# Chapter 9
# The Social Desirability Response Bias in Ethics Research

**Donna M. Randall and Maria F. Fernandes**

Since the early 1950s researchers in the organizational sciences (e.g., Edwards 1953) have expressed concern that self-report questionnaires may contain response sets biasing observed relationships between variables. One artifact, in particular, which may impair the results of survey research and policy capturing (Arnold and Feldman 1981; Bishop et al. 1986; Mazen 1990) is a social desirability response set. Social desirability (SD) is broadly understood as the tendency of individuals to deny socially undesirable traits and behaviors and to admit to socially desirable ones (Zerbe and Paulhus 1987).

Researchers in business ethics, a rapidly developing subdiscipline within the field of management, need to be particularly sensitive to the potential effects of a social desirability response bias. Observation and measurement of business ethics is difficult (Trevino 1986). While self-report questionnaires are very commonly used as an observation technique in business ethics research, empirical studies have noted a high degree of sensitivity on the part of managers to questions about ethics (e.g., Victor and Cullen 1988). Respondents are frequently asked to express their agreement or disagreement with a statement such as, "Ethical practices are good business in the long-run" (Brown and King 1982, p. 15) or express their opinion about a behavior such as "Rejection of qualified job applicant because he is Jewish" (Goodman and Crawford 1974, p. 182). The socially desirable answer in such statements is quite apparent.

Due to the sensitive nature of ethics research, the presence of a social desirability response bias may pose an even greater threat to the validity of findings in ethics research than in more traditional organizational behavior research topics. However, little effort has been directed toward determining the impact of a response bias in ethics research.

D.M. Randall (✉)
Office of the President, Albion College, Albion, MI 49224-1831, USA
e-mail: drandall@albion.edu

M.F. Fernandes
7445 W. Devonwood Drive, 83703 Boise, ID, USA
e-mail: ffernadeseidahopower.com

A.C. Michalos and D.C. Poff (eds.), *Citation Classics from the Journal of Business Ethics*, Advances in Business Ethics Research 2,
DOI 10.1007/978-94-007-4126-3_9, 

A recent review of business ethics research (Randall and Gibson 1990) revealed that self-report data were relied upon in almost 90% of empirical journal articles. However, only one of 96 empirical research articles since 1960 has attempted to assess the impact of a social desirability response bias. This paper seeks to examine the relationship between conceptually distinct measures of social desirability responding and self-reported ethical conduct.

## Social Desirability

Individuals may have some attributes that are negatively valued by general societal norms (e.g., abuse of alcohol, abuse of drugs, drunken driving) and other attributes that are positively valued (e.g., voting in elections, church attendance) (Groves 1989). In an effort to conform to societal norms, individuals may present themselves in a favorable light, regardless of their "true" feelings or "actual" behavior. Specifically, individuals may under-report those activities perceived to be socially or culturally undesirable and may over-report those activities deemed to be socially or culturally desirable (Ganster et al. 1983).

From the early 1930s, researchers have expressed interest in the effects of a social desirability response set (e.g., Bernreuter 1933; Humm and Humm 1944). Results from these early tests raised the suspicion that test-takers who scored high on these tests were "faking in order to look good." Ganster et al. (1983) viewed this tendency as problematic because it may mask the relationship between two or more variables (a suppressor effect), provide a false correlation between independent and dependent variables (a spurious effect), or moderate the relationship between those variables (a moderator effect).

## Dimensions of Social Desirability

Organizational science literature today generally depicts the social desirability construct as composed of two independent dimensions. When viewed as a personality characteristic, social desirability is frequently termed "need for social approval," and when viewed as an item characteristic, it is labelled "trait desirability" (Gove and Geerken 1977; Phillips and Clancy 1972).

### *SD as a Personality Characteristic*

When viewed as a personality characteristic, a social desirability response is frequently operationalized using the Marlowe-Crowne (M-C) Scale (Crowne and Marlowe 1960). The scale consists of items drawn from a set of behaviors which are "culturally sanctioned and approved, but which are improbable of occurrence," such as "Before voting, I thoroughly investigate the qualifications of all candidates"

(Crowne and Marlowe 1960, p. 23). Compared with low scorers, high scorers on the M-C scale respond more to social reinforcement, restrain aggressive feelings, and are more amenable to social influence. Their performance on tasks is strongly influenced by how others evaluate them. They prefer to engage in low-risk behaviors and avoid evaluations by others (Crowne 1979).

When the M-C scale was published in 1960, it was intended as a general measure of social desirability in self-reports and assumed to reflect a need for social approval. However, questions have been raised about the validity of the M-C scale as a measure of need for approval. For instance, a series of studies were conducted investigating the relationship between frequency and amount of cheating behavior and M-C scores (Jacobson et al. 1970; Millham 1974). As predicted, those who cheated scored higher on the M-C scale than non-cheaters when detection was perceived to be unlikely (Millham 1974). In addition, those who scored high on the M-C scale cheated only enough to avoid disapproval, but did not cheat when approval could be won. These findings supported Crandall's (1966) conclusion that the M-C scale appears to be more appropriate for measuring the impulse to avoid disapproval rather than the need to seek social approval.

More recently, Paulhus (1984) contended that the scale contains, and fails to differentiate between, two distinct factors: self-deception and impression management. Self-deception refers to an unconscious tendency to see oneself in a positive light and is manifested in self-descriptions that are socially desirable, biased, and believed to be true by the respondent (Zerbe and Paulhus 1987). In self-reporting behavior, the respondent is assumed to be motivated to protect self-beliefs, including self-esteem (Paulhus 1986). Conversely, impression management refers to a conscious presentation of a false front, manifested by deliberately falsifying test responses to create a positive impression (Zerbe and Paulhus 1987). The respondent's behavior is assumed to be instrumental (Paulhus 1986).

Due to perceived limitations of the M-C scale, Paulhus (1984, 1988) developed the Balanced Inventory of Desirable Responding (BIDR) scale which contains two distinct subscales for self-deception and impression management. The BIDR is a descendant of the Self- and Other-Deception Questionnaires developed by Sackeim and Gur (1978). While the original self-deception items were developed under the assumption that individuals with a propensity for self-deception tend to deny having psychologically threatening thoughts or feelings, the most recent version of the scale (Paulhus 1988) emphasizes exaggerated claims of positive cognitive attributes.

The impression management items were selected under the assumption that some respondents consciously tend to over-report performance of desirable behaviors and under-report undesirable behaviors. Because the claims involve overt behaviors (e.g., I have never dropped litter on the street), any distortion can be presumed to be a conscious lie (Paulhus 1989). It is assumed that self-deception is minimized because the questions are concerned with the behavior of the respondent rather than the respondent's thoughts.

The preceding discussion reveals that social desirability as a personality trait has been reconceptualized in two highly distinct ways – a propensity for self-deception and a propensity for impression management – and that each conceptualization is

independently related to self-reported conduct. Applying these findings to the ethics literature, the following hypotheses can be set forth:

$H_1$: The greater the propensity for self-deception, the greater the extent to which individuals will self-report ethical behaviors.

$H_2$: The greater the propensity for impression management, the greater the extent to which individuals will self-report ethical behaviors.

## *SD as an Item Characteristic*

The second approach to a social desirability response bias, perceived desirability of the item, considers various behaviors or traits to be more or less socially desirable and thus discusses social desirability in relation to particular items. Strong support exists for the conclusion that social desirability effects are heavily influenced by characteristics of the item (Groves 1989).

In an early study, Edwards (1953) provided empirical support for the relationship between the judged desirability of a response in a self-report study and the likelihood of an individual giving that response. Edwards' findings were replicated in a series of investigations focusing on the influence of perceived item desirability on responses to various personality measures (e.g., Cowen and Tongas 1959; Rosen 1956; Wiggins and Rumrill 1959). Several years later, Phillips and Clancy (1972) again demonstrated that respondents consistently reported themselves as possessing characteristics they saw as desirable.

As a consequence, one can hypothesize:

$H_3$: The greater the perceived desirability of behavior, the greater the extent to which individuals will self-report ethical behavior.

Phillips and Clancy (1972) also examined the joint influence of need for approval and item characteristic measures on self-reported attributes. Contrary to what might be expected, the two possible response determinants (need for approval and item desirability) were found to be generally unrelated to each other but independently related to individuals' responses to various sociological measures. Further, they determined that item desirability exerted a greater influence than need for approval on individuals' responses. While Phillips and Clancy's measurement of trait desirability has recently been criticized for its complexity by Gove and Geerken (1977), the latter researchers did support the finding that item desirability and need for approval are largely independent and their effects additive.

It appears that while the personality of some individuals may predispose them toward a general pattern of socially desirable responding, their answers will be more strongly influenced by the situational influences – their perception of the desirability of engaging in specific behaviors. Thus,

$H_4$: Perceived desirability of behavior will exercise a stronger influence on self-reported ethical behavior than will propensity for self-deception or impression management.

## Questioning the Existence of a Social Desirability Response Set

Several researchers (e.g., Groves 1989; Nunnally 1978; Paulhus 1988) have raised the possibility that any association between personality characteristics, item desirability, and self-reported behavior may accurately reflect the true state of affairs. That is, different levels of self-deception, impression management, or item desirability may be associated with actual differences in ethical conduct. This conduct is, in turn, accurately reported within surveys.

To test this alternative explanation, it would be desirable to observe actual behavior. As this is difficult to do, Phillips and Clancy (1972) developed an overclaiming scale to give researchers confidence that respondents actually do not perform claimed behaviors. Their scale seeks to measure the tendency to assert that one has accomplished some task when, in fact, this is not true. Phillips and Clancy asked respondents about their use of several new products, books, television programs, and movies (all of which were nonexistent), and the desirability of being the kind of person who uses these items. They found that those respondents who viewed the use of the items as highly desirable were more than twice as likely to give inaccurate responses (to overclaim) than those viewing the characteristics as highly undesirable. Phillips and Clancy concluded that the desirability of performing the behaviors influenced reporting of the behaviors and argued that the alternative explanation, that individuals' self-reported responses are accurate, cannot be fully supported.

However, Bradburn and his colleagues (1979) disagreed with Phillips and Clancy's conclusion. They determined that those individuals scoring high on need for approval actually behaved differently on a number of measures and that there are "pervasive real-world differences in the way persons with high and low MC scores behave and relate to other people" (p. 105). As DeMaio (1984) noted, findings from Phillips and Clancy's study and Bradburn et al.'s study can be reconciled as the former study is based on an item desirability measure of social desirability, while the latter is based on a personality characteristic measure of social desirability.

As a consequence, one can hypothesize:

$H_5$: Overclaiming will be more closely associated with the perceived desirability of ethical behavior than with propensity for self-deception or impression management.

Finally, two exploratory analyses will be conducted. On occasion, researchers have attempted to assess the influence of a social desirability response bias in ethics questionnaires (e.g., Stevens 1984) by incorporating the M-C scale into the study. As discussed above, despite the current popularity of the M-C scale, its validity has been called into question and other, more valid, scales (e.g., the BIDR) have been developed to identify the existence of a social desirability response bias. This study proposes to examine the relationship between scores on the M-C scale to measures of self-deception, impression management, item desirability, and self-reported ethical conduct.

Second, despite heavy reliance on self-report methodology, minimal effort in ethics research has been directed toward an assessment of the impact of a social desirability response bias on ethics scales commonly included in self-report questionnaires. This study also proposes to explore the impact of a SD response bias on one of the most popular ethics scale, the Ruch and Newstrom scale (1975).

## Methods

A survey instrument was designed which included the following measures: the Marlowe-Crowne scale, the BIDR (with self-deception and impression management subscales), perceived item desirability of unethical behavior, an overclaiming scale, self-report unethical conduct, and Ruch and Newstrom's (1975) ethics scale. Prior to administering the survey, a pretest of the instrument was conducted using 43 junior and senior level students at a large state university. Items identified as either confusing, through a debriefing of pretest subjects, or lacking variability, through analysis of pretest results, were modified.

The revised questionnaire was then administered to students enrolled in an introductory management class. The students were asked to fill out a brief questionnaire in exchange for extra credit. Numerous precautions were taken by the researchers to promote disclosure of unethical conduct: assuring anonymity, asking that no name or identification mark be put on the survey, requesting participants to write their identification number only on a separate cover sheet (to receive extra credit), providing clearly visible and separate public drop boxes for the cover sheet and survey, administering the survey in a public auditorium, and asking for no identifying information in the survey other than level in school and gender. The questionnaire was completed by 348 students (50% female and roughly equal numbers of juniors and seniors).

The survey instrument contained the following seven scales (discussed in the order in which they appeared in the questionnaire):

### *Overclaiming*

The procedure to detect overclaimers set forth by Phillips and Clancy (1972) was modified to allow for the use of an overclaiming scale with more items and items of greater relevance to students. On a scale of 1 to 5 (1 = not at all familiar, 3 = somewhat familiar, and 5 = very familiar), respondents were asked to rate their degree of familiarity with items in several categories: newly released movies, products, music albums, television programs, and designer label clothing. Each category contained five items, two of which were non-existent.

All affirmative responses to behaviors which could not possibly have occurred were tabulated to arrive at an overclaiming score. Possible values on the overclaiming

scale ranged from 10 to 50. A score of 10 indicated the respondent was not at all familiar with any of the 10 fake items. A score of 50 indicated the respondent reported being very familiar with each of the 10 fake items. The internal consistency coefficient (alpha) for the scale was 0.70. Appendix A presents the 10 items composing the scale.

Respondents are often unwilling to admit their ignorance in areas they believe themselves to be experts (Bradley 1981). In a study of fictitious public affairs issues, Bishop et al. (1986) hypothesized that people can be pressured into giving an opinion on a fictitious issue when the topic seemed familiar to them. However, they found the more a person knew about a subject, the less likely he or she was to make such a mistake. In fact, the more respondents knew about a subject, the more easily they could recognize what was familiar and what was not. Therefore, the less knowledgeable a subject is about a topic, the more easily the person can be confused and pressured to give an opinion about it. As a consequence, the overclaiming scale was rescored after controlling for respondents' alleged expertise in each of the five areas. This was done by dividing the familiarity rating of non-existent items by the total familiarity ratings within each topic area. The respondent's total score is the sum across the five topics. As the results were highly correlated with the original scoring method, all subsequent analyses of the overclaiming scale used the original scoring method.

### *Marlowe-Crowne Scale*

The Marlowe-Crowne Scale (Crowne and Marlowe 1960) was incorporated into the study for exploratory purposes. Respondents were asked to express their agreement by responding true or false to 33 items (15 of which were negatively coded). This instrument included such items as, "I am always courteous, even to people who are disagreeable," and "I never hesitate to go out of my way to help someone in trouble."

In past research, the average internal consistency coefficients of the Marlowe-Crowne Scale have ranged from 0.73 to 0.88 (Crowne and Marlowe 1964; Fisher 1967; Paulhus 1984; Tanaka-Matsumi and Kameoka 1986). A test-retest correlation of 0.88 over 1 month was reported by Crowne and Marlowe (1964), and 0.84 over a 1-week interval was reported by Fisher (1967). In the present study, the internal consistency coefficient (alpha) was 0.74.

### *Self-Report Behaviors*

Students were asked to report whether they had engaged in a series of 10 unethical behaviors by responding true or false to each of the behaviors. The 10 behaviors had been previously identified in a study of students by Stem and Steinhorst (1984) and included such behaviors as, "exchanging answers with another student during an

exam," "plagiarizing on a term paper," "receiving help on a take-home exam," and "turning in the same paper for two classes." The internal consistency coefficient (alpha) for the scale was 0.65.

### *Self-Deception and Impression Management*

The Balanced Inventory of Desirable Responding (BIDR) contains two 20-item subscales measuring self-deception (an honest positivistic bias) and impression management (purposeful self-presentation) (Paulhus 1989). Respondents were asked to rate how each of 40 statements applied to themselves on a seven-point scale (1 = not true at all, 4 = neither true or false, and 7 = very true). Ten items in each subscale were reverse coded. As set forth by Paulhus (1989), all 40 items on the BIDR were summed to yield an overall measure of socially desirable responding. In addition, separate subscale scores were computed by summing the 20 items composing the self-deception subscale and the 20 items composing the impression management subscale.

Past research has shown that the average internal consistency coefficient for the full BIDR is 0.83, 0.68–0.80 for the self-deception subscale, and 0.75–0.86 for the impression management subscale (see Paulhus 1988). Average test-retest correlations of 0.69 and 0.65 over a 5 week period were reported by Paulhus (1988) for the self-deception and impression management scales, respectively. The complete BIDR demonstrates concurrent validity in correlating 0.71 with the M-C scale (Paulhus 1988) and 0.80 with Jacobson et al.'s (1970) Multidimensional Social Desirability Inventory. In the present study, the internal consistency coefficient for the full BIDR was 0.79, 0.60 for the self-deception subscale, and 0.79 for the impression management subscale.

### *Item Desirability*

Following a procedure set forth by Dohrenwend (1966) and Phillips and Clancy (1972), item desirability was assessed by having respondents rate each self-reported unethical behavior on a nine-point scale of desirability (1 = very undesirable, 5 = neither undesirable or desirable, and 9 = very desirable). The less desirable respondents believed a behavior to be, the lower the assigned number. The internal consistency coefficient of the scale (alpha) in the present study was 0.83.

### *Ruch and Newstrom Scale*

One of the most frequently used scales in the ethics literature to measure perceptions of unethical conduct was developed by Ruch and Newstrom (1975) (e.g., Ferrell and Weaver 1978; Izraeli 1988; Kidwell et al. 1987; Krugman and

Ferrell 1981; Newstrom and Ruch 1975, 1976). In completing the scale, students were asked to report how unethical they perceived a series of 17 business practices to be using a five-point scale (1 very unethical, 2 = basically unethical, 3 = ethically neutral, 4 = basically ethical, and 5 = very ethical). Examples of scale items include: "Using company services for personal use," and "padding an expense account up to 10%" (Ruch and Newstrom 1975, p. 18). High values indicate that respondents consider the questionable practices to be highly ethical. The 17 items were summed to yield a scale score. The internal consistency coefficient (*alpha*) for the scale in the present study was 0.83.

As the intent of the research was to examine the impact of socially desirable responding on self-reported ethical conduct, the hypotheses were tested using a series of zero order correlations or multiple regressions.

## Results

### *Hypotheses*

The first hypothesis examined the relationship between propensity for self-deception and self-report ethical conduct. It was expected that the greater the propensity for self-deception, the greater the extent to which individuals will report ethical behavior. As reflected in the correlation matrix in Table 9.1, the correlation between scores on the self-deception subscale and self-reported ethical conduct was 0.10 ($p < 0.05$). The first hypothesis was supported.

**Table 9.1** Pearson correlations between social desirability measures, overclaiming, and self-reported ethical conduct[a]

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. Self-reported ethical behavior | – | | | | | | | |
| 2. Desirability of ethical behavior | 0.68*** | – | | | | | | |
| 3. M-C scale | 0.24*** | 0.26*** | – | | | | | |
| 4. BIDR | 0.42*** | 0.39*** | 0.64*** | – | | | | |
| 5. Self-deception subscale | 0.10* | 0.11* | 0.45*** | 0.76*** | – | | | |
| 6. Impression management subscale | 0.53*** | 0.49*** | 0.59*** | 0.88*** | 0.36*** | – | | |
| 7. Overclaiming scale | –0.07 | –0.09 | 0.18*** | 0.13** | 0.14** | 0.11* | – | |
| 8. Ruch and Newstrom's scale[b] | –0.16*** | –0.29*** | –0.17*** | –0.19*** | –0.05 | –0.24*** | 0.14** | – |

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

[a] $N = 319–341$

[b] High values = acceptance of unethical behavior

**Table 9.2** Magnitude of influence of item desirability and personality characteristic measures on self-report behavior

| Variable | *B* | *SEB* | *Beta* | *T* | *SigT* |
|---|---|---|---|---|---|
| Impression management subscale | 0.0372 | 0.0063 | 0.2895 | 5.91 | 0.00 |
| Self-deception subscale | –0.0127 | 0.0077 | –0.0706 | –1.65 | 0.10 |
| Desirability of ethical behavior | 0.0828 | –0.0072 | 0.5322 | 11.58 | 0.00 |
| (Constant) | 7.5121 | 0.7716 | | 9.74 | 0.00 |

Multiple *R* 0.71
*R* Square 0.50
Adjusted *R* Square 0.50
Standard Error 1.48

The second hypothesis examined the relationship between propensity for impression management and self-report ethical conduct. It was expected that the greater the propensity for impression management, the greater the extent to which individuals will report ethical behavior. The correlation between scores on the impression management subscale and self-reported ethical conduct was 0.53 ($p<0.001$). The second hypothesis was supported.

The third hypothesis set forth that the more desirable individuals assess ethical behavior, the greater they will self-report the behavior. The results in Table 9.1 indicated that the correlation between perceived desirability of ethical behavior and self- reported ethical conduct was 0.68 ($p<0.001$). The third hypothesis was supported.

The fourth hypothesis set forth that trait desirability will exercise a stronger influence on self-reported behavior than propensity for self-deception or propensity for impression management. To examine this hypothesis, self-reported ethical behavior was regressed on perceived item desirability, the impression management subscale, and the self-deception subscale. Table 9.2 reveals that the three independent variables explained 50% of the variation in self-reported unethical conduct. Perceived item desirability explained the largest share of that variance with the impression management subscale making a small, additional contribution. Further, addition of variables into the regression equation did not appreciably change the beta weights for those variables, reducing concern about the moderate level of association between independent variables. In sum, the fourth hypothesis was strongly supported.

The fifth hypothesis proposed overclaiming would be more strongly associated with perceived desirability of ethical behavior than with propensity for self-deception or impression management. The mean score on the overclaiming scale was 13.5. Of the 348 individuals responding, 103 (30%) did not overclaim on any of the 10 nonexistent items. As reflected in Table 9.1, overclaiming was significantly correlated with the self-deception subscale ($r=0.14$, $p<0.01$) and with the impression management subscale ($r=0.11$, $p<0.05$), but was not significantly correlated with the perceived desirability of engaging in ethical conduct ($r=-0.09$). Hence, Hypothesis 5 was rejected.

### *Exploratory Analyses*

#### Marlowe-Crowne Scale

The Marlowe-Crowne scale was significantly correlated with self-reported ethical behavior ($r=0.24$), the perceived desirability of ethical behavior ($r=0.26$), the self-deception subscale ($r=0.45$), the impression management subscale ($r=0.59$), the overclaiming scale ($r=0.18$), and Ruch and Newstrom's scale ($r=-0.17$). All correlations were significant at the 0.001 level. However, entering the M-C scale into the regression equation in Table 9.2 reveals that the scale failed to make a significant contribution to the prediction of self-reported ethical conduct over that of the three variables previously entered into the equation.

#### Ruch and Newstrom's Scale

Ruch and Newstrom's scale was significantly correlated with self-reported ethical behavior ($r=-0.16$, $p<0.001$), the perceived desirability of ethical behavior ($r=-0.29$, $p<0.001$), the Marlowe-Crowne scale ($r=-0.17$, $p<0.001$), the impression management subscale ($r=-0.24$, $p<0.001$), and the overclaiming scale ($r=0.14$, $p<0.01$). However, it was not significantly correlated with the self-deception subscale ($r=-0.05$).

## Discussion

As set forth in the first and second hypotheses, propensity for self-deception and impression management were significantly correlated with self-reported ethical behavior. As anticipated, the strength of the correlation varied measurably with the operationalization of the SD personality characteristic, confirming the importance of separating the two dimensions of the BIDR conceptually and analytically. The self-deception subscale was only weakly correlated with self-reported ethical behavior, while the impression management subscale of the BIDR reflected a stronger correlation. Thus, it appears that self-reported ethical conduct is more closely associated with a conscious over-reporting of desirable behaviors and under-reporting of undesirable behaviors, than it is associated with an unconscious tendency as measured by the self-deception subscale.

It was interesting to note that the self-report behavior scale had a relatively low internal consistency estimate (*alpha*=0.65). Subsequent item analysis revealed that the deletion of any item from the scale failed to increase the scale's consistency estimate. Hence, it appears that respondents differentiate between various unethical activities. Such a conclusion is consistent with the finding by Bradburn and colleagues

(1979) that social desirability effects are heavily influenced by characteristics of the item. As Grover (1990) has suggested, researchers may need to examine ethical conduct at the molecular level, i.e., study a single element of unethical behavior, as opposed to examining conduct at the molar level, i.e., viewing unethical behaviors as interchangeable.

As predicted in the third and fourth hypotheses, perceived desirability of ethical behavior and self-reported ethical conduct were positively and strongly correlated, with item desirability exerting a significantly stronger influence on self-reported ethical behavior than propensity for self-deception or impression management. In the present study, the addition of an impression management subscale to the regression equation minimally contributed to explained variation in self-reported behavior, and the addition of a self-deception subscale added very little to a regression equation containing an item desirability measure.

As discussed earlier, one could argue that self- reported ethical conduct accurately reflects the behavior of individuals with certain personality characteristics and with certain perceptions of item desirability. However, findings from the present study indicate that those who report what we know to be false familiarity with a series of items have higher levels of self-deception and impression management. Yet, unexpectedly, overclaiming was not significantly related to perceived desirability of ethical behavior, leading to the rejection of Hypothesis 5. It may be the case that the tendency to overclaim is more closely associated with a personality trait than with item desirability.

The magnitudes of the self-deception-overclaiming correlation and the impression management-overclaiming correlation are not large. Overclaiming clearly does not account for all variance in self-reported ethical behavior. A psychological patterning (ethical individuals provide socially desirable responses which agree with their behavior) appears to explain a majority of our findings, yet it does not explain all. At least a portion of the variance in self-reported ethical conduct appears to be due to a social desirability response bias.

In terms of exploratory analyses, the commonly used M-C scale revealed a weak correlation with self-reported ethical behavior. It appears that self-reported ethical conduct is more closely associated with impression management and perceived desirability of behavior than with a tendency to avoid disapproval as measured by the M-C scale. After taking into consideration other SD measures, the M-C scale adds little to explained variation in self- report ethical conduct.

Finally, exploratory analyses revealed that responses to Ruch and Newstrom's ethics scale were significantly correlated with all measures of social desirability except the self-deception subscale. Those individuals who identified a series of questionable business practices as highly unethical had significantly higher levels of impression management and perceived item desirability. Such a pattern of findings may indicate that responses to Ruch and Newstrom's scale, and possibly other ethics scales, may be influenced more by a conscious tendency to over-report desirable behaviors and a desire to project a particular image than by an unconscious tendency measured by the self-deception subscale.

## Implications

The examination of the effects of a social desirability response bias on self-reported ethical conduct has a number of research and practical implications. First, as past social science research has convincingly demonstrated that people tend to report behavior in light of what they feel others will expect is appropriate (e.g., Cicourel 1964; Friedman 1967; Riecken 1962; Rosenthal 1966), it is imperative that researchers dealing with such a value-laden topic as ethical conduct be sensitized to the possibility of a strong social desirability bias. Managers are often reluctant to have their ethics observed or measured, and few employees may agree to provide information to researchers that might be incriminating to them or to their friends.

Previous research has convincingly demonstrated that observed levels of socially desirable responding vary with the level of anonymity. The more anonymity seems assured, the less socially desirability responding is detected (Bradburn et al. 1979; Nederhof 1985; Paulhus 1984; Wiseman 1972). In an experiment contrasting anonymous and public conditions, Paulhus (1984) determined that impression management scales were more sensitive to situational changes in anonymity than self-deception scales.

However, it is likely that steps that ethics researchers commonly take to minimize a social desirability response bias (e.g., asking that names not be placed on the survey instrument itself or assuring respondents that their names will never be associated with their findings) will not completely reduce the influence of a social desirability response bias. Although generalizations may be speculative, our research demonstrated that a social desirability bias persists even if a survey is administered in a non-threatening situation. Despite the numerous precautions we took to assure a totally anonymous survey administration, a significant social desirability response bias, largely due to impression management, was still observed. (However, one might argue that, as reflected by scores on the overclaiming scale, 30% of the respondents did not give any evidence of a social desirability bias.)

It is possible that had we not followed these precautionary measures, stronger relationships between the measures of social desirability (particularly, the impression management subscale) and individuals' responses would have been observed. As business ethics researchers do not commonly build in as many precautions as we did, we might argue that our findings reflect a floor for the influence of a social desirability response bias. Hence, the results presented here raise serious questions about the validity of self-report ethics research.

If a social desirability response bias persists after anonymous testing conditions, it may be possible to eliminate the influence of a social desirability response bias through the use of alternative methodologies such as randomized response methods, forced-choice items, proxy subjects, or computer administration (for a discussion of these techniques see Aupperle 1984; Lautenschlager and Flaherty 1990; Martin and Nagao 1989; Nederhof 1985; Paulhus 1984). However, if the use of these techniques is not possible or advisable, business ethics researchers could control for or partial out socially desirable response effects from hypothesized

relationships (unless the influence of a social desirability response bias is of theoretic interest, see Zerbe and Paulhus 1987). To aid in this effort, Ganster et al. (1983) have detailed a procedure for identifying how the social desirability response set influences the relationship between independent and dependent variables. The results of the present study provide a useful complement to Ganster et al.'s research in that the latter did not explore how the SD response bias should be measured, only the type of influence a SD response bias (as measured by the Marlowe-Crowne scale) has on the relationship between independent and dependent variables.

From the present study one might conclude that further use of the M-C scale is not advisable. Zerbe and Paulhus (1987) also contended that studies using the M-C scale may underestimate the actual relationship between social desirability responding and organizational behavior constructs. Hence, business ethics studies which have incorporated the M-C scale in an effort to detect a social desirability bias and failed to find such an effect (e.g., Stevens 1984) may have utilized a weak measure. Findings from the present study indicate that if researchers desire to investigate social desirability as a personality characteristic, use of the impression management subscale of the BIDR is preferable. Furthermore, since item desirability consistently exerted a stronger influence than any personality characteristic measure in our study, an assessment of perceived item desirability of the dependent variable appears to be preferable in future research.

Inclusion of an overclaiming scale into a questionnaire may provide a less cumbersome method of detecting a social desirability response bias than an item desirability assessment when the number of items to be rated is large. An overclaiming scale constitutes a direct and unambiguous measure of an individual's attempt to deceive on a questionnaire (as the items are known to be non-existent). However, in the present study the scale failed to explain any additional variation over that of item desirability and impression management when entered into a multiple regression equation. Further, as overclaiming was only correlated with SD personality characteristics, overclaiming may not tap a social desirability response bias due to perceived item desirability.

The lack of a strong correlation between overclaiming and the various measures of social desirability incorporated into the present study may be attributable to either questionable construct validity of the other SD scales (for the overclaiming scale does ultimately represent a direct measure of deception) or to unreliability of the overclaiming scale itself (as the coefficient alpha is moderately low at 0.70). Before further empirical work with the overclaiming scale is conducted, it is desirable that theoretical linkages between various SD personality characteristic measures, item desirability, and overclaiming be specified. On the one hand, it may be that overclaiming mediates the linkage between personality characteristics and self-reported behavior and the linkage between item desirability and self-reported behavior. On the other hand, overclaiming may have a direct and unique impact as an independent variable on unethical behavior. More research needs to explore why individuals claim knowledge when they lack it and whether overclaiming is an unconscious or conscious tendency.

As noted earlier, studies on overclaiming report discrepant findings. Whereas Phillips and Clancy (1972) determined a clear linkage between overclaiming and item desirability and Bradburn and colleagues (1979) found no linkage between overclaiming and the M-C scale, we found a significant linkage between overclaiming and all SD personality characteristic measures, but no significant linkage between overclaiming and item desirability. Findings from the present study would indicate that overclaiming is more closely associated with a personality characteristic than with item desirability. Thus, overclaiming may only mediate the personality characteristic-self-reported behavior relationship.

Whenever a student sample is used, a caveat regarding the issue of generalizability is necessary. Due to the nature of the present sample, extrapolation of these results to employees in different organizations and to research on other sensitive topics should be undertaken with some caution. Nevertheless, the study was designed to provide a high degree of realism in regard to survey design and administration. Future research is needed to develop a nomological net for the overclaiming construct, to validate the scale developed within this study, and to replicate findings from our study with different samples and with a different set of unethical behaviors. The results reported here highlight the need for and can facilitate further investigation into the influence of social desirability in business ethics research.

## Appendix A: Overclaiming Scale

How familiar are you with each of the following newly released movies?

1. Turned to Gold
2. Katherine's Mistake

How familiar are you with each of the following products?

1. Microsoft Statistical Assistant
2. New Life Spices

How familiar are you with each of the following albums?

1. Cosmic Being
2. Offender After Dark

How familiar are you with each of the following TV programs?

1. The Adventures of Johnnie
2. Chicago Heat

How familiar are you with each of the following designer labels?

1. Ocean City
2. Jones L. A.

All answers provided with a five point Likert scale (1 = not at all familiar, 3 = somewhat familiar, and 5 = very familiar).

## References

Arnold, J.H., and D.G. Feldman. 1981. Social desirability response bias in self-report choice situations. *Academy of Management Journal* 24: 377–385.

Aupperle, K.E. 1984. An empirical measure of corporate social orientation. *Research in Corporate Social Performance and Policy* 6: 627–654.

Bernreuter, R.G. 1933. Validity of the personality inventory. *Personality Journal* 11: 383–386.

Bishop, G.F., A.J. Tuchfarber, and R.W. Oldendick. 1986. Opinions on fictitious issues: The pressure to answer survey questions. *Public Opinion Quarterly* 50: 240–250.

Bradburn, N.M., S. Sudman, and Associates. 1979. *Improving interview method and questionnaire design: Response effects to threatening questions in survey research*. San Francisco: Jossey-Bass.

Bradley, J.V. 1981. Overconfidence in ignorant experts. *Bulletin of the Psychonomic Society* 17(2): 82–84.

Brown, D.J., and K.B. King. 1982. Small business ethics: Influences and perceptions. *Journal of Small Business Management* 20: 11–18.

Cicourel, A.V. 1964. *Method and measurement in sociology*. New York: Free Press.

Cowen, E., and P. Tongas. 1959. The social desirability of trait descriptive terms. *Journal of Consulting Psychology* 3: 361–365.

Crandall, V. 1966. Personality characteristics and social achievement behaviors associated with children's social desirability response tendencies. *Journal of Personality and Social Psychology* 4: 447–486.

Crowne, D.P. 1979. *The experimental study of personality*. Hillsdale: Erlbaum.

Crowne, D.P., and D. Marlowe. 1960. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology* 24: 349–354.

Crowne, D.P., and D. Marlowe. 1964. *The approval motive*. New York: Wiley.

Demaio, T.J. 1984. Social desirability and survey measurement: A review. In *Surveying subjective phenomena*, vol. 2, ed. C.F. Turner and E. Martin, 257–282. New York: Russell Sage.

Dohrenwend, B.P. 1966. Social status and psychiatric disorder: An issue of substance and an issue of method. *American Sociological Review* 23: 361–365.

Edwards, A. 1953. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology* 37: 90–93.

Ferrell, O.C., and K.M. Weaver. 1978. Ethical beliefs of marketing managers. *Journal of Marketing* 42(3): 69–73.

Fisher, G. 1967. Normative and reliability data for the standard and the cross-validated Marlowe-Crowne social desirability scale. *Psychological Reports* 20: 174.

Friedman, N. 1967. *The social nature of psychological research*. New York: Basic.

Ganster, D.C., H.W. Hennessy, and F. Luthans. 1983. Social desirability response effects: Three alternative models. *Academy of Management Journal* 26(2): 321–331.

Goodman, C.S., and M. Crawford. 1974. Young executives: A source of new ethics? *The Personnel Journal* 53: 180–187.

Gove, W.R., and M.P. Geerken. 1977. Response bias in surveys of mental health: An empirical examination. *The American Journal of Sociology* 82: 1289–1317.

Grover, S. L. 1990. The influence of professional role conflict and moral maturity on professionals' reporting accuracy. Paper presented at the 1990 Academy of Management meetings, San Francisco, CA, August 12–15.

Groves, R.M. 1989. *Survey errors and survey costs*. New York: Wiley.

Humm, D.G., and K.A. Humm. 1944. Validity of the Humm-Wadsworth temperament scale: With consideration of the effects of subjects' response-bias. *Journal of Psychology* 18: 55–64.

Izraeli, D. 1988. Ethical beliefs and behavior among managers: A cross-cultural perspective. *Journal of Business Ethics* 7: 263–271.

Jacobson, L.I., S.E. Berger, and J. Millham. 1970. Individual differences in cheating during a temptation period when confronting failure. *Journal of Personality and Social Psychology* 15: 48–56.

Kidwell, J.M., R.E. Stevens, and A.L. Bethke. 1987. Differences in ethical perceptions between male and female managers: Myth or reality? *Journal of Business Ethics* 6: 489–493.
Krugman, D.M., and O.C. Ferrell. 1981. The organizational ethics of advertising. *Journal of Advertising* 10(1): 21–30.
Lautenschlager, G.J., and V.L. Flaherty. 1990. Computer administration of questions: More desirable or more social desirability? *Journal of Applied Psychology* 75(3): 310–314.
Martin, C.L., and D.H. Nagao. 1989. Some effects of computerized interviewing on job applicant responses. *Journal of Applied Psychology* 74(1): 72–80.
Mazen, A.M. 1990. The moderating role of social desirability, age, and experience in human judgment: Just how indirect is policy capturing? *Organizational Behavior and Human Decision Processes* 45: 19–40.
Millham, J. 1974. Two components of need of approval score and their relationship to cheating following success and failure. *Journal of Research in Personality* 8: 378–392.
Nederhof, A.J. 1985. Methods of coping with social desirability bias: A review. *European Journal of Social Psychology* 15: 263–280.
Newstrom, J.W., and W.A. Ruch. 1975. The ethics of management and the management of ethics. *MSU Business Topics* 23(Winter): 29–37.
Newstrom, J.W., and W.A. Ruch. 1976. Managerial values underlying intra-organizational ethics. *Atlanta Economic Review* 26 (May–June): 12–18.
Nunnally, J.C. 1978. *Psychometric theory*, 2nd ed. New York: McGraw-Hill.
Paulhus, D.L. 1984. Two-component models of socially desirable responding. *Journal of Personality and Social Psychology* 46: 598–609.
Paulhus, D.L. 1986. Self-deception and impression management in test responses. In *Personality assessment via questionnaire: Current issues in theory and measurement*, ed. A. Angleitner and J.S. Wiggins. Berlin: Springer.
Paulhus, D.L. 1988. Development of the balanced inventory of desirable responding. Unpublished manuscript, University of British Columbia.
Paulhus, D.L. 1989. Measurement and control of response bias. In *Measures of social-psychological attitudes*, ed. J.P. Robinson, P.R. Shaver, and L. Wrightsman. New York: Academic.
Phillips, D.L., and K.J. Clancy. 1972. Some effects of 'social desirability' in survey studies. *The American Journal of Sociology* 77(5): 921–940.
Randall, D.M., and A.M. Gibson. 1990. Methodology in business ethics research: A review and critical assessment. *Journal of Business Ethics* 9(6): 457–472.
Riecken, H.W. 1962. A program for research on experiments in social psychology. In *Values and groups*, ed. N.F. Washburn. New York: Pergamon.
Rosen, E. 1956. Self-appraisal, personal desirability, and perceived social desirability of personality traits. *Journal of Social Psychology* 52: 151–158.
Rosenthal, R. 1966. *Experimental effects in behavioral research*. New York: Appleton.
Ruch, W.A., and J.W. Newstrom. 1975. How unethical are we?, *Supervisory Management,* pp. 16–21.
Sackeim, H.A., and R. Gur. 1978. Self-deception, self- confrontation and consciousness. In *Consciousness and self-regulation: Advances in research*, vol. 2, ed. G.E. Schwartz and D. Shapiro, 139–197. New York: Plenum.
Stem Jr., D.E., and R.K. Steinhorst. 1984. Telephone interview and mail questionnaire applications of the randomized response model. *Journal of the American Statistical Association* 79(387): 555–564.
Stevens, G.E. 1984. Business ethics and social responsibility: The responses of present and future managers. *Akron Business and Economic Review* 15(3): 6–11.
Tanaka-Matsumi, J., and V.A. Kameoka. 1986. Reliabilities and concurrent validities of popular self-report measures of depression, anxiety, and social desirability. *Journal of Consulting and Clinical Psychology* 54: 328–333.
Trevino, L.K. 1986. Ethical decision making in organizations: A person-situation interactionist model. *Academy of Management Review* 11(3): 601–611.
Victor, B., and J.B. Cullen. 1988. The organizational bases of ethical work climates. *Administrative Science Quarterly* 33: 101–125.

Wiggins, J., and C. Rumrill. 1959. Social desirability in the MMPI and Welsh's factor scales A and R. *Journal of Consulting Psychology* 23: 100–106.

Wiseman, F. 1972. Methodological bias in public opinion surveys. *Public Opinion Quarterly* 36: 105–108.

Zerbe, W.J., and D.L. Paulhus. 1987. Socially desirable responding in organizational behavior: A reconception. *Academy of Management Journal* 12(2): 250–264.