Reducing the Social Desirability Bias of Self-Reported Value Orientation By Measuring Values Via a Picture-Based Scale

Andrea Groeppel-Klein, Saarland University, Germany
Sarah Kobel, Saarland University, Germany

Many consumers behave differently from the way they would be expected to behave according to their value orientation. One reason may be the social desirability bias (SDB). Consumers over- or under-report their real attitudes in ways they believe will be viewed favorably by others. Pictorial scales can reduce SDB.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/1023677/volumes/v45/NA-45

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/.
Reduction of the Social Desirability Bias of Self-Reported Value Orientation by Measuring Values via a Picture-Based Scale
Andrea Groeppel-Klein, Saarland University, Germany
Sarah Kobel, Saarland University, Germany

EXTENDED ABSTRACT

For decades, values have been said to be “one of the most powerful explanations of, and influence on, consumer behavior” (Clawson and Vinson 1978) and can guide people’s actions (Schwartz 1992). According to Inglehart (1971, 2008), an intergenerational value shift should take place in society: Materialistic values (e.g., safety or wealth) should be replaced by post-materialistic values (e.g., quality of life). However, Inglehart’s predicted shift of values has been criticized (Brym et al. 2016; Tranter and Western 2010). Even though the importance of post-materialistic values has increased, materialistic values are still important (Segev, Shoham, and Gavish 2015) owing to political or economic incidents. Inglehart and Norris (2016) explain this phenomenon as a backlash against cultural change. Another reason for Inglehart’s misprediction may be the social desirability bias (SDB) of surveys, which can lead to underreporting of materialistic values. The SDB is considered either to reduce the validity of measured insights into consumers’ value orientations (King and Bruner 2000), or as an additional source of value reflection (Fisher and Katz 2000), indicating that faking allegedly esteemed values can hint at the relevance of a topic. In our article, we focus on the first perspective by developing and validating a new measurement tool.

Empirical research within the last five decades has used different value measurement approaches. What all the methods have in common is their pure verbal measurement. The Rokeach Value Survey (RVS) (1973) seeks to obtain insights into people’s value hierarchy by asking them to rank 36 values. However, this method is criticized for being ipsative, for the strict value rank-order and the item choice itself (Braithwaite and Law 1985). Schwartz (1992) developed the theory of basic human values, suggesting 10 basic values as trans-situational goals, varying in importance and serving as guiding principles in a person’s life. He developed a survey (SVS) inviting participants to rate a broad range of single values (different varieties of each value) as guiding principles, based on a complex nine-point scale, ranging from -1 (opposed to my values) to 7 (very important). Hence, a high level of abstract thought is necessary for using the SVS successfully (Schwartz et al. 2001). Therefore, the survey might be vulnerable for SDB. For similar reasons, Kahle’s (1983) more parsimonious “List of Values” seems to be affected by SDB as it also uses rating scales (Fisher and Katz 2000; Lee, Soutar, and Louviere 2007). Verbal measurement techniques, especially if they require abstract thinking skills (Torelli and Kaikati 2009) may induce rationalization processes in respondents and therefore diminish the possibility of obtaining intuitive and spontaneous answers. Individuals tend to reflect well on themselves (Randall and Fernandes 1991) and are therefore likely to give socially desirable answers (Zerbe and Paulhus 1987). Many verbal measures, especially if consumers only have to change the rating scale figure, face this problem. In the context of value measurement, studies (e.g. Fisher and Katz 2000) indicate that lots of values are affected by this problem. The degree to which value self-reports are affected by SDB depends on the value’s potential to present oneself in a favorable light (Fisher and Katz 2000). Over-reporting is likely to occur if values seem socially acceptable, whereas under-reporting is expected for socially undesirable values (Needle and Cronley 2004; Roxas and Lindsay 2012). In today’s research, universalism values, for example, play an important role (Doran 2009; Vermeir and Verbeke 2006; Yamoah et al. 2016) as ethical consumption is requested as a social standard (Auger and Devinney 2007), thus increasing the risk of SDB (over-reporting). A materialistic orientation on the other hand faces the problem of being underreported, as negative social consequences are expected (Mick 1996). For many years, researchers have been interested in the scale of SDB occurrences and who over- and underreports (De Jong, Fox, and Steenkamp 2015). The “pure” indirect questioning technique (Schwartz et al.’s (2001) PVC approach, in which consumers indicate to what extent the description of another person is congruent with their own personality is a mixture of direct and indirect technique) uses questions like, “What would the typical other do?” (Fischer and Tellis 1998) and is problematic if personal values and hierarchies have to be discovered. Many researchers therefore use the Marlowe-Crowne Social Desirability Scale (or modifications) to detect SDB as a personality variable and ascertain whether or not a respondent is generally concerned with social approval and conforming to societal conventions (which may affect the validity of the survey). However, clever consumers might see through this verbal scale or misconstrue it so that inconsistencies appear (King and Bruner 2000).

Nevertheless, consumer research has a strong interest in obtaining insights into consumers’ real value orientations because values can guide behavior. Manufacturers and retailers wish to adjust marketing strategies according to the values of their target groups. Enlarging the share of fair trade or regional products, for example, may be a consequence of consumers’ increased interest in universalistic values, whereas more security officers in shopping centers may be a result of growing safety needs. However, this requires a valid method to measure values. Pictorial scales are expected to reduce SDB by facilitating intuitive answers without cognitive filtering and evaluation of the social effects (Groeppel-Klein et al. 2010). Furthermore, pictorial scales could increase enjoyment and response rates.

Development of a Picture-Based Scale and Validation Studies

To map out a pictorial scale, different steps are necessary. We paid particular attention to careful validation (Groeppel-Klein et al. 2010), and to the scale’s ability to reduce SDB in a self-reporting questionnaire.

Step 1: Identification of values and value varieties

Both materialist (Segev et al. 2015) and post-materialist (Ladhari and Tchetgna 2015) values play a part in contemporary marketing research and practice. Beyond the relevance of materialism (especially power) and post-materialism (especially universalism) values, such as stimulation or hedonism might be important, especially in consumer behavior. Schwartz’ (1992) 10 basic human values served as a theoretical foundation. However, some of these values are abstract and complex. Schwartz et al. (2012) therefore deliver a detailed description of these abstract dimensions. Our goal in step 1 was to discover subtle facets of values (rather than broad supercategories) that are more apperceptive and can specifically be used in a consumption context. Results of a thorough literature review and the descriptions of Schwartz et al. (2012) served as a basis for qualitative focus group
discussions (one at Saarland University, three in grocery stores in Germany; \( n = 20 \) (five participants per group); \( M_{\text{age}} = 50.2 \) years, SD = 25.7). According to the results of the focus groups, we identified 26 subtle value facets that might be relevant for consumer behavior.

**Step 2: Selection of suitable pictures for value visualization**

Five coders checked various commercial image databases to find pictures that represent the values identified in step 1 (screening of 10,000 pictures). In search for suitable and unambiguous pictures, the coders were guided by each value’s description offered by Schwartz et al. (2012). Particular attention was paid to the pictures’ diversity. This was necessary because in step 3 we compare whether single pictures or picture collages (representing different facets) are more effective in identifying values. Twenty pictures were selected for each value, later reduced to five or six by expert judgement. So, 135 single pictures were chosen for further investigation.

**Step 3: Validation studies**

**Verbal Validation (3a).** Sixty-nine participants (students and employees of Saarland University in Germany, 54% female; \( M_{\text{age}} = 25.61 \) years, SD = 7.54) were asked to allocate either single pictures or picture collages to a value list. Participants were told that the list showed different aspects that are more or less important in peoples’ lives. Each picture respectively each collage could only be allocated to one value in the list. All but one picture collages were assigned correctly to the value list (the collage for “considering the future” was assigned correctly to the value list by only 55.3% of the respondents whereas the percentages of respondents assigning all the other collages correctly to the corresponding value ranged between 76.3% and 100% and were on average much higher than in the single picture condition, demonstrating a superiority of picture collages). Next, 106 further participants (61% female; \( M_{\text{age}} = 24.76 \) years, SD = 7.59) were asked to provide unaided labeling of the stimuli. Again, the collages outperformed the single pictures with respect to correct value identification (20 of 26 values labeled correctly in the collage condition, whereas in the single picture condition, only 18 values were labeled correctly). As values are abstract in nature (Torelli and Kaikati 2009), picture collages seem to represent them more accurately than single pictures do. To reduce the image material, recipients completing the labeling task for the collages were also asked to rank the pictures in each collage according to which picture best represents the corresponding value. A clear order emerged for all of the 20 correctly labeled collages.

**Stack Sort Analysis (3b).** To test the validity of the picture collages more comprehensively (and to further increase it) and also to control for the six values that were not identified in the verbal validation task, a stack sort analysis was conducted with 160 participants (60% female; \( M_{\text{age}} = 24.74 \) years, SD = 8.93) at Saarland University. Step 3a shows that picture collages represent value varieties better than single pictures. The pictures in the collages used in step 3a were the results of step 2. To discover whether consumers regard the same pictures as related, participants were asked to sort a number of pictures by perceived coherence. They were completely free to choose the number of stacks and the number of pictures per stack. To avoid cognitive overload, the number of pictures was subdivided into three units, so each participant was only shown pictures representing one-third of the 26 value varieties. Using co-assignment rates of all pairs of pictures for hierarchical cluster analysis (Ward’s procedure), 22 of the 26 value categories were replicated, which means 22 clusters consisted of the same pictures that were chosen via step 2 and used in step 3a. The values “tolerance,” “diversity” and “purposefulness” could not be replicated. For these categories, pictures were replaced and tested subsequently until the values were identified correctly.

Owing to the results of the stack sort task, “possession of money” and “prestige” formed one instead of two clusters; this was not changed.

**Comparison of Labeled Versus Unlabeled Picture Collages**

Based on these studies, the final pictorial scale for measuring peoples’ value orientations consists of 25 picture collages, each with three pictures. In a further step, labeled versus unlabeled collages were tested. On the one hand, labeling the collages might help dispel ambiguity and make the scale easier to use. Additionally, time needed to capture the intended value and therefore to fulfill the task might be shorter for labeled collages. On the other hand, labeling the collages with verbal stimuli may again lead to a SDB. As reducing this bias is one of the major goals of a pictorial scale, the unlabeled scale should be at least as good as the labeled scale with respect to performance, handling and enjoyment. To examine this, 52 students (between-subjects-design, \( n = 26 \) per group; 55.7% female; \( M_{\text{age}} = 22.31 \) years, SD = 3.01) were asked to attribute the 25 value categories (either with or without labeling) to a stimulus. A famous football club (FC Bayern Munich) and the pharmacist’s profession were chosen as stimuli. Based on expert judgement and website information, a broad variety of values compatible with the two stimuli were defined (appreciation, possession of money / prestige, success, dominance, purposefulness, friendship, excitement, quality, tolerance, diversity, and trust for Bayern Munich; security, quality, family, health, and trust for the pharmacist’s profession). The two scale formats did not differ with regard to time needed to fulfill the task (\( t(50) = 0.993, \) NS) or with regard to the number of correct and incorrect attributed values. Afterwards, participants were asked to evaluate the collages (either with or without labeling) with regard to handling and enjoyment. Results show that the different scale formats did not significantly differ with regard to handling, and unlabeled collages were more pleasant for participants than labeled ones (\( t(50) = 2.273, p = .027 \)).

**Comparing Verbal Versus Pictorial Scales Concerning SDB**

In the last step, the final unlabeled pictorial scale was tested against a verbal scale in relation to SDB. For the verbal scale, we used parts of the SVS (Schwartz 1992). In a group discussion with 11 participants at Saarland University as a pretest, seven (of the 25) values were identified as potentially sensitive, that is they might trigger social desirability concerns (Tourangeau and Yan 2007). Six of these seven values appear to have a high self-presentation potential (e.g., social justice or openness to something new), thus leading to expected over-reporting by recipients. One value (dominance) tends to be viewed more negatively, so under-reporting was expected. Eighty-one participants (50.4% female, \( M_{\text{age}} = 32.19 \) years, SD = 13.75) were asked to rate the importance of these values as guiding principles on either a verbal or a pictorial seven-point scale. Results show that the six values for which over-reporting was expected scored higher in the verbal (vs. pictorial) scale condition (e.g., openness to something new: \( M_{\text{verbal}} = 5.48, M_{\text{pictorial}} = 3.71 \)), whereas dominance, for which under-reporting was expected, scored lower in the verbal than the pictorial scale condition (\( M_{\text{verbal}} = 1.65, M_{\text{pictorial}} = 2.38 \)). It can therefore be concluded that verbal value scales are, compared to pictorial scales, more affected by over- (or under-)reporting of sensitive items, and therefore by SDB.

After the value rating, people read different scenarios, each presenting a typical behavior associated with the tested values. The description was either mirroring the specific value or was contradictory to it. For instance, participants read a scenario about themselves in which they want to get rid of old CDs, actually knowing that CDs...
are not allowed to be thrown in the garbage, but have to be disposed of properly in special waste. This takes time and effort. Participants were asked to indicate the likelihood of performing the behavior described in the scenario on a seven-point scale (one single statement per described behavior for likelihood measurement, e.g. “How likely is it that you would dispose of the CDs in the garbage nevertheless?”, 1 = not at all likely, 7 = very likely). Different studies (Eyal et al. 2009; Torelli and Kaikati 2009) use this method to reveal significant correlations between values and the likelihood of showing a value-congruent behavior. As the likelihood of performing the described behavior is measured with verbal scales and the behaviors explained in the scenarios seem to have high self-presentation potential (as values do, because the described behavioral reactions refer to specific values), this verbal self-reported likelihood of showing the described value-congruent behavior is expected to be influenced by SDB for the same reason value self-reports are. I.e., if somebody over-reports the importance of, for example, environmentalism values for reasons of social desirability, this person is also likely to over-report the likelihood of showing a described environmentally friendly behavior. Therefore, if the correlation between those measures is high, SDB seems obvious. Hence, in a further step, we tested whether significant correlations between value importance and likelihood of performing a described value-congruent behavior only occur for value importance measured with verbal, but not with pictorial scales. Results show that for five of the seven values tested, the expected effect occurred in the verbal condition, reinforcing the assumption that value measurement by using pictorial instead of verbal scales reduces the risk of SDB.

**General Discussion**

This article illustrates the development of a pictorial scale for science and practice to measure peoples’ value orientations with substantially reduced SDB. A thorough literature review and qualitative focus group discussions were conducted to identify subtle variants of values with importance for consumer research. Several steps were taken to assure the validity of the developed pictorial scale. Our study shows that picture collages outperform single pictures and do not need additional labeling for comprehension and handling. Additionally, our results show that verbal scales (compared to pictorial scales) have a higher risk to be fraught with over-reporting (or under-reporting if the expected social consequences are negative) and are therefore biased. Furthermore, significant correlations between value importance and the self-reported likelihood of performing a verbally described value-congruent behavior only occurred if value importance was measured with verbal scales. This leads again to the assumption that pictorial scales are less affected by SDB than verbal scales. However, an alternative explanation is also imaginable: correlations between verbally measured values and verbally measured likelihoods of performing a described value-congruent behavior might always be higher (because of the same modality) than correlations between pictorial and verbal stimuli (different modalities). Future research could address this problem, e.g. by observing actual behavior (instead of measuring the likelihood of showing a hypothethical behavior described in a scenario). But as values can be more or less abstract, they might influence either near or distant future behavior. Studies comparing values’ influence on actual behavior therefore should regard the potentially moderating effect of temporal distance (Torelli and Kaikati 2009).

The importance of investigation into ethical consumption is increasing. People claim to care for the environment, but organic products still only account for a minor share of turnover. Simultaneously, the (seemingly) increasing concern for ethical values might lead to an under-reporting of materialistic values (which are opposed to universalism values and therefore might be associated with negative social consequences), that might nevertheless be important for consumers’ buying decisions. Under-reporting of these values might therefore bias the results of verbal surveys. Pictorial scales provide insights into consumers’ true value orientations without SDB, which might contribute to deeper understanding of this phenomenon and enable manufacturers and retailers to focus more accurately on consumers’ actual needs to get true data to use for aligning their marketing strategies.

The new measurement approach could be used to analyze the relevance of values for customer-retailer- as well as for customer-brand-relationships. Two studies conducted in Germany using the here developed picture scale on a touchscreen monitor (figure 1) showed promising results that a match of consumers’ and retailers’ values can be beneficial for both. Manufacturers and retailers could therefore try to match their positioning concerning diverse values with people’s most relevant values to share a common value basis and encourage long-term relationships. To do that, true (unbiased) data is needed – and could be provided by pictorial scales. Beyond that, the two studies conducted revealed that using picture scales is enjoyable for respondents – even if the interviews lasted 45 up to 60 minutes. So the developed approach seems to enable researchers to analyze comprehensive and complex issues without causing a fatigue effect.

**REFERENCES**


