ORIGINAL ARTICLE

Social desirability dimensionality: one or two continua?

Krzysztof Fronczyk (ID) *1 · A,B,C,D,E,G*, *Ewa Witkowska* (ID) *2 · D,E,F*

1: Faculty of Psychology, University of Warsaw, Warsaw, Poland
2: Institute of Psychology, The Maria Grzegorzewska University, Warsaw, Poland

BACKGROUND

The objective of this paper is to examine the structure of social desirability. This variable is important in self-report questionnaires. Theoretical analyses indicate the existence of at least two dimensions of this construct, but there is no unambiguous empirical evidence for such a structure of social desirability. The present research aimed to determine the factor structure of the social desirability construct and the number of its dimensions.

PARTICIPANTS AND PROCEDURE

The inductive research strategy was applied and a new pool of questionnaire items was generated. This approach allows for detachment from the existing social desirability assessment instruments. This study focuses on the concepts of social desirability that examinees have, rather than on the content of existing items measuring social desirability. The verification of the social desirability structure was performed in three stages. At each stage, the number of items was gradually reduced. The successive phases of the study involved three different samples, respectively, consisting of 657, 415, and 1,709 examinees.

RESULTS

Initially, at the first stage of the study, two dimensions were detected, but in the two subsequent stages, the second dimension represented acquiescence. The procedure of partialling the mean out of the correlation matrix was applied to eliminate acquiescence. This dimension is related, as expected, to agreeableness and conscientiousness. It also appeared that the social desirability structure can be confounded by acquiescence in longer questionnaires.

CONCLUSIONS

It is concluded that every item reflects both social desirability and acquiescence, albeit to different extents. A one-dimensional construct of social desirability was finally obtained.

KEY WORDS

factor analysis; social desirability; acquiescence; balanced scale; inductive strategy

BACKGROUND

Social desirability is usually defined as the tendency to present oneself in a manner that will be viewed favorably by others. This psychological construct has been present within the field of interest of researchers for many years (Crowne & Marlowe, 1960; Edwards, 1957; Furnham, 1986; Krumpal, 2013; Paulhus, 2002), as it often manifests itself when self-report measures are used to assess personality traits, attitudes, or disorders. Social desirability might manifest itself in the form of over-reporting one's desirable characteristics and denying the existence of undesirable behavior – generally, describing oneself in an extremely positive manner (Crowne & Marlowe, 1960; Edwards, 1957; Holden & Passey, 2009; Krumpal, 2013; Paulhus, 2002). Moreover, context influences what is considered social desirable or not – e.g., some behaviors are desirable during a job interview, others with friends or family but not with strangers.

It was originally assumed that social desirability is a one-dimensional construct (Crowne & Marlowe, 1964; Edwards, 1957). Crowne and Marlowe (1964) defined it as a continuous extent that an individual behaves in a manner aimed to achieve specific reinforcements, which may include the approval of other people, their support and sympathy, as well as avoiding social rejection. However, several variants of the Crowne and Marlowe measure are characterized by poor psychometric properties (Loo & Thorpe, 2000). Besides this conceptualization of the social desirability, there are reports indicating its multidimensional structure, with a lack of full consensus among researchers regarding the number of dimensions and their psychological content. Thus the structure of the social desirability construct remains unclear.

The two-dimensional model of social desirability proposed by Paulhus (1984) distinguishes positive self-deception and impression management as separate aspects of this construct. Paulhus (1984) defines self-deceptive positivity as the tendency to give self-reports that are honest but positively biased, and defines impression management as deliberate self-presentation to the audience. In his early work on this model, Paulhus (1984) demonstrated that these two dimensions are orthogonal, although several studies (Loo & Loewen, 2004; Ramanaiah, Schill, & Leung, 1977; Ventimiglia & MacDonald, 2012) have repeatedly indicated that the structure of social desirability is somewhat different and that these dimensions are correlated (Musch, Brockhaus, & Bröder, 2002; Ones, Viswesvaran, & Reiss, 1996), so the orthogonal model fits the data poorly (Lanyon & Carle, 2007).

Paulhus persisted in his studies and in a joint work published with Reid (1991), he demonstrated a two-dimensional structure of social desirability in exploratory factor analysis. However, the factors that emerged do not fully match the original dimensions. The first factor was intended to measure impression management, in terms of both denial and attribution. The denial items of the self-deception scale also loaded on this factor. The second factor comprised only self-deception attribution items. This is not the only existing two-factor model. According to the alternative approach, social desirability can be considered as consisting of two dimensions: (1) denial of negative characteristics and (2) the enhancement of positive traits (Loo & Loewen, 2004; Ramanaiah et al., 1977). Such a structure was revealed in a large cross-cultural study by He et al. (2015) in a pooled sample of nearly 3,500 subjects from 20 countries. These studies are consistent with data reported by Espinosa and van de Vijver (2014) that yielded two factors for Marlowe-Crowne's scale, comprising positive and negative characteristics in Mexican society.

Pearson and Francis (1989) applied exploratory factor analysis (without rotation) to three different Eysenk's lie scales and obtained a two-factor structure of social desirability representing either psychological submissiveness or social conformity (i.e., component A) or "pure" propensity to present oneself in a positive manner to others (i.e., component B). The two-factor model was also identified by Ventimiglia and MacDonald (2012), but they also found some evidence supporting a one-dimensional model in their Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960) analysis. Nevertheless, they concluded on the superiority of the two-factor model (i.e., Impression Management and Self-deceptive Enhancement) of social desirability, basing this conclusion on the change in χ^2 for both factors. The factors of Impression Management and Self-Deception also emerged as result of principal component analysis applied to two language versions of an 11-item questionnaire designed by Shultz and Chavez (1994). One of these versions was English and the other one was Spanish. The two-dimensional structure was much less evident in the Spanish sample, which could indicate not only linguistic but also cultural differences pertaining to the understanding of this construct. In this case, less of the variance was explained by Impression Management and Self-Deception than in the English sample. Some psychological content was not covered by these two dimensions.

There are also reports of a more complex factorial structure of social desirability. Leite and Beretvas (2005), studying the Paulhus model of this construct, stated that in an investigated sample of students, Impression Management is two-dimensional and Self-Deception is one-dimensional. Kroner and Weekes (1996) also reported that three factors had emerged in their exploratory factor analysis. In this study, Paulhus's self-deception items divided and formed two separate factors: (1) denial of negative characteristics and (2) attribution of positive qualities. The third fac-

tor consisted of impression management items. These results were confirmed in an independent study by Cervellione, Lee, and Bonanno (2009), but rejected by Li and Li (2008) in a study of a Chinese sample. Loo and Thorpe (2000) obtained a similar solution in a replication study; however, these dimensions explain only 17% of the variance. Thus, some variance representing specific content is not explained by any factor.

Moral de la Rubia, Garcia-Cadena, and Antona Casas (2012) proposed an even more elaborate model comprising four dimensions: (1) positively-keyed impression management; (2) negatively-keyed impression management; (3) positively-keyed self-deception; and (4) negatively-keyed self-deception. A very similar solution was obtained by Li, Li, and Wang (2015) for a Chinese sample. From several alternative models, the model in which impression management and self-deception split into denial and enhancement fits the data best.

As the results were non-conclusive, Stöber (1999, 2001) decided to return to the original idea that social desirability is one-dimensional. He developed the Social Desirability Scale with a theoretical basis similar to Crowne and Marlowe's (1960) concept of social desirability. His approach was strictly theoretical. In the context of Crowne and Marlowe's theory, it means that items were generated assuming the one-dimensional nature of social desirability. Subsequent empirical studies indicate quite high levels of reliability and validity of this new instrument (Stöber, 1999, 2001).

Single measures of social desirability were applied in all the above-mentioned studies. This may mean that the variety of factor solutions is the result of differences in specific content items in each measure. To avoid this limitation, some researchers have jointly factor-analyzed items from more than one measure. Paulhus (1984), for example, factor analyzed the results of six different social desirability measures and obtained the familiar two-dimensional structure consisting of self-deception and impression management. In a similar study by Holden and Fekken (1989), items from three different social desirability measures were used and two dimensions emerged: (1) sense of own general capability and (2) interpersonal sensitivity. In Helmes and Holden's (2003) study, factor analysis of a number of different measures of social desirability revealed a factor that reflects sensitivity to social demands and a second factor that reflects self-sufficiency and personal adjustment. All of these studies yielded a two-dimensional structure; however, the psychological content of the obtained dimensions is different, and these results remain inconclusive.

There are authors who suggested that cultural differences could be the reason for different factor solutions. The main concern is that one social group may see some self-descriptions as desirable while another may not. Especially, at least in the context of European-Asian comparisons, Li et al. (2015) pointed to Eastern people's greater inclination to accept contradictions as compared to Westerners, who want to be more internally consistent in their self-descriptions. These authors suggest that this is the cause of the aforementioned splitting of the dimensions of self-deception and impression management into two separate factors of denial and enhancement in Chinese society. Espinosa and van de Vijver (2014), based on Schwartz (2011) and Marín, Triandis, Kashima, and Betancourt (1983), indicate that conceptions concerning what is socially desirable vary across cultures depending on specific needs, traits, and values; thus, instead of adaptation existing tools for use in diverse cultures or languages, emic measures would be more valid.

It should be kept in mind that the majority of popular social desirability scales still in use were developed decades ago (e.g. Paulhus' Balanced Inventory of Desirable Responding), or even half a century ago, e.g. Crowne and Marlowe's scale or Edwards' scale. Since the socio-cultural conditions in those days differed considerably from those prevailing today, this may lead to inadequacy in defining social desirability. In other words, it could be difficult to unambiguously indicate what is universally socially desirable and undesirable in different countries at different times.

The issue of social desirability is complicated not only because of the ambiguous structure of this variable but also because of its relationship with other response styles, for instance, the acquiescent response style, i.e. the tendency to agree with items irrespective of their content (Martin, 1964; Ray, 1983). Acquiescence and social desirability are viewed as variables that can be modeled as common factors (Ferrando, Lorenzo-Seva, & Chico, 2009). Hofstee, Ten Berge, and Hendricks (1998) suggested that acquiescence and social desirability are positively related because they share a tendency towards conformity or compliance. In the study by He and van de Vijver (2013), social desirability and acquiescence loaded on one factor with positive and negative loadings, respectively. Some other empirical evidence suggests, however, that they are essentially uncorrelated (Greenwald & Clausen, 1970; Stricker, 1963).

Regardless of the social desirability model, this variable is associated with certain personality traits. Upon comparison with dimensions of the Big Five, the strongest relationship is observed for agreeableness and conscientiousness (Borkenau & Ostendorf, 1992; Paulhus, 2002; Stöber, 2001). The captured correlations are usually interpreted as exaggerations of people's self-ratings of agreeableness components such as friendliness, cooperativeness, other-focus, and conscientiousness components such as reliabil-

ity, attentiveness, responsibility, organization and efficiency and so on. Thanks to such exaggerations, people with high social desirability can be viewed more favorably. Leary and Toner (2015) formulated an alternative interpretation, according to which agreeableness or conscientiousness is maintained, at least in part, by the desire to be positively viewed. In other words, people who make favorable impressions behave more conscientiously and are more conciliatory in an effort to promote a positive public image. In correlation research design, however, it is not possible to decide about the accurateness of any of these two interpretations. There is no correlation, or a much weaker correlation, between social desirability and the remaining Big Five dimensions – extraversion, neuroticism and openness to experience (Leary & Toner, 2015).

The aim of the present research was to identify the number and psychological content of the social desirability dimensions in the Polish population. This paper attempts to resolve the problem of the internal structure of social desirability, since the theoretical considerations, as well as empirical findings, proved to be ambiguous. As mentioned earlier, existing questionnaires are outdated and created in specific sociocultural and historical contexts so none of them could be used. Moreover, the theoretical basis of existing social desirability scales is ambiguous. Therefore a new questionnaire, using an inductive strategy, was developed.

In this strategy, thanks to the technique referred to as "matrix staring", the constructors usually manage to discover existing subdimensions and to devise an adequate label for each of them. As Burisch (1986) emphasizes, one starts with the collection of individual items and lets the data "speak for itself". In the inductive strategy, items are generated without any previous assumptions regarding the nature and the structure of construct of interest, so there is no need to rely on any definition underlying a scale carefully describing its structure, and interrelations among its subdimensions. We decided to generate culture-specific items, referring to content that is particularly important in the Polish population in the second decade of the 21st century.

The additional aim of the study was to examine the concurrent and discriminant validity of the final structure of social desirability with respect to trait scales from the Five Factor Model of personality: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. Certain aspects of personality may inherently involve efforts to present oneself in a particular way. Some personality characteristics can influence the person's concerns with social evaluation and self-presentation. Then personality variables can serve to examine the validity and correlate with the tendency to respond in a socially desirable way (Holden & Passey, 2009).

PARTICIPANTS AND PROCEDURE

PARTICIPANTS

In the first stage, a sample of 657 participants was investigated, comprising 25% males and 75% females. Moreover, 63% of the participants were 20-24 years old. In the second stage, a sample of 415 participants was investigated, comprising 61% males and 39% females. In the third stage, a sample of 1,709 participants was investigated, comprising 37% males and 63% females.

The study was approved by the ethics committee of the Faculty of Psychology at the University of Warsaw. Verbal informed consent was obtained from all participants. The participants were mainly university students, teachers, and medical doctors – all of whom were Polish citizens. The results were obtained in neutral conditions without any situational pressure to distort answers (like in job selection situations or instructions to "fake good" or "fake bad") where participants responded anonymously. Participants were volunteers and were not financially remunerated.

PROCEDURE

The study consisted of three stages, and three independent subject samples participated in each stage. We used this procedure to ensure that our results are independent of the specific research sample and can be generalized. The social desirability structure obtained in the first stage was verified in subsequent stages of the study on different samples. In subsequent stages of the study, the number of items was reduced and the previously obtained social desirability structure was verified in the above manner.

In order to determine the dimensions of social desirability, an inductive research strategy was used (Burisch, 1986). In this approach (also called "internal", "internal consistency" or "itemetric") to inventory construction, the number and nature of extracted scales result from data analysis. In order to uncover the nature of the assessed variable, a large item pool is formed by inventing new items (Burisch, 1986). The answers to these items are than collected. The sample should be as large as possible. The next step is an exploratory analysis of the accumulated data to reveal its underlying structure.

INSTRUMENT

Having decided to adopt the inductive strategy, the first step was to create an original pool of items, addressing a universe of behaviors, in which social desirability manifests itself. Diverse candidate items, reflecting alternative perspectives on social desirability,

were developed to improve construct elaboration. It was decided to involve non-specialists in proposing new items, as it was assumed that everyone implicitly knows what is socially desirable in everyday life. The aim was to avoid an *a priori* structure imposed by a pool of items sampled from existing measures.

For this purpose, 120 students, participating in the psychometrics course, were asked to create a few items, each referring to different behaviors, thoughts, and feelings that, according to them, are manifestations of social desirability. The students were first introduced to the general concept of social desirability and then instructed to create items on a wide variety of certain aspects of life (e.g., hygiene, lifestyle, healthy eating, polite behavior, minor offenses, daily habits, norms and social beliefs, etc.).

In order to limit the risk that students with their limited life experience, mainly due to age and not necessarily full self-sufficiency, would reduce the breadth of daily life content in generated items, the item pool was designed fifty-fifty by full-time students and extramural students. In Poland, extramural students usually work (sometimes for a long period of time, preceding studies), they are also older, on average, than full-time students, and they have already started their own families. This means that their life experience is much more diverse than that of full-time students. As a result, extramural students were able to generate items that would specify much more varied items than if they were created only by full-time students.

The items were generated independently, and this method yielded an initial pool of about 400 items. Such a procedure of developing new items is an inherent feature of the inductive strategy. Previous studies indicate that novice item writers (i.e., students without research experience) can design personality scales with worthwhile validities (Jackson, 1975). Nevertheless, it is very important to note that not every student-generated item was a linguistically and psychologically perfect item for our instrument. The authors of the study had to assess in the discussion the linguistic and psychological aspects of every student-generated item. In some cases, students' items were a kind of "food for thought" inspirations, rather than ready-made items. In these cases, we had to rewrite such items in a linguistically acceptable form. Moreover, we wanted to avoid items specific only for students' life, so every item of this type was converted into a form appropriate for non-students. We did not decide to draw any items from preexisting measures (although such a possibility is sometimes used in the inductive strategy) to avoid any previous theoretical and sociocultural influences.

The next step focused on elimination of redundant content items. Such items were removed and the initial pool of 156 items, all positively worded, was shaped into a balanced scale with four-point, Likert format answers (i.e., *strongly disagree, disagree, agree,*

strongly agree). The true-false response format of previous social desirability measures (Crowne & Marlowe, 1960) was replaced with an expanded one in the present study. An even number of answer options (i.e., without a neutral middle category) was chosen to avoid hesitation in answering, encourage participants to give a definite response, and minimize midpoint response style (Bishop, 1987; Kalton, Roberts, & Holt, 1980).

Each half of the items measured one direction of social desirability. The sum of appropriately reversed item scores is expected to be a measure of social desirability that is reasonably free from acquiescence, whereas the mean (or the sum) of the item scores, without reversing, is considered to be a measure of acquiescence (Ray, 1983). This large number of items was deliberately kept to avoid losing any of the manifestations of social desirability. The items were written in Polish. The question order was fixed across participants. This initial pool of items (Polish originals and their English translations) is accessible at the Open Science Framework repository: https://osf.io/emg64/?view_only=40e ad4da8ff84eb1a149055c2704baf4.

Principal component analysis (PCA) was chosen as the primary method of data analysis. The following indices were used to determine the number of components: the Kaiser criterion (eigenvalue > 1) (Kaiser, 1960), the scree test (Cattel, 1966) and the parallel analysis (Horn, 1965). This last procedure is a Monte Carlo simulation method known to be more efficient than the Kaiser criterion and the scree test in determining the number of components. It was decided to use the 95th quantile of simulated eigenvalues as the point of comparison when deciding whether the observed eigenvalues explain more variance than random data.

In order to eliminate the influence of acquiescence, Ten Berge (1999) proposed a method of partialling the mean out of the correlation matrix, in which the correlations of items with individual averages are treated as acquiescence factor loadings. In this method, a factor analysis of the questionnaires starts with extracting the acquiescence factor from the correlation matrix. Finally, PCA is performed on this residual matrix. This procedure was applied in each stage of the present study, preceded by classical PCA.

The Ten Berge (1999) method can be applied only to a balanced pool of items. Therefore, elimination of items was based on the factor loadings and sustainability of the balanced scale. This was achieved by separately selecting the same number of the highest loadings from two groups: (1) a group of reverse-scored items and (2) a group of straightforward-scored items.

However, the Ten Berge (1999) method results in a singular correlation or covariance residual matrix. Only PCA, not a factor analysis, can be applied to such a matrix. Thus, the Ten Berge (1999) method was followed by PCA despite its limitations, of which the authors were aware. Specifically, the PCA

solution, compared to the factor analysis solution, usually seems to explain more variance. Since it is well known that in PCA, the components are determined so as to account for maximum variance of all observed variables, this method is appropriate for identification of the main sources of variance; as Reyment and Jöreskog (1996) pointed out, "principal component analysis is variance-oriented, whereas factor analysis is correlation-oriented". It has to be mentioned that PCA is not "selective" in the variance it analyzes, which is a plausible limitation of this method in the identification of psychological constructs, as PCA refers to all the variance.

Additionally, in the third stage, 1240 participants filled in the Polish adaptation (Strus, Cieciuch, & Rowiński, 1994) of the 50-item set of the International Personality Items Pool Big Five Factor Markers questionnaire to measure extraversion, agreeableness, conscientiousness, stability (equivalent of neuroticism), and intellect (equivalent of openness) (Goldberg, 1992). The reliability and validity of the Polish version were tested on a large sample, showing high internal consistency (α ranging from .73 to .91), adequate factor structure and correlations (from .47 to .70) with the scales from the other Big Five measures (Strus et al., 2014).

At each stage, internal consistency of obtained dimensions was estimated. Cronbach's alphas (and 95% confidence interval) were computed for items constituting each of the dimensions.

RESULTS

STAGE 1

In the first stage, the entire data set of 156 items was analyzed. The scree plot of PCA is presented at the top of Figure 1. It is clear that two components emerged.

Together, the two obtained components explain about 24% of the variance, with eigenvalues of 16.00 and 8.98. The first component was recognized as representing a bipolar continuum, with one extreme corresponding to the denial of negative features (i.e., negative item loadings) and the opposite extreme assigning positive properties (i.e., positive item loadings). Quite considerable difficulties emerged in the interpretation of the second component. This dimension is loaded by items with different content, comprising both desirable and undesirable mental characteristics. All of these loadings are positive, suggesting that the discussing component reflects acquiescence.

A considerable number of components have an eigenvalue exceeding the value of one (Kaiser criterion). Altogether there are 20 such components. Their eigenvalues were similar to one another and at the same time they were much smaller than the first or second component (visual inspection of the scree-plot clearly demonstrates this). For better understanding of this problem, a parallel analysis was performed. Four of these 20 components have an ei-
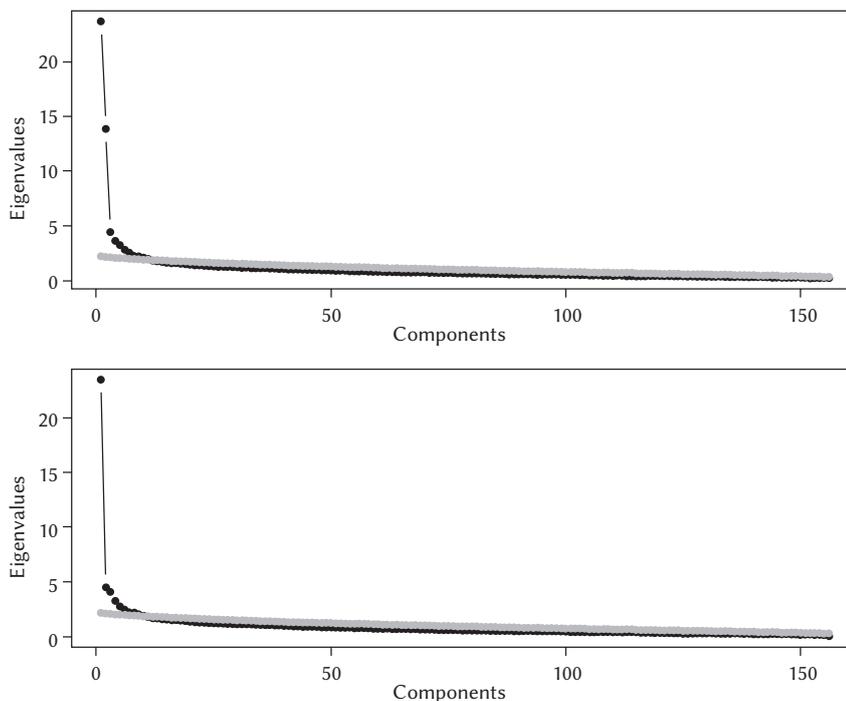
Figure 1. Scree plot with parallel analysis results of the principal component analysis (PCA) of 156 items before partialling the mean (top) and after partialling the mean (bottom).

genvalue exceeding the value suggested by the parallel analysis. It was very difficult to indicate what psychological content represents these four dimensions, because factor loadings of most items are low. This suggests the existence of white noise in the analysis data as well as certain poorly represented psychological contents, which are difficult to interpret.

In order to better interpret the meaning of the components, the correlation between component scores and the acquiescence score was calculated and presented in Table 1. The second component, in particular, is strongly associated with the acquiescence score and, as such, reflects response style, which supports the above suggestion regarding this component. To avoid this and remove the influence of response style, the procedure of Ten Berge (1999) was applied, and PCA was performed once again. This time, the dimensionality was different. It is presented at the bottom of Figure 1.

The first dimension explains about 16.52% of the variance, while the second explains only about 3.14%. The corresponding eigenvalues are 15.17 and 3.03, respectively. This time, the second component almost disappeared. As before, the eigenvalues of 19 components exceeded the Kaiser criterion. However, they were relatively small. Four of these 19 components were characterized by eigenvalues exceeding the values determined by parallel analysis. However, psychological interpretation of these component was impossible due to low factor loadings.

The correlations between the scores of the two solutions (see Table 2) illustrate that the first component is almost identical in both solutions, but the second component represents a somewhat different meaning.

Table 1

Relationship between principal components and acquiescence score

Principal components obtained			
Before partialling the mean		After partialling the mean	
Component 1	Component 2	Component 1	Component 2
First stage			
.16**	−.95**	.08*	−.34**
Second stage			
.53**	.84**	.47**	.05
Third stage			
−.22**	−.91**	−.10	.14

Note. $^*p < .05$, $^{**}p < .001$.

Table 2

Correlation between scores before and after partialling the mean

Principal components obtained		
Before partialling the mean	After partialling the mean	
	Component 1	Component 2
First stage		
Component 1	.99**	−.04
Component 2	.10**	.45**
Second stage		
Component 1	.99**	−.04
Component 2	.06	.03
Third stage		
Component 1	.99**	−.11
Component 2	−.01**	−.43

Note. $^{**}p < .001$.

Table 3

Reliability indices

		First stage		Second stage		Third stage	
		Comp. 1	Comp. 2	Comp. 1	Comp. 2	Comp. 1	Comp. 2
Before partialling the mean	Number of items	98	58	40	10	8	2
	Cronbach's α	.96	.92	.95	.78	.74	.59
	Lower bound	.95	.91	.94	.75	.71	.54
	Upper bound	.96	.93	.96	.81	.77	.63
After partialling the mean	Number of items	144	12	47	3	9	1[†]
	Cronbach's α	.96	.46	.94	.54	.79	
	Lower bound	.96	.40	.94	.48	.77	
	Upper bound	.97	.52	.95	.60	.81	

Note. [†]Computation of the internal consistency index is impossible for one item.

Reliability indices (Cronbach's alphas) for all components obtained before and after partialling the mean are presented in Table 3. The reliability of the first component is very high, whereas the reliability of the second component diminished considerably after partialling the mean.

Summing up the results of the first stage of the study, it can be stated that 2 dimensions were detected: social desirability and acquiescence.

STAGE 2

In order to verify this structure, the second stage of the study was performed. This stage was performed using 50 selected items with the greatest loadings on both the dimensions obtained in the first stage before partialling the mean. The procedure was similar to that applied in the first stage. Two PCAs were performed, before and after partialling the mean. The relevant scree plots of these analyses are depicted in Figure 2.

Visual inspection of these graphs leads to the conclusion that, after partialling the mean, the two-factorial solutions that emerged before partialling the mean disappeared. The eigenvalues of the first two components of the first analysis – that is, before partialling the mean – were 9.59 (explained variance of 29.71%) and 2.12 (explained variance of 8.28%). The corresponding eigenvalues after partialling the mean were 6.68 (explained variance of 30.7%) and 1.15 (explained variance of 3.94%). There are eleven eigenvalues in the first analysis and ten eigenvalues in the second analysis that are greater than one. This time, however, the
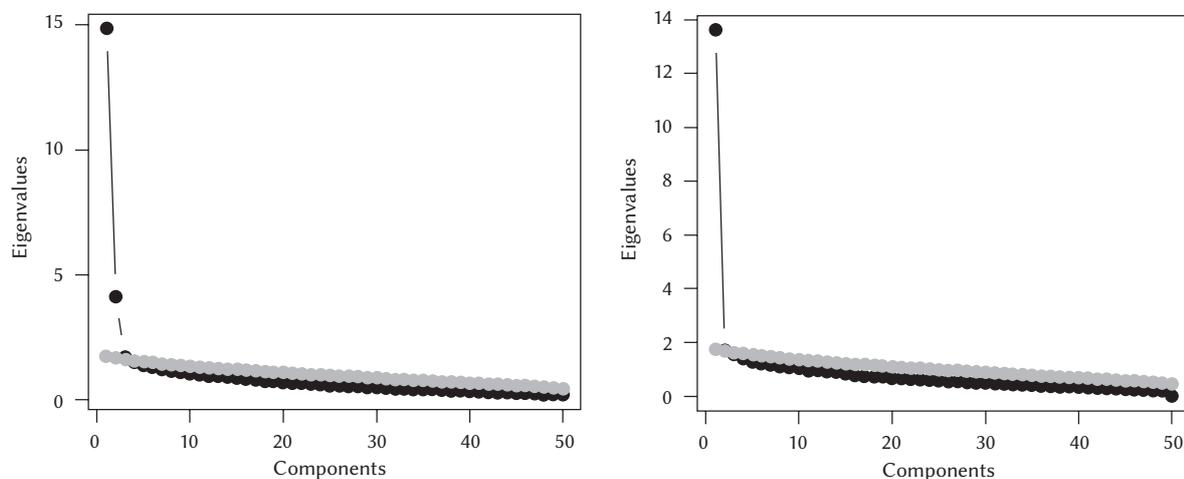
Figure 2. Scree plot with parallel analysis results of the principal component analysis (PCA) of 50 items before partialling the mean (left) and after partialling the mean (right).

parallel analysis clearly indicated that there are two dimensions (before partialling the mean) or one dimension (after partialling the mean) since eigenvalues of all these remaining components do not exceed the values determined in parallel analysis.

The correlation pattern, presented in Table 1, suggests that the second component of the first analysis reflected the response style above all. After partialling the mean, in the second analysis, this component was removed completely. However, the correlation of the first component with the acquiescence score almost did not change between the two analyses.

The data presented in Table 2 indicate that the first component is almost identical in both analyses (before and after partialling the mean), while the second component is completely different in each analysis.

Reliability indices are presented in Table 3. Reliability of the first component is high before partialling the mean and practically does not change after partialling the mean. The reliability of the second component is lower and becomes unacceptably low after partialling the mean.

The analyses of the second stage lead to the conclusion that social desirability is a one-dimensional construct and an optional second dimension represents acquiescence.

STAGE 3

Maintaining a large number of items obtained in stage 2 was unjustified. In this situation, the item pool was reduced to ten and selection was based on absolute value loading obtained in stage two. (Selected items, both in the Polish original version and their English translations, are available in the Open Science Framework repository: https://osf.io/emg64/?view_only=40

ead4da8ff84eb1a149055c2704baf4). These items were tested on a new subject sample, and a similar method of data analysis to that used in stages one and two was applied. In this case, one dimension emerged in PCA, as shown at the top of Figure 3.

The eigenvalue of this dimension is 1.93. The remaining eigenvalues are considerably smaller and rather similar. The variance explained by the first component is 33.07%. The scree plot, after partialling the mean, is presented at the bottom of Figure 3. Clearly this one dimension emerged, characterized by a similar eigenvalue of 1.93 (explained variance of 34.86%). In both cases, parallel analysis clearly indicated the existence of one dimension.

As shown in Table 1, the first component is still correlated with acquiescence to some extent, even after partialling the mean. Interestingly, before partialling the mean, a very high correlation between the second component and acquiescence is observed.

The first components obtained in both analyses (before and after partialling the mean) are highly correlated (see Table 2). There were also strong correlations among the second components.

At this stage of study, reliability of the first component was somewhat lower (Table 3). The reliability of the second component was not assessed after partialling the mean because only one item's loading was higher on the second component than on the first. It is of course impossible to compute an internal consistency index for one item.

An additional analysis carried out at this stage of the study was the correlation of obtained dimensions with personality traits. The obtained results are presented in Table 4. Almost all correlations are statistically significant due to the large research sample. The relatively high correlations of the first component with agreeableness and conscientiousness are
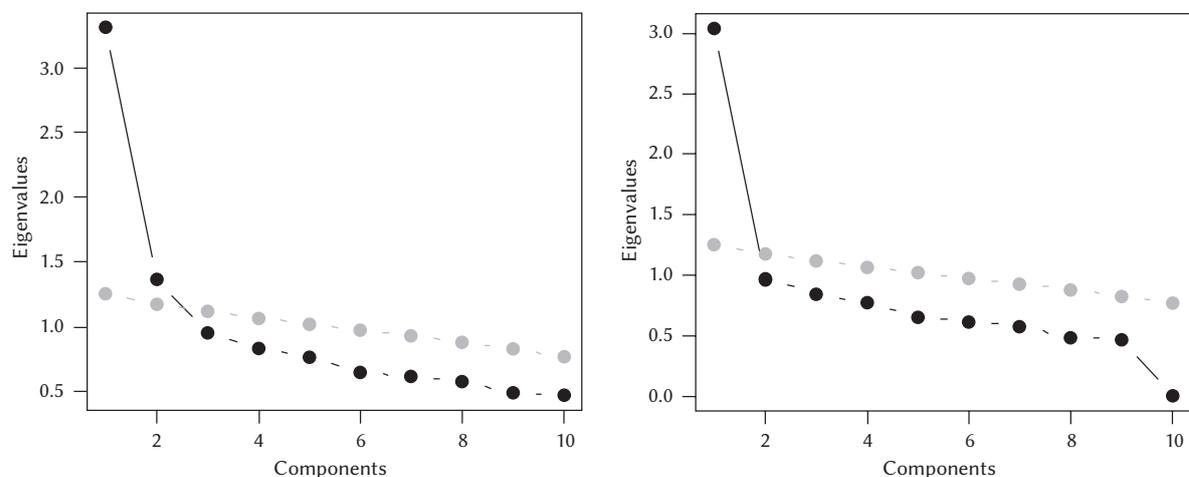
Figure 3. Scree plot with parallel analysis results of the principal component analysis (PCA) of 10 items before partialling the mean (left) and after partialling the mean (right).

Table 4

Correlations of Big Five dimensions and third stage dimensions

| | Before partialling the mean | | After partialling the mean | |
	Component 1	Component 2	Component 1	Component 2
Extraversion	.35**	−.30**	.36**	.01
Agreeableness	.53**	−.32**	.54**	.00
Conscientiousness	.60**	−.25**	.60**	.16**
Stability	.36**	−.21**	.36**	.12**
Intellect	.24**	−.35**	.26**	.04

Note. **p < .001.

noticeable. The magnitude of these correlations does not change after partialling the mean.

In the third stage of the study, two dimensions emerged, social desirability and acquiescence, but after the procedure of partialling the mean, we obtained one dimension that is completely free of acquiescence. In both analyses, social desirability is strongly related to agreeableness and conscientiousness.

DISCUSSION

In the first stage of the study (156 items), an orthogonal two-factor solution was obtained. These two dimensions refer to social desirability and a preference for acquiescence.

The first impression of our results might be that the obtained single dimension of social desirability reflects its internally undifferentiated character, so we should regard this variable as the global characteristic describing people's tendency to present themselves in a positive light in different aspects of their life. But a careful examination of the content of the items comprising the final version of our measure leads to a somewhat different conclusion. It seems that the content of items of our instrument pertains to the highest extent to interpersonal relations. Our final instrument focuses mainly on the behaviors and beliefs that people display in social situations and relations with others, so the closest term for our measure would be "fairness", but it shall not be identified as entirely overlapping with agreeableness for the Big Five model. This instrument measuring a specific facet of social desirability can be used as such by other researchers.

In addition to content analysis of the measure that emerged as a result of this study, it is worth considering its high correlations with agreeableness and conscientiousness. The social desirability dimension obtained in our study resembles self-deception more than impression management, since agreeableness and conscientiousness show a stronger correlation with self-deception than impression management in the previous study by Paulhus (2002), although impression management is also correlated with these two personality dimensions but to a much lesser extent. Thus the dimension revealed in our study is not a "clear" measure of self-deception.

However, taking jointly into account these correlations and the content of the obtained dimension, it can be assumed that our measure concerns self-deception in the field of social relations. Therefore, it concerns the internal conviction rather than the consciously manipulated image of the respondents about the positive nature of their relationships with other people.

It could be interpreted that social relations are the main domain in which people want to present themselves in a favorable view. Other aspects of the positive personal image that have been included in the initial item pool (e.g., hygiene, fitness lifestyle, polite behavior, breaking minor social norms, and so on) appeared not so important in the Polish population. Probably the specific nature of Polish culture in the aspect of social desirability manifests itself mainly in social relations that are crucial for all human beings.

Such a structure of the presented measure that is referred to as the "Social Desirability Questionnaire" does not correspond to the two-dimensional structure proposed by Paulhus (1984, 2002) or Ramanaiah et al. (1977). The result from Ramanaiah et al. (1977) proved to be attributable to the method variance caused by the direction of keying in Marlowe-Crowne's scale (Ramanaiah & Martin, 1980), indicating that the two-dimensional structure was strongly influenced by the response style. Our failure to replicate the two-factor structure of social desirability is therefore congruent with Ramanaiah and Martin's (1980) findings.

The orthogonal two-factor solution, derived from the first stage of the study, was not fully replicated in the second stage where 50 selected items with the greatest factor loadings on both dimensions were used. The situation became clearer in the third

phase of the study. In this case, the second dimension proved to be generally difficult to distinguish from the first one.

All combined analyses of the three stages of this study led to the conclusion that the inductive approach has revealed one dimension of social desirability. The new items developed by non-experts, reflecting what people consider socially desirable or undesirable and what is closer to everyday experience, have created a single dimension. However, one possible study limitation is that choice of the inductive strategy method may lead to overly narrowed content when very rare content is rejected in the factor analysis. We cannot completely exclude the possibility that very rare manifestations of social desirability are omitted in our final factor solution.

Interestingly, the inductive approach used in this study has led to a one-way factor solution, in contrast to the studies using theoretically oriented tools. This may mean that existing psychological theories assume greater diversity of subjects in terms of positive self-presentation than the actual diversity in reality. It is also possible, however, that existing measures provide examinees with the opportunity to present themselves in more differentiated ways, requiring more than one dimension. This alternative can be an issue for further studies and analysiThe acquiescence dimension emerged in the first phase, especially when the largest number of items in the measure was used. In the second phase, where a smaller number of items in the measure was used, this factor was disclosed to a lesser extent. During the third stage, when the smallest number of items in the measure was used, acquiescence did not exist at all. The number of items is, therefore, of great importance for the emergence of acquiescence. According to one of the anonymous reviewers of the first version of this paper, satisficing rather than acquiescence is a better description of this phenomenon. Indeed, a large number of items can cause fatigue and a consequent reduction in the cognitive resources necessary to provide reliable answers. As Krosnick (1999) pointed out, in this situation two processes can operate. One of them affects the retrieval and integration of information required to formulate the answer, and the other affects response expression. In the first case, the process of generating reasons to disagree with the item's assertion usually needs more cognitive capacities because of the confirmatory bias that typically leads most people to seek reasons to agree rather than disagree. The person would thus be inclined to agree. In the second case, people do not evaluate the item's assertion at all, but simply choose to agree because it seems like the commanded and polite action to take (Krosnick, 1999).

A one-dimensional structure was detected in all three stages of the study after partialling the mean. Therefore, elimination of acquiescence has left a "clear" dimension of social desirability only. Interestingly, however, even this "clear" dimension correlates to some extent with the acquiescence score. This correlation exists in spite of the fact that the used scale is balanced. This result suggests that even items that primarily measure social desirability can be impacted to some extent by acquiescence. Probably, the content of socially desirable items leads at the same time to greater acquiescence, and this is a property of language that is difficult to eliminate. Moreover, some people may acquiesce for reasons of social desirability (Knowles & Condon, 1999). Simply, people may consider saying "no" as something undesirable. It means that social desirability and acquiescence are not fully different from a theoretical point of view, as well as in the light of our empirical data.

However, one social desirability dimension of our study may be confused with the tendency to acquiescence. In order to avoid this, the procedure of partialling out the mean has to be applied. This limits the utility of the measure to scientific research, because in individual diagnosis, the abovementioned procedure would be too difficult to apply.

The lack of confirmation of the two-dimensional social desirability structure is not unusual. Paulhus (1986), the great proponent of the two-factor structure of social desirability, also encountered difficulties in proving empirically a clear structure of social desirability and obtained some results wherein the two conceptually different responding styles – deceptive self-denial and impression management – created one factor. Ventimiglia and MacDonald (2012), however, identified a two-factor model of social desirability, which also found evidence for a one-dimensional model in their Marlowe-Crowne Social Desirability Scale analysis (Crowne & Marlowe, 1960). The explanation for this one-factor result would be the fact that the need for self-presentation diminishes in anonymously surveyed participants and deceptive self-denial and impression management refer to similar characteristics.

The main advantage of the measure described in this paper is its relevance for contemporary Polish socio-cultural conditions. Popular tools that help measure social desirability, such as Edwards' (1957) Social Desirability Scale or Marlowe-Crowne's Social Desirability Scale (Crowne & Marlowe, 1960), were developed over half a century ago. The Balanced Inventory of Desirable Responding, by Paulhus (1984), is newer; still, it was developed over 30 years ago. Taking into account the socio-cultural changes associated with the passage of time, it should be assumed that these measures are no longer adequate measures of social desirability. Social desirability is strongly linked to social norms (Malham & Saucier, 2016) that change over time, so items of old inventories were suitable indicators of what was socially desirable many years ago.

It should be noted, however, that the alphas, which are generally high, may be the result of the large numbers of items used, especially in the first and second stage of the study. Moreover, the internal consistency of the 10-item scale in stage 3 is quite low, which could be a limitation of using this scale in practice.

It seems that the question of the social desirability structure remains open. The question of its intercultural universality is especially interesting. Our study, in contrast to most of the studies cited – which were conducted in North America – was carried out in Poland. Perhaps our results capture some specific aspects of Polish culture. This would manifest itself in the existence of a single dimension of social desirability reflecting mainly social relations. In addition, an important dimension would be acquiescence. But whether the one-dimension structure of social desirability is applicable in other cultures is questionable. This issue requires further verification. In order to generalize the obtained results to cultures other than Polish, it is necessary to conduct similar studies using a similar research scheme in other cultures. In such studies, the inductive method would be used and the procedure would start with culture-specific items.

CONCLUSIONS

The obtained results lead to the conclusion that social desirability is a unidimensional construct and it can be confounded by acquiescence, especially in a longer questionnaire. Every item of our questionnaire reflects both social desirability and acquiescence, albeit to different extents. However, the procedure of partialling the mean out of the correlation matrix eliminates acquiescence.

FUNDING

References

Bishop, G. F. (1987). Experiments with the middle response alternative in survey questions. *Public Opinion Quarterly, 51*, 220–232. https://doi.org/10.1086/269030

Borkenau, P., & Ostendorf, F. (1992). Social desirability scales as moderator and suppressor variables. *European Journal of Personality, 6*, 199–214. https://doi.org/10.1002/per.2410060303

Burisch, M. (1986). Methods of personality inventory development – a comparative analysis. In A. An-
gleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires* (pp. 109–120). Berlin, Heidelberg: Springer

Cattel, R. B. (1966). The scree test for the numbers of factors. *Multivariate Behavioral Research, 1*, 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Cervellione, K. L., Lee, Y. S., & Bonanno, G. A. (2009). Rasch modeling of the self-deception scale of the balanced inventory of desirable responding. *Educational and Psychological Measurement, 69*, 438–458. https://doi.org/10.1177/0013164408322020

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–354. https://doi.org/10.1037/h0047358

Crowne, D. P., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence.* New York, NY: Wiley.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* Fort Worth, TX: Dryden Press.

Espinosa, A. D. C., & van de Vijver, F. J. (2014). An Indigenous Social Desirability Scale. *Measurement & Evaluation in Counseling & Development, 47*, 199–214. https://doi.org/10.1177/0748175614522267

Ferrando, P. J., Lorenzo-Seva, U., & Chico, E. (2009). A general factor-analytic procedure for assessing response bias in questionnaire measures. *Structural Equation Modeling, 16*, 364–381. https://doi.org/10.1080/10705510902751374

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences, 7*, 385–400. https://doi.org/10.1016/0191-8869(86)90014-0

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*, 26–42. https://doi.org/10.1037/1040-3590.4.1.26

Greenwald, H. J., & Clausen, J. D. (1970). Test of relationship between yeasaying and social desirability. *Psychological Reports, 27*, 139–141. https://doi.org/10.2466/pr0.1970.27.1.139

He, J., & van de Vijver, F. J. R. (2013). A general response style factor: Evidence from a multi-ethnic study in the Netherlands. *Personality and Individual Differences, 55*, 794–800. https://doi.org/10.1016/j.paid.2013.06.017

He, J., van de Vijver, F. J. R., Espinosa, A. D., Abubakar, A., Dimitrova, R., Adams, B. G., ...Villieux, A. (2015). Socially desirable responding: Enhancement and denial in 20 countries. *Cross-Cultural Research, 49*, 227–249. https://doi.org/10.1177/1069397114552781

Helmes, E., & Holden, R. R. (2003). The construct of social desirability: One or two dimensions? *Personality and Individual Differences, 34*, 1015–1023. https://doi.org/10.1016/S0191-8869(02)00086-7

Hofstee, W. K. B., Ten Berge, J. M. F., & Hendricks, A. A. J. (1998). How to score questionnaires. *Personality*

and Individual Differences, 25, 897–909. https://doi.org/10.1016/S0191-8869(98)00086-5

Holden, R. R., & Fekken, G. C. (1989). Three common social desirability scales: Friends, acquaintances, or strangers? *Journal of Research in Personality, 23*, 180–191. https://doi.org/10.1016/0092-6566(89)90022-6

Holden, R. R., & Passey, J. (2009). Social desirability. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 441–454). New York, NY: Guilford Press.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179–185. https://doi.org/10.1007/B.F.02289447

Jackson, D. N. (1975). The relative validity of scales prepared by naive item writers and those based on empirical methods of personality scale construction1. *Educational and Psychological Measurement, 35*, 361–370. https://doi.org/10.1177/001316447503500214

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141–151. https://doi.org/10.1177/001316446002000116

Kalton, G., Roberts, J., & Holt, D. (1980). The effects of offering a middle response option with opinion questions. *Statistician, 29*, 65–78. https://doi.org/10.2307/2987495

Knowles, E. S., & Condon, C. A. (1999). Why people say "yes": a dual-process theory of acquiescence. *Journal of Personality and Social Psychology, 77*, 379–386. https://doi.org/10.1037/0022-3514.77.2.379.

Kroner, D. G., & Weekes, J. R. (1996). The balanced inventory of desirable responding: Factor structure, reliability, and validity with an offender sample. *Personality and Individual Differences, 21*, 323–333. https://doi.org/10.1016/0191-8869(96)00079-7

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50*, 537–567. https://doi.org/10.1146/annurev.psych.50.1.537

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity, 47*, 2025–2047. https://doi.org/10.1007/s11135-011-9640-9

Lanyon, R. I., & Carle, A. C. (2007). Internal and external validity of scores on the balanced inventory of desirable responding and the Paulhus deception scales. *Educational and Psychological Measurement, 67*, 859–876. https://doi.org/10.1177/0013164406299104

Leary, M. R., & Toner, K. (2015). Self-processes in the construction and maintenance of personality. In M. Mikulincer & P. R. Shaver (Eds.), *APA handbook of personality and social psychology* (Vol. 4, pp. 447–467). Washington, DC: APA Publishing.

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. https://doi.org/10.1177/0013164404267285

Li, F., & Li, Y. (2008). The Balanced Inventory of Desirable Responding (BIDR): a factor analysis. *Psychological Reports, 103*, 727–731. https://doi.org/10.2466/pr0.103.3.727–731

Li, F., Li, Y., & Wang Y. (2015). *Socially desirable responding in Chinese university students: Denial and enhancement? Psychological Reports, 116*, 409–421. https://doi.org/10.2466/03.07.pr0.116k18w8

Loo, R., & Loewen, P. (2004). Confirmatory factor analyses of scores from full and short versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Applied Social Psychology, 34*, 2343–2352. https://doi.org/10.1111/j.1559-1816.2004.tb01980.x

Loo, R., & Thorpe, K. (2000). Confirmatory factor analyses of the full and short versions of the Marlowe-Crowne Social Desirability Scale. *The Journal of Social Psychology, 140*, 628–635. https://doi.org/10.1080/00224540009600503

Malham, P. B., & Saucier, G. (2016). The conceptual link between social desirability and cultural normativity. *International Journal of Psychology, 51*, 474–480. https://doi.org/10.1002/ijop.12261

Marín, G., Triandis, H. C., Kashima, Y., & Betaneourt, H. (1983). Ethnic affirmation versus social desirability: Explaining discrepancies in bilinguals' responses to a questionnaire. *Journal of Cross-Cultural Psychology, 14*, 173–186. https://doi.org/10.1177/0022002183014002003

Martin, J. (1964). Acquiescence – measurement and theory. *British Journal of Clinical Psychology, 3*, 216–225. https://doi.org/10.1111/j.2044-8260.1964.tb00430.x

Moral de la Rubia, J., García-Cadena, C. H., & Antona Casas, C. J. (2012). Traducción y validación del inventario balanceado de deseabilidad social al responder en una muestra probabilística de estudiantes universitarios mexicanos [Translation and validation of the balanced inventory of desirable responding in a probability sample of Mexican university students]. *Revista de Psicología GEPU, 3*, 54–72.

Musch, J., Brockhaus, R., & Bröder, A. (2002). Ein inventar zur erfassung von zwei faktoren sozialer erwünschtheit [An inventory for the assessment of two factors of social desirability]. *Diagnostica, 48*, 121–129. https://doi.org/10.1026//0012-1924.48.3.121

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*, 660–679. https://doi.org/10.1037//0021-9010.81.6.660

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

Paulhus, D. L. (1986). Self-deception and impression management in tests responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via*

questionnaire (pp. 143–165). Berlin, Heidelberg: Springer.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Lawrence Erlbaum Associates.

Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology, 60*, 307–317. https://doi.org/10.1037/0022-3514.60.2.307

Pearson, P. R., & Francis, L. J. (1989). The dual nature of the Eysenckian lie scales: Are religious adolescents more truthful? *Personality and Individual Differences, 10*, 1041–1048. https://doi.org/10.1016/0191-8869(89)90255-9

Ramanaiah, N. V., Schill, T., & Leung, L. S. (1977). A test of the hypothesis about the two-dimensional nature of the Marlowe-Crowne Social Desirability Scale. *Journal of Research in Personality, 11,* 251–259. https://doi.org/10.1016/0092-6566(77)90022-8

Ramanaiah, N. V., & Martin, H. J. (1980). On the two-dimensional nature of the Marlowe-Crowne Social Desirability Scale. *Journal of Personality Assessment, 44*, 507–514. https://doi.org/10.1207/s15327752jpa4405_11

Ray, J. J. (1983). Reviving the problem of acquiescent response bias. *Journal of Social Psychology, 121*, 81–96. https://doi.org/10.1080/00224545.1983.9924470

Reyment, R. A., & Jöreskog, K. G. (1996). *Applied factor analysis in the natural sciences.* New York: Cambridge University Press.

Schwartz, S. H. (2011). Studying values: Personal adventure, future directions. *Journal of Cross-Cultural Psychology, 42*, 307–319. https://doi.org/10.1177/0022022110396925

Shultz, K. S., & Chávez, D. V. (1994). The reliability and factor structure of a social desirability scale in English and in Spanish. *Educational and Psychological Measurement, 54*, 935–940. https://doi.org/10.1177/0013164494054004009

Stöber, J. (1999). Die Soziale-Erwünschtheits-Skala-17 (SES-17): Befunde zu Reliabilität und Validität [The Social Desirability Scale-17 (SDS-17): Development and first results on reliability and validity]. *Diagnostica, 45*, 173–177. https://doi.org/10.1026//0012-1924.45.4.173

Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment, 17*, 222–232. https://doi.org/10.1027//1015-5759.17.3.222

Stricker, L. J. (1963). Acquiescence and social desirability response styles, item characteristics, and conformity. *Psychological Reports, 12*, 319–341. https://doi.org/10.2466/pr0.1963.12.2.319

Strus, W., Cieciuch, J., & Rowiński, T. (2014). Polska adaptacja kwestionariusza IPIP-BFM-50 do pomiaru pięciu cech osobowości w ujęciu leksykalnym [Polish adaptation of IPIP-BFM-50 measuring five personality traits in a lexical approach]. *Roczniki Psychologiczne, 17*, 327–366.

Ten Berge, J. M. F. (1999). A legitimate case of component analysis of ipsative measures, and partialling the mean as an alternative to ipsatization. *Multivariate Behavioral Research, 34*, 89–102. https://doi.org/10.1207/s15327906mbr3401_4

Ventimiglia, M., & MacDonald, D. A. (2012). An examination of the factorial dimensionality of the Marlowe Crowne Social Desirability Scale. *Personality and Individual Differences, 52*, 487–491. https://doi.org/10.1016/j.paid.2011.11.016