991) Balanced Inventory of Desirable Responding (BIDR 6) that could facilitate comprehension and control of socially desirable responding thanks to a considerably reduced administration time.

A long research tradition starting from the 1950s and culminating in Paulhus’ works highlighted two main types of SDR, known as self-deceptive enhancement (SDE) and impression management (IM). SDE refers to an unconscious tendency to provide honest but positively biased self-reports with the aim of protecting positive self-esteem; it is a predisposition to see oneself in a favorable light, is positively related to narcissism, and to a form of “rigid overconfidence.” On the one hand, SDE was shown to be negatively correlated with the Big Five personality trait Neuroticism, while, on the other, it was positively correlated with Conscientiousness and Extraversion. Moreover, SDE showed positive correlation with perceived control, optimism, self-controlling behaviour, self-esteem, a general sense of capability, self-efficacy, and individualism (the Horizontal Individualism dimension or HI. For details see: Triandis & Gelfand, 1998. See also: Gravdal & Sandal, 2006; Lalwani et al., 2006; Paulhus, 1991, 1998a, 1998b; Scheier & Carver, 1985; Stöber, Dette, & Musch, 2002). High SDE people tend to exaggerate knowledge of various persons, events, and objects (Paulhus, Harms, Bruce, & Lysy, 2003). IM refers to the habitual and conscious presentation of a favorable public image. It is closely related to faking, correlates highly and positively with traditional lie scales, and was positively associated with dissimulation, including deceiving others, responsiveness to situational demands for positive self-presentation, Agreeableness, Conscientiousness and Openness Big Five traits, and to Collectivism (the Horizontal Collectivism dimension or HC) (Holden, Book, Edwards, Wasylkiw & Starzyk, 2003; Lalwani et al., 2006; Paulhus, 2002; Schlenker & Britt, 1999; Stöber et al., 2002).

The 40-item BIDR 6, designed by Paulhus (1991) in order to measure SDE (20 items) and IM (20 items) is, after Crowne and Marlowe’s scale (MC-SDS; Crowne & Marlowe, 1960), one of the most widely used instruments to detect SDR. It supplied reliable scores (with Cronbach’s alpha ranging from .68 to .80 for SDE, from .75 to .86 for IM, and close to .80 when all 40 items are used) (Laynon & Carle, 2007; Stöber et al., 2002), to have satisfactory convergent and discriminant validity, and to be independent of psychopathology (Paulhus, 1991, 1998a). Moreover, the BIDR is

made up of an equal number of attribution and denial items for each of the two 20-item scales, and is therefore balanced. The correlation between SDE and IM is moderate, ranging from .05 to .40. Usually, men tend to show higher SDE scores than women but lower or the same level of IM (Helmes & Holden, 2003; Li & Bagger, 2007; Paulhus, 1991, 1998a; Riketta, 2005; Stöber et al., 2002). Recently, cross-nations (United States vs. Singapore), cross-culture (HI vs. HC), and cross-condition (standard vs. faking) measurement equivalence of the BIDR was tested by Li and Reb (2009). Support was found across groups formed on the basis of their HI and HC scores, but limited support was found across national and motivational conditions and therefore the issue of using the BIDR without proper controls is raised. As an example, Li and Reb suggested that noninvariant items in the scale could be carefully identified and excluded, but paying attention not to damage the construct validity of the scale.

Despite an ongoing debate on the nature and dimensionality of social desirability construct (Barger, 2002; Leite & Beretvas, 2005) and some issues on scale length, item wording, scoring method, impact of administration settings discussed in the literature (Fisher, 2000; Helmes & Holden, 2003; Li & Reb, 2009; Stöber et al., 2002), defining and measuring SDE and IM could be still considered noteworthy. Many studies aiming to investigate the factor structure of responses to SDR scales have been criticized for not using the appropriate statistical techniques to look at dimensionality issues and thus studies adopting a confirmatory factor analysis approach are needed (Leite & Beretvas, 2005).

The topic of scale length is particularly relevant when a researcher wants to detect if any of the scales included in a questionnaire are biased in a socially desirable manner: the costs/benefits ratio of adding a long scale may become excessive, especially if the 40-item BIDR is added to several other instruments more crucial for the study purposes. As regards scoring, both dichotomous and continuous methods are allowed by Paulhus, with a 5- or 7-point Likert-type answer scale. Recently Stöber et al. (2002) showed that the BIDR 6 scores derived from continuous 7-point scoring have advantages in comparison to those derived from dichotomous scoring in terms of reliability, convergent correlations, and sensitivity to instructional variations (fake-good vs. fake-bad conditions).

It was our aim to devise a short version of BIDR 6 and to collect data useful to support its reliability and validity. Few similar attempts are known in the literature leading to interesting and promising results, even if the dimensionality of short versions should be explored further (Leite & Beretvas, 2005; Musch, Brockhaus, & Bröder, 2002).

Samples of non-student adults and university students were involved in the three correlational studies described below. In the first, the short version was achieved by means of exploratory, confirmatory, and reliability analyses. Moreover, the invariance of factorial structure was tested via the multi-sample procedure comparing the male and female subgroups within the adult sample, and then the whole adult group with the university student group. Differences for gender and level of education on SDE and IM scores were examined. In the second study, the short version devised in Study 1 was included in a Web-based questionnaire to collect data supporting soundness of psychometric properties and factorial structure with a different administration method of growing popularity (Hancock & Flowers, 2001). Gender differences were also computed. Traditional (paper-and-pencil) and Web-based forms of personality scale are usually found to be largely equivalent, but this fact could not be taken for granted a priori. In fact, recommendations in the literature suggest conducting appropriate control studies (for a short over-

view, see also: Bobbio, Fochesato, Manganelli Rattazzi, & Crivellari, 2005). In the third study, part of a broader research project in an organizational context, further data were collected in order to support the validity of BIDR short version. In fact, another important aspect to evaluate scale validity is to explore relationships with constructs that are expected to show different correlation patterns, in our case with SDE and IM. The scale was administered to a group of adults along with internal locus of control, general self-efficacy, work alienation, and hopefulness measures, and a short version of Crowne and Marlowe's Social Desirability scale. Internal locus of control refers to whether individuals believe that their behaviour or, more precisely, the reinforcements from behaviour, are under their own control or not (Rotter, 1966). General self-efficacy is the expression of people's belief in their capabilities to produce designated levels of performance and to exercise influence over events and activities that affect their life (Schwarzer & Jerusalem, 1995). Work alienation represents a generalized, indifferent outlook toward work that indicates absence of enthusiasm and involvement (Maddi, Kobasa, & Hoover, 1979). Hopefulness or Hope is generally defined as a cognitive set comprising agency (belief in one's capacity to initiate and sustain actions) and pathways (belief in one's capacity to generate routes) to reach goals (Snyder et al., 1996). Crowne and Marlowe's scale (MC-SDS; Crowne & Marlowe, 1960), even though showing positive correlations with both SDE and IM (Stöber et al., 2002) seems mainly to capture the need for social approval or impression management (IM) (Helmes & Holden, 2003) and therefore could supply information about BIDR 6 factor convergent validity.

Even if it must be stated beforehand that only moderate correlations should be observed between social desirability scales and the other measures, given that too high correlation coefficients would constitute a serious threat for their construct validity, we hypothesized different patterns of results for SDE and IM dimensions. As for locus of control, self-efficacy, and hopefulness we predicted positive correlations with both SDE and IM with those involving SDE being higher than those involving IM. In fact, all these constructs mainly concern self-concept and have stronger implications for self-esteem, which is what SDE is supposed to have. On the contrary, work alienation is a domain-specific measure reflecting a low level of engagement in the work role, conceptually associated with diminished levels of positive psychological activation pertaining to the work endeavour and setting (Seeman, 1991) and consequently we argued that it does not appear to be functional to a positive self-presentation strategy in the work context. Thus, negative correlations were predicted with both SDE and IM, even if higher for IM. Finally, concerning the short version of the MC-SDS scale, we hypothesized positive correlations with SDE and IM, even if higher for IM than for SDE, because MC-SDS is supposed to capture mainly impression management strategies.

## STUDY 1

### Method

#### *Participants*

Two convenience samples took part in the research on a voluntary basis, without any form of compensation, financial or otherwise. Participants in the first sample were 704 non-

student adults, aged between 18 and 66 years, 327 males and 375 females (2 missing data, 0.3%), individually recruited with the collaboration of students attending introductory courses in Psychology. Mean age was 36.32 years ( $SD = 10.32$ ); 38.14 years ( $SD = 10.52$ ) for the male subgroup and 34.76 years ( $SD = 9.93$ ) for the female,  $t(699) = 4.34, p < .0001$ . The majority lived in Northern Italy (97%). As regards the level of education, 16.11% had completed compulsory schooling, 59.42% held a high school diploma, and 24.16% a university degree or more (3 missing data, 0.41%).

Participants in the second sample were 150 Psychology students from the University of Padua, aged between 19 and 42 years, 29 males and 118 females (3 missing data, 2%), recruited in class, after lessons. Mean age was 22.34 years ( $SD = 3.34$ ); 23.15 ( $SD = 3.27$ ) for the male group and 22.13 for the female ( $SD = 3.32$ ):  $t(144) = 1.43, ns$ . The majority lived in Northern Italy (82%).

### *Materials and Procedure*

The BIDR 6 was translated and adapted to the Italian context via the backtranslation method by two independent researchers, and with the support of a native English speaker. When a complete agreement on item translation and back-translation was reached, a pre-test was organized to determine comprehension of items, ease of understanding, and ability to respond. Two groups of adults and university students were involved, each one made up of 15 people. Their feedback helped us to improve item translation and wording. A 6-point Likert-type response scale was selected, ranging from 1 (*strongly disagree*) to 6 (*strongly agree*), without a central neutral point in order to force respondents to express judgments about themselves (Krosnick, 2002).

The 40-item BIDR 6 was included in a self-administered anonymous questionnaire and a few socio-demographic questions were also added (gender, age, geographic area, level of education). No special instructions were given to participants other than they were asked to answer all items as accurately as possible. Voluntary participation was highlighted and emphasised. Completion of the questionnaire was taken as consent to participate in the study. The maintenance of confidentiality with regard to individual findings was clearly assured as quantitative findings were to be reported as a group.

### *Data Analyses*

A two-stage approach was followed. In the first (Part I), conducted on the non-student adult group only, the short version was devised by means of exploratory and confirmatory factor analysis (estimation method: maximum likelihood, ML, applied to variance-covariance matrices). In the second stage (Part II) the short version factorial invariance was tested on the university student sample. Invariance of factor structure was checked via the multi-sample procedure (Byrne, 1998) which allows the simultaneous testing of data from different samples, constraining some or all parameters to be invariant between groups. Data analyses were carried out with SPSS 17 and LISREL 8.72. Goodness-of-fit was checked taking into account multiple criteria:  $\chi^2$ ,  $\chi^2/df$ , RMSEA, SRMR, and CFI. Recommended rule of thumb cutoff criteria are as follows:  $\chi^2$  not sig-

nificant, even if this rule should be considered with caution because it strongly depends on sample size;  $\chi^2/df \leq 2$  (very good), or between 2 and 3 (acceptable); RMSEA  $\leq .05$  (very good), or between .05 and .08 (acceptable); SRMR  $\leq .05$  (very good), or between .05 and .10 (acceptable); CFI  $\geq .97$  (very good), or between .95 and .97 (acceptable); in some analyses we also considered the AIC (Akaike Information Criterion) in order to compare competitive models (Schermele-Engel, Moosbrugger, & Müller, 2003).

Using PRELIS 2.54 kurtosis, skewness, and Mardia's (1970) multivariate kurtosis indices were computed: kurtosis and skewness are satisfactory if falling between  $-1.00$  and  $+1.00$  (Bollen, 1989) in order to sustain univariate normality of item distribution; Mardia's coefficient must vary between  $-1.96$  and  $+1.96$  to support data distribution multivariate normality. Moreover, the presence of univariate outliers was checked following indications by Tabachnick and Fidell (2001) ( $z > 3.29$ ).

Finally, Cronbach's alpha were computed and 95% confidence intervals (95% CIs) for the reliability estimates were calculated with the SPSS language developed by Fan and Thompson (2003).

## Results

### *Part I*

The group of 704 non-student adults was randomly split into two equal subgroups ( $n = 352$ ), the *calibration* and *validation* samples (Byrne, 1998). For all items skewness and kurtosis ranged between  $-1.00$  e  $+1.00$ ; Mardia's coefficient of relative multivariate kurtosis was 1.08 for the calibration sample and 1.08 for the validation sample; no univariate outliers were identified.

Considering only the calibration sample and following Gerbing and Hamilton (1996), who maintained that exploratory factor analysis can be used prior to analysis techniques aiming to confirm hypotheses on data structure, a principal component analysis (extraction criterion: eigenvalue  $> 1$ , varimax rotation) was performed on the correlation matrix among BIDR items to determine the number of factors to extract. Eleven components emerged accounting for 56.46% of the total variance. The first two, made up of seven and six items, with factor loadings ranging between  $|.44|$  and  $|.71|$ , expressed IM only; the third and fourth, composed of five and four items, with factor loadings between  $|.44|$  and  $|.73|$ , referred to SDE only. The other seven factors, made up of two or three items, resulted in a mixture of SDE and IM items. Both Cattell's scree test and results of a parallel analysis (O'Connor, 2000) suggested the extraction of three factors and the analysis was repeated using a factor analysis method (maximum likelihood procedure). In order to verify factors correlation oblimin rotation was applied. The percentage of total variance explained decreased to 21.75%. The first factor mainly expressed IM (10 items with factor loadings  $> .40$ ), the second expressed only SDE (10 items), and the third only IM (five items); the first and the third factor were found to be correlated (.30). Taking into account these results as well as the fact that Paulhus' model postulated two dimensions of SDR, we repeated the factor analysis procedure imposing the extraction of only two factors. Goodness-of-fit was  $\chi^2(208) = 516.90$ ,  $p \cong .000$  (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Percentage of total variance explained was 17.44. Ten items composed the SDE dimension and 11 the IM dimension (four IM

items showed factor loadings lower than .40). A preliminary check on Cronbach's alphas gave acceptable results: alpha was .74 for SDE, .76 for IM, and .75 for the whole scale (21 items).

A confirmatory factor analysis (CFA) was then performed to test a model with two latent variables ( $\xi$ ) and 21 observed variables ( $x$ ). The correlation between measurement errors was not allowed. The goodness-of-fit was not acceptable:  $\chi^2(188) = 533.21$ ,  $p \cong .000$ ,  $\chi^2/df = 2.84$ ; RMSEA = .07, CFI = .85, SRMR = .07. To improve the model, standardized residuals and correlations between error terms were taken into account, following guidelines in the literature (e.g., modification index > 3.84; Byrne, 1998; Jöreskog & Sörbom, 1996). Once detected, the most problematic items were deleted one by one and a new model was tested at each stage.

The final model with two-correlated factors and 16 observed variables showed a good data fit:  $\chi^2(103) = 184.72$ ,  $p \cong .000$ ,  $\chi^2/df = 1.79$ ; RMSEA = .048, CFI = .94, SRMR = .05, AIC = 250.68. Standardized coefficients were between  $|.39|$  and  $|.59|$  for SDE and between  $|.39|$  and  $|.65|$  for IM, all significant ( $p < .01$ ); the correlation between latent factors ( $\phi_{12}$ ) was .16 ( $p < .04$ ). Alpha for SDE (eight items) was .70 (95% CI: .64, .74); for IM (eight items) .73 (95% CI: .69, .77), and .71 (95% CI: .67, .75) for the entire scale (16 items) (see Appendix for items and details). Goodness-of-fit of a one-factor model with 16 indicators, that could constitute a more parsimonious solution, turned out to be not acceptable:  $\chi^2(104) = 629.26$ ,  $p \cong .000$ ,  $\chi^2/df = 6.05$ ; RMSEA = .12, CFI = .72, SRMR = .11, AIC = 693.26. It should be noted that for the SDE factor all "reversed" items were excluded from the short-version identified.

The same model was tested on the *validation sample* and showed a good fit:  $\chi^2(103) = 183.13$ ,  $p \cong .000$ ,  $\chi^2/df = 1.78$ ; RMSEA = .05, CFI = .94, SRMR = .06. Standardized coefficients ranged between  $|.31|$  and  $|.55|$  for SDE and between  $|.32|$  and  $|.67|$  for IM, all significant ( $p < .01$ ); the correlation between latent factors ( $\phi_{12}$ ) was .15 ( $p < .05$ ). Also in this case, goodness-of-fit of a one-factor model with 16 indicators turned out to be not acceptable:  $\chi^2(104) = 649.56$ ,  $p \cong .000$ ,  $\chi^2/df = 6.25$ ; RMSEA = .12, CFI = .73, SRMR = .11, AIC = 713.56. Alpha for SDE (eight items) was .70 (95% CI: .64, .74); for IM (eight items) .75 (95% CI: .70, .78); and .72 (95% CI: .68, .76) when all the 16 items were considered.

To examine invariance of factor structure between the calibration and validation samples the *multi-sample* procedure was applied with four consequential hypotheses: i) number of underlying factor equivalent (Baseline = B); ii) invariance of standardized coefficients ( $\lambda_x$ ) (Model 1, M<sub>1</sub>); iii) invariance of standardized coefficients and error variances ( $\theta_\delta$ ) (Model 2, M<sub>2</sub>); iv) invariance of standardized coefficients, error variances, and correlation between latent factors ( $\phi$ ) (Model 3, M<sub>3</sub>). As summarized in Table 1, none of the  $\chi^2$  difference tests were significant and all  $\Delta CFI$  were smaller than  $-.01$  (Cheung & Rensvold, 2002), and so we concluded in favour of model complete invariance.

Invariance of factor structure between the adult male ( $n = 327$ ) and female ( $n = 375$ ) subgroups was tested in a similar way. First of all, model goodness-of-fit was tested separately. For the male group it was:  $\chi^2(103) = 180.82$ ,  $p \cong .000$ ,  $\chi^2/df = 1.76$ ; RMSEA = .048, CFI = .93, SRMR = .063 (factor correlation was .26,  $p < .01$ ). For the female group it was:  $\chi^2(103) = 173.95$ ,  $p \cong .000$ ,  $\chi^2/df = 1.69$ ; RMSEA = .043, CFI = .95, SRMR = .05 (factor correlation was nonsignificant). As displayed in Table 2, all the  $\chi^2$  difference tests, apart from the one comparing Model 3 and Model 2, are not significant, and therefore we can conclude in favor of model complete invariance even if factor correlation turned out to be higher for the male than for the female group.

TABLE 1  
 Study 1  
 Multi-sample analysis. Calibration and validation sample, two-correlated factors model

Model	Fit indices	Test of invariance
Baseline (B)	$\chi^2(206) = 367.82, p \cong .000$ $\chi^2/df = 1.79$ RMSEA = .05 CFI = .94 SRMR = .06	
Model 1 (M <sub>1</sub> )	$\chi^2(220) = 378.82, p \cong .000$ $\chi^2/df = 1.72$ RMSEA = .05 CFI = .94 SRMR = .06	M <sub>1</sub> — B $\Delta\chi^2(14) = 11.00$ $p > .05$ $\Delta CFI = .00$
Model 2 (M <sub>2</sub> )	$\chi^2(236) = 395.17, p \cong .000$ $\chi^2/df = 1.67$ RMSEA = .04 CFI = .94 SRMR = .06	M <sub>2</sub> — M <sub>1</sub> $\Delta\chi^2(16) = 16.35$ $p > .05$ $\Delta CFI = .00$
Model 3 (M <sub>3</sub> )	$\chi^2(239) = 395.13, p \cong .000$ $\chi^2/df = 1.65$ RMSEA = .04 CFI = .94 SRMR = .06	M <sub>3</sub> — M <sub>2</sub> $\Delta\chi^2(3) = -.04$ $p > .05$ $\Delta CFI = .00$

TABLE 2  
 Study 1  
 Multi-sample analysis. Adult male and female groups, two-correlated factors model

Model	Fit indices	Test of invariance
Baseline (B)	$\chi^2(206) = 354.78, p \cong .000$ $\chi^2/df = 1.72$ RMSEA = .05 CFI = .94 SRMR = .05	
Model 1 (M <sub>1</sub> )	$\chi^2(220) = 364.25, p \cong .000$ $\chi^2/df = 1.66$ RMSEA = .04 CFI = .94 SRMR = .05	M <sub>1</sub> — B $\Delta\chi^2(14) = 9.47$ $p > .05$ $\Delta CFI = .00$
Model 2 (M <sub>2</sub> )	$\chi^2(236) = 375.67, p \cong .000$ $\chi^2/df = 1.59$ RMSEA = .04 CFI = .94 SRMR = .05	M <sub>2</sub> — M <sub>1</sub> $\Delta\chi^2(16) = 11.42$ $p > .05$ $\Delta CFI = .00$
Model 3 (M <sub>3</sub> )	$\chi^2(239) = 389.36, p \cong .000$ $\chi^2/df = 1.63$ RMSEA = .04 CFI = .94 SRMR = .06	M <sub>3</sub> — M <sub>2</sub> $\Delta\chi^2(3) = 13.69$ $p < .05$ $\Delta CFI = .00$

At the end of this first part, we reported goodness-of-fit indices for the overall adult group (*calibration + validation sample*,  $N = 704$ ):  $\chi^2(103) = 226.27$ ,  $p \cong .000$ ,  $\chi^2/df = 2.20$ ; RMSEA = .041, CFI = .95, SRMR = .045. Cronbach's alpha was .70 (95% CI: .66, .73) for SDE, .74 (95% CI: .71, .77) for IM, and .71 (95% CI: .69, .75) for the whole scale.

To examine the overlap between BIDR 6 short version and the original 40-item one we computed correlation coefficients according to the formula devised by Silverstein (1971) between the original SDE (alpha = .68; 95% CI: .61, .72) and the eight-item SDE obtaining .78. As regards the original IM (alpha = .77; 95% CI: .74, .81) and the eight-item IM the correlation was .79. Between BIDR 6 general score (alpha = .74; 95% CI: .68, .77) and that of the short version (16 items) it was .80.

### Part II

For the university student sample skewness and kurtosis ranged between  $-1.00$  and  $+1.00$  for all the items; Mardia's coefficient of relative multivariate kurtosis was 1.04; no univariate outliers were identified. Goodness-of-fit of the two-correlated factors and 16-indicators model was acceptable:  $\chi^2(103) = 141.69$ ,  $p \cong .000$ ,  $\chi^2/df = 1.38$ ; RMSEA = .05, CFI = .92, SRMR = .07. The correlation between measurement errors was not allowed. Standardized coefficients fell between  $|.32|$  and  $|.60|$  for SDE and between  $|.30|$  and  $|.64|$  for IM, all significant ( $p < .01$ ); the correlation between latent factors ( $\phi_{12}$ ) was .33 ( $p < .01$ ). Alpha for SDE (eight items) was .74 (95% CI: .66, .79) and .70 (95% CI: .58, .74) for IM (eight items); for the entire scale it was .71 (95% CI: .64, .77).

Multi-sample analysis allowed us to support the complete invariance of factor structure between the adult ( $N = 704$ ) and the student ( $N = 150$ ) sample (Table 3).

Finally, in order to examine the existing overlap between BIDR 6 short version and the original 40-item version in the university student sample we again computed correlation coefficients according to Silverstein's (1971) formula. Between the original SDE (alpha = .67, 95% CI: .61, .69) and the eight-item SDE it was .78. Between the original IM (alpha = .75; 95% CI: .72, .77) and the eight-item IM it was .80. Between the BIDR 6 general score (alpha = .72; 95% CI: .65, .75) and the short version one (16 items) it was .80.

### Differences between Groups

Differences between adults' and students' SDE and IM scores were examined through a multivariate analysis of variance (MANOVA) with a two-level between-factor (adults vs. students). The multivariate effect was significant,  $F(2, 851) = 27.26$ ,  $p < .0001$ ,  $\eta_p^2 = .06$ . Analysis of univariate effects showed that the adult group had the higher IM score (4.04 vs. 3.43),  $F(1, 852) = 54.30$ ,  $p < .0001$ ,  $\eta_p^2 = .06$ . The general BIDR score was higher for the adult than the student sample (3.86 vs. 3.51),  $t(852) = 6.31$ ,  $p < .0001$ .

Gender differences were investigated via a multivariate analysis of variance with gender as between-factor. In the case of adults, age was considered as a covariate for two reasons: i) it turned out to be correlated with SDE and IM, with  $r$  coefficients equal to .13 ( $p < .01$ ) and .27 ( $p < .0001$ ) respectively; ii) adult males and females had a significantly different mean age.

TABLE 3  
 Study 1  
 Multi-sample analysis. Adults and university students, two-correlated factors model

Model	Fit indices	Test of invariance
Baseline (B)	$\chi^2(206) = 367.96, p \cong .000$ $\chi^2/df = 1.79$ RMSEA = .04 CFI = .94 SRMR = .08	
Model 1 (M <sub>1</sub> )	$\chi^2(220) = 390.74, p \cong .000$ $\chi^2/df = 1.78$ RMSEA = .04 CFI = .94 SRMR = .08	M <sub>1</sub> — B $\Delta\chi^2(14) = 22.78$ $p > .05$ $\Delta CFI = .00$
Model 2 (M <sub>2</sub> )	$\chi^2(236) = 412.94, p \cong .000$ $\chi^2/df = 1.75$ RMSEA = .04 CFI = .94 SRMR = .09	M <sub>2</sub> — M <sub>1</sub> $\Delta\chi^2(16) = 22.20$ $p > .05$ $\Delta CFI = .00$
Model 3 (M <sub>3</sub> )	$\chi^2(239) = 415.14, p \cong .000$ $\chi^2/df = 1.74$ RMSEA = .04 CFI = .94 SRMR = .09	M <sub>3</sub> — M <sub>2</sub> $\Delta\chi^2(3) = 2.20$ $p > .05$ $\Delta CFI = .00$

As expected, the multivariate effect of age was significant,  $F(2, 697) = 36.11, p < .0001, \eta_p^2 = .09$ , as was that of gender,  $F(2, 697) = 19.07, p < .0001, \eta_p^2 = .05$ . Regarding gender, the univariate significant effects concerned both SDE,  $F(1, 698) = 11.56, p < .002, \eta_p^2 = .02$  (males = 3.79 vs. females = 3.58), and IM,  $F(1, 698) = 21.97, p < .0001, \eta_p^2 = .03$  (males = 3.92 vs. females = 4.14). Summarizing these results we could argue that, controlling for the influence of age, men showed higher SDE scores than women; vice versa for IM.

In the university student sample the multivariate effect of gender was significant,  $F(2, 142) = 4.27, p < .02, \eta_p^2 = .06$ , and the only significant univariate effect concerned IM,  $F(1, 143) = 7.52, p < .008, \eta_p^2 = .05$  (males = 3.09 vs. females = 3.51). Again, the female group showed higher IM scores.

Only for the adult group was a further analysis of variance performed, with level of education as a between factor (compulsory school degree,  $n = 113$ ; secondary school degree,  $n = 418$ ; university degree or more,  $n = 170$ ), and age as a covariate. The multivariate effects of age and level of education were significant,  $F(2, 695) = 32.174, p < .0001, \eta_p^2 = .09$  and  $F(4, 1392) = 3.81, p < .005, \eta_p^2 = .01$ . The univariate effect of level of education was significant only for SDE,  $F(2, 696) = 5.47, p < .005, \eta_p^2 = .02$ . Post-hoc analysis (Scheffé) showed that SDE mean scores for the university degree subgroup were lower than those of the compulsory and secondary school diploma subgroup, which did not differ from each other (3.50 vs. 3.72 and 3.78).

## Conclusion

Thanks to the use of exploratory and confirmatory factor analysis a short version of Paulhus' BIDR 6 was achieved. In sum, the 16-item scale proved to be reliable and its factorial structure to be invariant between distinct groups and subgroups of participants. The correlation between the original BIDR 6 and the BIDR 6 short version was high in all samples. Differences for gender and level of education were addressed. Non-student adults showed higher IM score compared to university students. With reference to gender differences, in the adult sample men showed higher SDE scores compared to women, while women scored higher on IM. In the university student sample the only difference between male and female respondents regarded the IM dimension, where women had the higher score. Finally, in the case of the adult group, we observed that the lower the level of education the higher the SDE score.

## STUDY 2

### Method

#### *Participants*

A convenience sample took part in the research on a voluntary basis, without any form of compensation, financial or otherwise. Participants were 690 individuals, aged between 14 and 62 years, 206 males (29.86%) and 484 females (70.14%). Mean age was 25.45 years ( $SD = 7.34$ ); 26.8 years ( $SD = 7.75$ ) for the male subgroup, and 24.93 years ( $SD = 7.06$ ) for the female one,  $t(699) = 3.36, p < .0001$ .

#### *Materials and Procedure*

A Web-based questionnaire was developed using Question Mark Perception 4.1 (Question Mark Computing, 2005). This facilitated on-line completion and provided further guarantee of anonymity, as the responses were submitted to an independent member of the University's technical staff for collation. The questionnaire included the 16-item BIDR 6 short version devised in Study 1 followed by a few socio-demographic questions (gender, age). Participants were recruited in two ways: i) an email invitation including the Web link to the questionnaire was sent to all students attending an advanced course in Psychology at the University of Padua. Participants were invited to take part in a new online research in the field of social psychology by filling out the questionnaire and then forwarding the invitation and the Web link to all their email contacts, asking them to do the same; ii) the Web link to the questionnaire, along with instructions and requests, was posted by an independent member of the University's administrative staff on popular social networks ([www.facebook.com](http://www.facebook.com), [www.linkedin.com](http://www.linkedin.com)). A brief description of the actual aims of the study along with acknowledgement for participation was individually presented to each participant in the last page of the online questionnaire, that appeared immediately after submitting all answers. Voluntary participation was highlighted and emphasised. Completion of the online survey was taken as consent to participate in the study. The researchers clearly assured respondents of confidentiality with regard to individual results as quantitative findings were to be reported as a group.

---

Differently from Study 1, a 5-point Likert-type response scale was adopted for all measures included in the questionnaire, ranging from 1 = *strongly disagree*, to 5 = *strongly agree* (3 = *neither in disagreement nor in agreement*).

## Results

Before performing data analyses, univariate and multivariate normality of item distributions was checked. Skewness and kurtosis for all item distributions fell between  $-1.00$  and  $+1.00$ . Mardia's (1970) coefficient was 1.17; no univariate outliers were identified. The goodness-of-fit of BIDR short version turned out to be suitable (estimation method: ML):  $\chi^2(103) = 262.36$ ,  $p \cong .000$ ,  $\chi^2/df = 2.55$ ; RMSEA = .047, CFI = .93, SRMR = .047. The correlation between measurement errors was not allowed. Standardized coefficients fell between  $|.39|$  and  $|.64|$  for SDE and between  $|.32|$  and  $|.63|$  for IM, all significant ( $p < .01$ ) (see Appendix for details); the correlation between latent factors was low but significant ( $\phi_{12} = .11$ ,  $p < .01$ ). Cronbach's alpha for SDE was .69 (95% CI: .65, .72); it was .73 (95% CI: .69, .75) for IM and .69 (95% CI: .65, .72) for the whole scale.

Only IM resulted weakly but significantly correlated with age ( $r = .16$ ,  $p < .0001$ ). A multivariate analysis of variance with gender as between-factor and age as covariate highlighted the expected significant multivariate effect of age,  $F(2, 685) = 12.12$ ,  $p < .0001$ ,  $\eta_p^2 = .03$ , and the significant multivariate effect of gender,  $F(2, 685) = 11.56$ ,  $p < .0001$ ,  $\eta_p^2 = .03$ . Analysis of univariate effects revealed a significant difference only for the IM score,  $F(1, 689) = 19.79$ ,  $p < .0001$ ,  $\eta_p^2 = .03$  (male mean = 2.70,  $SD = .64$ ; female mean = 2.91,  $SD = .69$ ). For the SDE factor,  $F(1, 689) = 2.06$ , *ns* (male mean = 3.18,  $SD = .59$ ; female mean = 3.10,  $SD = .57$ ).

## Conclusion

The 16-item short BIDR version developed in Study 1 was administered via a Web-based survey adopting a 5-point response scale where the central neutral point was allowed. Results supported the scale's factorial structure and reliability; moreover, the correlation between the two latent factors — low but significant — and gender differences for the IM factor — with women scoring higher than men on IM when the influence of age was controlled as well — were consistent with results of Study 1 and with data in the literature.

## STUDY 3

### Method

#### *Participants*

Two-hundred and eighty-six adult workers aged between 25 and 50 years, took part in the research on a voluntary basis without any form of compensation. There were 140 males (49%) and 146 females (51%), all living in the North-East part of Italy. Mean age was 39.12 years ( $SD = 9.48$ ); 40.99 years ( $SD = 9.4$ ) for the male group and 37.34 years ( $SD = 9.27$ ) for the

female one,  $t(284) = 3.38$ ,  $p < .002$ . As regards level of education, the majority had a secondary school diploma ( $n = 149$ , 52.37% of the participants), both within the male (55.45%) and the female subgroup (49.35%). Most respondents were employed as white-collar workers (49.58%); 55.59% had been working in the same company for over 10 years.

### *Materials and Procedure*

All participants completed an anonymous questionnaire including the 16-item BIDR 6 short version and other measures, all already validated in the Italian context in previous studies. When the research questionnaire was collected, every participant was briefly informed about the aim of the study. Voluntary participation was highlighted and emphasised. Completion of the questionnaire was taken as consent to participate in the study. Confidentiality with regard to individual findings was assured because the researchers clearly stated that only quantitative findings for the whole group would be reported for scientific purposes. The questionnaire included the following measures, devised in previous studies.

*Internal locus of control.* This was measured with six items of the Internal IPC subscale by Levenson (1974), in the Italian adaptation by Nigro and Galli (1988). A sample item is: “When I get what I want, it’s usually because I worked hard for it.”

*General self-efficacy.* We used seven items from the General Self-Efficacy scale by Schwarzer and Jerusalem (1995) validated in the Italian context by Bobbio, Manganelli Rattazzi, and Zorzan (2008). A sample item is: “I am confident that I could deal efficiently with unexpected events.”

*Work alienation* was measured using seven items selected both from the Work Alienation Scale (Seeman, 1967) and the Alienation Test (Maddi et al., 1979). A sample item is: “I wonder why I work at all.”

*Hopefulness* was measured through five items from the Snyder State Hope Scale (Snyder et al., 1996). A sample item is: “I can think of many ways to reach my current goals.”

Both work alienation and hopefulness scale were validated in the Italian context by Bobbio et al. (2008).

Then a nine-item short version of Crowne and Marlowe’s Social Desirability scale developed by Manganelli Rattazzi, Canova, and Marcorin (2000) for the Italian context, was added. The nine items were taken from a short version proposed by Reynolds (1982) and are numbers 3, 6, 12, 19, 21, 26, 28, 30, and 33 of Crowne and Marlowe’s (1960) scale. A sample item is: “I sometimes try to get even, rather than forgive and forget.”

Finally, socio-demographic data were requested (gender, age, geographic area, level of education, job position, and job seniority). For all instruments except the BIDR 6 short version a 7-point Likert-type response scale was used. As for the 16-item BIDR 6, a 6-point Likert-type scale was adopted in order to force respondents to express their judgements. The response scales ranged from 1 = *strongly disagree*, to 6 or 7 = *strongly agree*.

### *Results*

Before performing data analyses, univariate and multivariate normality of item distributions was checked via PRELIS. Skewness and kurtosis for all item distributions fell between –

1.00 and + 1.00. Mardia's (1970) coefficients were comprised between 1.12 and 1.30: being not significant ( $< 1.96$ ) they gave support to the multivariate normality of item distributions. No univariate outliers were identified. The goodness-of-fit of BIDR short version turned out to be acceptable (estimation method: ML):  $\chi^2(103) = 184.38, p \cong .000, \chi^2/df = 1.79$ ; RMSEA = .05, CFI = .93, SRMR = .06. The correlation between measurement errors was not allowed. Standardized coefficients fell between  $|.35|$  and  $|.61|$  for SDE and between  $|.34|$  and  $|.68|$  for IM, all significant ( $p < .01$ ) (see Appendix for details); the correlation between latent factors ( $\phi_{12} = .07$ ) was not significant. In Table 4 Cronbach's alpha coefficients, all acceptable, and Pearson's correlation coefficients are summarized.

TABLE 4  
 Study 3  
 Cronbach's alpha, 95% CIs for reliability estimates and correlations between measures

Measures	A	95% CIs	SDE	IM
1. SDE (eight items)	.69	.59 – .71	–	–
2. IM (eight items)	.77	.73 – .81	.07	–
3. Internal Locus of Control	.72	.68 – .74	.31**	.05
4. General Self-Efficacy	.86	.82 – .88	.42**	.13*
5. Alienation	.76	.71 – .80	.07	–.24**
6. Hopefulness	.80	.76 – .83	.28**	.05
7. MC-SDS	.71	.65 – .74	.06	.60**

\*\*  $p < .01$ ; \*  $p < .05$ .

SDE is positively correlated with internal locus of control, general self-efficacy, and hopefulness. No significant correlation emerged between SDE, work alienation, and MC-SDS. Concerning IM, a negative correlation was found with work alienation, which could be seen as the opposite of a positive self-presentation strategy in the work context, and a positive correlation was found with MC-SDS. No significant correlations emerged between IM, internal locus of control, general self-efficacy, and hopefulness. Therefore, these constructs appeared to be more connected with individual and inner beliefs rather than with impression management strategies.

### Conclusion

The 16-item BIDR 6 version developed in Study 1 was applied to a group of adult participants in an organizational setting, along with internal locus of control, general self-efficacy, hopefulness, work alienation scales, and a short version of the MC-SDS scale. Again, results supported the scale's factorial structure and reliability. The correlations confirmed the possibility of distinguishing between the two main types of social desirability, SDE and IM: the former more focused on the protection of a positive self-esteem and the latter oriented toward the strategic presentation of a favorable public image.

---

## GENERAL DISCUSSION

A short version of Paulhus's (1991) BIDR 6 was developed and tested in three studies and its factorial structure was cross-validated using different samples. Data supporting its validity were also collected. The 16-item two-latent factor BIDR 6 short version showed satisfactory psychometric properties with paper-and-pencil and a Web-based administration method, adopting a Likert-type response scale including or excluding a central neutral point (6-point vs. 5-point). Moreover, its factorial structure proved to be invariant between non-student adult males versus females and between the whole adult versus the university student group. Importantly, the 16-item version turned out to be highly correlated with the original 40-item BIDR 6 Italian translation and alpha coefficients suggested that the reliability was not weakened by the item selection and reduction procedure. The validation of short versions of the BIDR 6 scale is an important issue because the cost of including a long scale along with other measures of primary interest to a researcher can become prohibitive.

Regarding differences for SDE and IM scores, first of all we noted that the non-student adults scored higher than university students on IM. This difference may be interpreted as a consequence of the higher concern for presenting a favorable public image usually correlated with age that is reported in the literature (e.g., Dijkstra et al., 2001). Then, in the adult group, men showed the higher SDE score while women the higher IM scores. Within the student group and the online group, differences were limited to the IM factor, and always in favor of women. This data mirror what is reported by Paulhus (1991), Riketta (2005), and Gravdal and Sandal (2006).

Furthermore, in the non-student group, the higher the level of education the lower the SDE score turned out to be. In interpreting this trend we may argue that SDE could be used as a sort of compensation or buffering strategy for people who see their education or perceive their social status as comparatively low. This hypothesis, along with the age-related IM tendency, could be an interesting hypothesis for future experimental studies.

Finally, SDE was positively correlated with internal locus of control, general self-efficacy, and hopefulness measures, while IM was negatively correlated with the work alienation measure and positively correlated with the short form of Crowne and Marlowe's Social Desirability scale. Altogether it is possible to conclude that our hypotheses found support. We had confirmation that while SDE mainly referred to a positive bias in item responses protecting self-esteem (declare to be able to master one's actions and fate and to be confident in positive events to come), on the other hand, IM was primarily referred to dissimulation of item responses in order to make a favorable impression on others (consciously hiding potential weakness, such as doubts about the meaning of everyday's work activities).

However, some limitations suggest caution and the need for supplementary studies. First, all of our studies were conducted on convenience samples recruited on a voluntary basis. So, they do not represent the overall Italian population. Second, even though great care was taken in the translation of the original BIDR, it is well-known that the literal translation of questionnaires may sometimes fail to capture the same psychological content as the original (Van de Vijver & Hambleton, 1996): this could be at least one of the reasons determining the lack of "denial" item for the SDE sub-scale. As a matter of fact, the final SDE 8-item scale is composed only by "attribution" items and so it potentially misrepresents the original balanced item pool, albeit it showed a strongly coherent correlational pattern with both the original SDE and the measures introduced in

Study 3. Third, strictly speaking the 6- and 5-point Likert-type answer formats chosen for our studies do not allow to generalize the findings in the case of 7-point or dichotomous response scales. Fourth, it should also be noted that standardized coefficients ( $\lambda_x$ ) for both SDE and IM dimensions are sufficiently high rather than very high, as are the reliability coefficients. Nevertheless, these results are consistent with those in the literature and seem also typical of social desirability measures that are made up of items with quite a heterogeneous content (Leite & Beretvas, 2005; Paulhus, 1991; Stöber et al., 2002).

In conclusion, the BIDR 6 short version presented here could be a useful research instrument for social scientists and methodologists, at least in the Italian context. The impact of the potential weaknesses discussed here could be lessened by extending the research to different populations and different countries, and seeing how the results correspond. Particularly for its shortness, the BIDR 6 short version could lead to a better comprehension and control of the intriguing response style defined as socially desirable responding, because it could be easily added to both paper-and-pencil and Web-based questionnaires.

#### ACKNOWLEDGMENTS

The authors wish to thank Alessia Coccioli, Chiara Dainese, Elisa Fanton, and Andrea Zorzan for their assistance in collecting the data, and are grateful to Dr. Simone Fluperi and Dr. Mauro Sarrica for technical support with the Web-based questionnaire. Special thanks to Broni Falinska for several useful suggestions.

#### REFERENCES

- Barger, S. D. (2002). The Marlowe-Crowne affair: Short forms, psychometric structure, and social desirability. *Journal of Personality Assessment*, 79, 286-305.
- Bobbio, A., Fochesato, M., Manganelli Rattazzi, A. M., & Crivellari, F. (2005). La scala di controllo sociopolitico di Zimmerman e Zahniser (SPCS, 1991). Un contributo all'adattamento italiano con metodi carta e matita e Web-based [The Zimmerman and Zahniser' Sociopolitical Control Scale (SPCS, 1991). A contribution to the adaptation to the Italian context using paper-and-pencil and Web-based forms]. *TPM*, 12, 61-82.
- Bobbio, A., Manganelli Rattazzi, A. M., & Zorzan, A. (2008). Measuring psychological empowerment in the workplace. A test of Spreitzer's model. *Bollettino di Psicologia Applicata*, 255, 11-20.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS and SIMPLIS: Basic concepts, applications and programming*. London-New Jersey: Lawrence Erlbaum Associated Publishers.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Dijkstra, W., Smit, J. H., & Comijs, H. C. (2001). Using social desirability scales in research among the elderly. *Quality and Quantity*, 31, 107-115.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of factor analysis in psychological research. *Psychological Methods*, 4, 272-299.
- Fan, X., & Thompson, B. (2003). Confidence intervals about score reliability coefficients. In B. Thompson (Ed.), *Score reliability: Contemporary thinking on reliability issue* (pp.69-84). Thousand Oaks, CA: Sage.
- Ferrari, J. R., Bristow, M., & Cowman, S. (2005). Looking good or being good? The role of social desirability tendencies in student perceptions of institutional mission and values. *College Student Journal*, 39, 7-13.
- Fisher, R. (2000). The future of social-desirability bias research in marketing. *Psychology and Marketing*, 17, 73-77.

- Gerbing, D. W., & Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling*, 3, 62-72.
- Gravdal, L., & Sandal, G. M. (2006). The two-factor model of social desirability: Relation to coping and defense, and implication for health. *Personality and Individual Differences*, 40, 1051-1061.
- Hancock, D. R., & Flowers, C. (2001). Comparing social desirability responding on World Wide Web and paper-administered surveys. *Educational Technology Research and Development (ETR&D)*, 49, 5-13.
- Heerwig, J. A., & McCabe, B. J. (2009). Education and social desirability bias: The case of a black presidential candidate. *Social Science Quarterly*, 90, 674-686.
- Helmes, E., & Holden, R. R. (2003). The construct of social desirability: One or two dimensions? *Personality and Individual Differences*, 34, 1015-1023.
- Holden, R. R., Book, A. S., Edwards, M. J., Wasylkiw, L., & Starzyk, K. B. (2003). Experimental faking in self-report psychopathology: Unidimensional or multidimensional? *Personality and Individual Differences*, 35, 1107-1117.
- Holtgraves, T. (2004). Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin*, 30, 161-172.
- Jöreskog, K. J., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software Inc.
- Krosnick, J. A. (2002). The causes of no-opinion responses to attitude measures in surveys: They are rarely what they appear to be. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 64-85). New York: Wiley.
- Lalwani, A., Shavitt, S., & Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology*, 90, 165-178.
- Laynon, R. I., & Carle, A. C. (2007). Internal and external validity of scores on Balanced Inventory of Desirable Responding and Paulhus Deception scales. *Educational and Psychological Measurement*, 67, 859-876.
- Leite, W. L., & Beretvas, S. N. (2005). Validation of scores on the Marlowe-Crowne Social Desirability Scale and the Balanced Inventory of Desirable Responding. *Educational and Psychological Measurement*, 65, 140-154.
- Levenson, H. (1974). Activism and powerful others: Distinctions within the concept of Internal-External Control. *Journal of Personality Assessment*, 38, 377-383.
- Li, A., & Bagger, J. (2006). Using the BIDR to distinguish the effects of Impression Management and Self-Deception on the criterion validity of personality measures: A meta-analysis. *International Journal of Selection and Assessment*, 14, 131-141.
- Li, A., & Bagger, J. (2007). The Balanced Inventory of Desirable Responding (BIDR). A reliability generalization study. *Educational and Psychological Measurement*, 67, 525-544.
- Li, A., & Reb, J. (2009). A cross-nation, cross-cultures, and cross-conditions analysis on the equivalence of the Balanced Inventory of Desirable Responding. *Journal of Cross-Cultural Psychology*, 40, 214-233.
- Maddi, S. R., Kobasa, S. C., & Hoover, M. (1979). An alienation test. *Journal of Humanistic Psychology*, 19, 73-76.
- Manganeli Rattazzi, A. M., Canova, L., & Marcorin, R. (2000). La desiderabilità sociale. Un'analisi di forme brevi della scala di Marlowe e Crowne [Social Desirability. An analysis of short forms of the Marlowe-Crowne Social Desirability Scale]. *TPM*, 7, 5-17.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.
- Musch, J., Brockhaus, R., & Bröder, A. (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit [An inventory for the measurement of two factors of social desirability]. *Diagnostica*, 48, 121-129.
- Nigro, G., & Galli, I. (1988). *La fortuna, l'abilità, il caso. Introduzione allo studio del locus of control* [Luck, skill, chance. Introduction to the study of Locus of Control]. Torino, IT: Centro Scientifico Torinese.
- Nuevo, R., Montorio, I., Márquez-González, M., Cabrera, I., Izal, M., & Pérez-Rojo, G. (2009). Diferencias asociadas a la edad en el efecto de la deseabilidad social en el autoinforme del estado emocional [Age-related differences in the effect of social desirability on self-reported emotional state]. *Revista Española de Geriatria y Gerontología*, 44, 85-89.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32, 396-402.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1, pp. 17-59). San Diego, CA: Academic Press.
- Paulhus, D. L. (1998a). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology*, 74, 1197-1208.

- Paulhus, D. L. (1998b). *Paulhus Deception Scales: Users' manual*. North Tonawanda, NY: Multi-Health System.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 67-88). Hillsdale, NJ: Erlbaum.
- 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*, 890-904.
- Question Mark Computing Ltd. (2005). *Questionmark Perception* (Version 4.1). London: Question Mark Computing.
- Ray, J. J., & Lovejoy, F. H. (2003). Age-related social desirability responding among Australian women. *The Journal of Social Psychology, 143*, 669-671.
- Reynolds, W. N. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 38*, 119-125.
- Riketta, M. (2005) Gender and socially desirable responding as moderators of the correlation between implicit and explicit self-esteem. *Current Research in Social Psychology, 11*, 14-28.
- Roccato, M. (2003). *Desiderabilità sociale e acquiescenza* [Social desirability and acquiescence]. Milano: LED.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, 1-28.
- Scheier, M. E., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4*, 219-247.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*, 23-74.
- Schlenker, B. R., & Britt, T. M. (1999). Beneficial impression management: Strategically controlling information to help friends. *Journal of Personality and Social Psychology, 76*, 559-573.
- Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: Nfer-Nelson.
- Schwartz, S. H., Verkasalo, M., Antonovsky, A., & Sagiv, L. (1997). Value priority and social desirability: Much substance, some style. *British Journal of Social Psychology, 36*, 3-18.
- Seeman, M. (1967). On the personal consequences of alienation in work. *American Sociological Review, 32*, 273-285.
- Seeman, M. (1991). Alienation and anomie. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Vol. I, pp. 291-371). San Diego, CA: Academic Press.
- Silverstein, A. B. (1971). A corrected formula for assessing the validity of WAIS, WISC, and WPPSI short forms. *Journal of Clinical Psychology, 27*, 212-213.
- Snyder, C. R., Sympson, S. C., Ybasco, F. C., Borders, T. F., Babyak, M. A., & Higgins, R. L. (1996). Development and validation of the State Hope Scale. *Journal of Personality and Social Psychology, 70*, 321-335.
- Stöber, J., Dette, D. E., & Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment, 78*, 370-389.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn & Bacon.
- Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology, 74*, 118-128.
- Van de Vijver, F. J. R., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist, 1*, 89-99.

APPENDIX

Italian Version of the BIDR 6 Short Form

		Standardized coefficients ( $\lambda_x$ )				
		C N = 352	V N = 352	S N = 150	W N = 690	O N = 286
<i>Self-Deceptive Enhancement (SDE)</i>						
1.	La prima impressione che ho sulla gente di solito si rivela corretta [My first impressions of people usually turn out to be right] ( <sup>‡</sup> )	.39	.31	.36	.40	.35
9.	So sempre perché qualcosa mi piace [I always know why I like things] ( <sup>‡</sup> ) ( <sup>§</sup> )	.53	.46	.54	.49	.36
11.	Una volta che ho preso una decisione, raramente gli altri riescono a farmi cambiare idea [Once I've made up my mind, other people can seldom change my opinion]	.42	.53	.58	.51	.40
13.	Ho il completo controllo del mio destino [I am fully in control of my own fate] ( <sup>§</sup> )	.43	.50	.54	.39	.41
17.	Non mi pento mai delle mie decisioni [I never regret my decisions] ( <sup>§</sup> )	.49	.55	.60	.54	.61
23.	Sono una persona completamente razionale [I am a completely rational person] ( <sup>‡</sup> ) ( <sup>§</sup> )	.51	.43	.32	.32	.39
27.	Sono molto sicuro/a dei miei giudizi [I am very confident in my judgements] ( <sup>‡</sup> ) ( <sup>§</sup> )	.59	.55	.57	.64	.46
31.	Se capita che io non piaccia a qualcuno, mi sento comunque a posto con me stesso/a [It's all right with me if some people happen to dislike me] ( <sup>§</sup> )	.43	.44	.56	.40	.41
<i>Impression Management (IM)</i>						
2.	A volte mento, se devo farlo [I sometimes tell lies, if I have to] (R) ( <sup>‡</sup> )	.52	.63	.39	.59	.61
6.	Ci sono state occasioni in cui ho approfittato di qualcuno [There have been occasions when I have taken advantage of someone] (R) ( <sup>‡</sup> )	.65	.67	.38	.63	.68
12.	Obbedisco sempre alle leggi, anche quando è improbabile che venga scoperto/a [I always obey laws, even if I'm unlikely to get caught] ( <sup>§</sup> )	.39	.32	.30	.32	.34
14.	Ho parlato male di un/a amico/a alle sue spalle [I have said something bad about a friend behind his or her back] (R)	.49	.49	.54	.52	.56
22.	Non ho mai gettato sporcizia per la strada [I have never dropped litter on the street] ( <sup>§</sup> )	.58	.53	.64	.41	.62
28.	Ho fatto delle cose che non racconterei ad altri [I have done things that I don't tell other people about] (R) ( <sup>‡</sup> )	.46	.49	.42	.48	.51
32.	Mi sono assentato/a per malattia, dal lavoro o da scuola, anche se non ero veramente malato/a [I have taken sick-leave from work or school even though I wasn't really sick] (R) ( <sup>‡</sup> ) ( <sup>§</sup> )	.55	.55	.45	.54	.61
35.	Ho alcune pessime abitudini [I have some pretty awful habits] (R)	.42	.50	.44	.47	.47

Note. Item number indicates the presentation order in the questionnaire. (R) = reversed score. (<sup>‡</sup>) = items in the short form by Musch et al. (2002); (<sup>§</sup>) = items in the short form by Leite and Beretvas (2005). Standardized coefficients are referred to the CFA completely standardized solutions, and respectively: C = Calibration sample; V = Validation sample; S = Student sample; W = Web-based sample; O = Organizational sample.