Journal of Business and Psychology, Vol. 17, No. 2, Winter 2002 (©2002)

UNDERSTANDING SELF-REPORT BIAS IN ORGANIZATIONAL BEHAVIOR RESEARCH

Stewart I. Donaldson
Claremont Graduate University
Elisa J. Grant-Vallone
California State University, San Marcos

ABSTRACT: Self-report and mono-method bias often threaten the validity of research conducted in business settings and thus hinder the development of theories of organizational behavior. This paper outlines a conceptual framework for understanding factors that influence the motivation of an employee to bias his or her responses to questions posed by organizational researchers. Using a longitudinal, multitrait-multimethod dataset, we illustrate various aspects of the problem and argue that traditional approaches for controlling self-report bias do not adequately prevent the problem. The results suggest the need for developing a theory of method effects and companion analytic techniques to improve the accuracy of psychological research in business settings.

KEY WORDS: self-report bias; organizational behavior research; method bias; construct validity.

Theoretical advances in organizational behavior and psychology are highly dependent upon empirical confirmation and disconfirmation (cf. McCall & Bobko, 1990; Sackett & Larson, 1990). That is, theoretical perspectives supported by numerous empirical studies become dominant in the field. Theories associated with a collection of null or mixed empirical

This work was supported by a grant from the National Institute of Mental Health to Dr. Stewart Donaldson (R03 MH 50230-02). We would like to thank Dr. John Graham for his assistance with the conceptualization and design of Project WORKWELL. Special thanks to Lynne Mellett for her outstanding contributions during the execution of this study, and to Kelly Higgins and the blind reviewers for smoothing out the rough edges.

Address correspondence to Dr. Stewart I. Donaldson, Department of Psychology, Claremont Graduate University, 123 E. Eighth Street, Claremont, CA 91711; Stewart.Donaldson@cgu.edu.

0889-3268/02/1200-0245/0

results become controversial and often fade away. This pattern is desirable if the empirical studies are accurate. However, there continues to be concern about the accuracy of some of the most commonly used methods in organizational research.

Accurate measurement of organizational behavior is essential for advancing the field. Despite its importance, measurement in organizational settings is often referred to as one of the main shortcomings of organizational behavior research (Donaldson, 1995; Donaldson, Ensher & Grant-Vallone, 2000; Mersman & Donaldson, 2000). This is because researchers must rely to a large extent on self-reports. Such measures are common because they are relatively easy to obtain and are often the only feasible way to assess constructs of interest.

Sackett and Larson (1990) found that over a third of all studies published in mainstream organizational behavioral journals between 1977 and 1987 were questionnaire-based. It was found that 83% of these studies used a cross-sectional design and 52% relied solely on self-report measures. Studies which rely on self reports as the only measure of organizational behaviors have come under attack recently for two primary reasons: 1) self-reports are prone to many kinds of response bias (see Campbell & Fiske, 1959; Donaldson, Thomas, & Graham, 2002; Graham, Collins, Donaldson, & Hansen, 1993; Schwartz, 1999; Stone et al., 2002), and 2) inferences about correlational and causal relationships may be inflated by the problem of common method variance (Borman, 1991; Donaldson, Thomas, Graham, Au, & Hansen, 2000; Spector, 1994).

Schmitt (1994) suggested the need for a theory of method bias in organizational behavior research in order to understand how to prevent and control for it. He provided a sample matrix illustrating ways motivational method biases might affect a variety of measures. Spector and Brannick (1995) argued that because the nature of measurement bias varies based on properties of the construct of interest, measurement bias will be best understood when it is studied in relation to specific constructs or possibly a particular research domain (e.g., health promotion in business settings). The purpose of this paper is to take a step in that direction by applying what has been learned about self-report and mono-method bias in business settings to constructs often assessed in employee assistance and corporate health promotion research (see Donaldson, Gooler, & Weiss, 1998; Donaldson & Klein, 1997; Donaldson & Weiss, 1998). This will be accomplished by (a) examining a general conceptual framework for understanding potential motivational biases in self-report questions used to gather factual data in workplace research, and (b) applying this framework specifically to constructs relevant to the domain of corporate health promotion.

SELF-REPORT BIAS

In general, research participants want to respond in a way that makes them look as good as possible. Thus, they tend to under-report behaviors deemed inappropriate by researchers or other observers, and they tend to over-report behaviors viewed as appropriate. Self-report bias is particularly likely in organizational behavior research because employees often believe there is at least a remote possibility that their employer could gain access to their responses.

This tendency for individuals to respond in socially desirable ways has been studied extensively (cf. Moorman, & Podsakoff, 1992; Zerbe, & Paulhus, 1987). In a meta-analysis, Moorman and Podsakoff (1992) found that social desirability, as measured by the Crowne-Marlowe Scale (Crowne & Marlowe, 1964), was related to several commonly used constructs in organizational behavior research (e.g., general job satisfaction, role conflict, role ambiguity, and organizational commitment). Borman (1991) reported that on average, peer and supervisor performance ratings were more accurate than self-ratings because of differential leniency effects in self-reports. Furthermore, Spector (1994) discussed the problem of self-reports and underscored the importance of considering the specific questions being asked and the type of research questions one wants to answer when using self-reports.

THE COMMON METHOD VARIANCE PROBLEM

The validity of conventional studies relying on one source (method) of data has been questioned quite often in recent times (e.g., see Donaldson, 1995; Glick, Jenkins, & Gupta, 1986; Howard, 1994; Podsakoff, & Organ 1986; Schmitt, 1994; Shadish, 1993; Spector, 1994; Williams & Brown, 1994). The problem of self-report bias is compounded by the fact that when all variables in an organizational behavior study are based on one method of measurement, substantive findings are likely to be contaminated by shared method variance. It is important to point out that the common method variance problem is not unique to self-report measures.

In a policy oriented editorial, a former editor of the *Journal of Applied Psychology* stated that the sole use of self-report measures is an unacceptable methodology in most areas of organizational research (Schmitt, 1989). However, the extent to which common method variance (e.g., studies based solely on self reports) affects research conclusions is still hotly debated. For example, Spector (1987) concluded that there is little evidence of the problem in a study that examined the relationship between perceived working conditions and affect. Williams, Cote, and

Buckley (1989) claim that his conclusion was wrong and the result of improper analytic procedures. Several years later, the *Journal of Organizational Behavior* published an invited set of essays outlining the importance of the problem for the field (Howard, 1994; Schmitt, 1994; Spector, 1994), and encouraged researchers to conduct empirically-based studies to understand the effects of method variance in organizational behavior research.

PRESENT STUDY

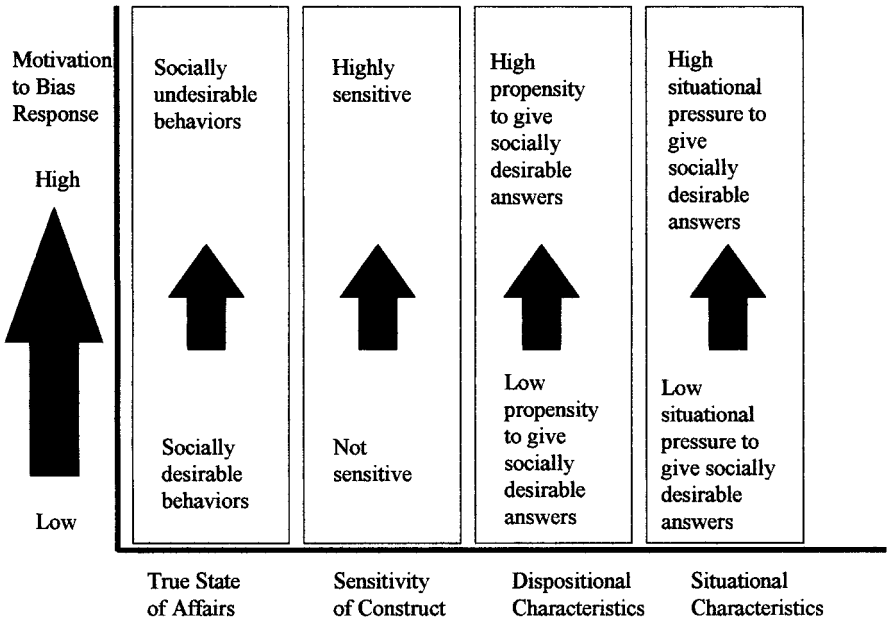
The aim of the present study is to develop knowledge about self-report bias in relation to constructs relevant to psychological research in business settings. First, we propose that it is useful to conceptualize motivational self-report bias as the function of the following four general factors:

1. The "true state of affairs" (i.e., the true score on the construct versus random or systematic error)
2. The nature of the construct of interest (e.g., sensitivity of the information)
3. Dispositional characteristics of the respondent
4. Situational characteristics of the measurement situation or environment (see Figure 1)

For example, a research participant who is abusing drugs at work (true state of affairs), reporting on his or her level of drug use (a potentially sensitive construct), who has a propensity to give socially desirable responses (a dispositional characteristic), and who is in a situation where he or she believes that true responses could cause him or her to be punished (e.g., fired) by his or her employer (a situational characteristic) is highly likely to bias responses on a research questionnaire. Whereas, a participant who abusing drugs at work with different dispositional characteristics and/or under different conditions may be more likely to provide accurate answers.

The purpose of the following analyses is to begin to examine the problem of self-report bias in relation to specific constructs. First, we will examine how well self and co-worker reports of the same variables converge across a number of constructs often assessed in psychological research carried out in business settings. This analysis will reveal whether there is variability in the agreement about the nature of constructs, and if this relative agreement or disagreement is stable across time. Second, we will examine the effects of an employee's propensity to give socially desirable responses (a dispositional characteristic) and his

Figure 1
Four Factors that Influence Self-Report Bias



or her level of fear of reprisal (a situational construct or a construct dependent of the "true state of affairs") on responses to questions about organizational behavior and employee well-being. Third, we will illustrate how mono-method bias can distort conventional parameter estimates of the substantive relationships between constructs measured in organizational behavior research. Finally, we will explore the question of how to determine which source of data to use (e.g., self or co-worker) when there is low agreement about the construct of interest.

DATA SOURCE

Project WORKWELL

Project WORKWELL was designed to shed light on the nature of response bias in organizational behavior and occupational health psychology research. Using a multitrait-multimethod (MTMM) dataset, the project intended to demonstrate some of the practical value initially foreseen by Campbell and Fiske (1959). In short, by collecting data about an employee's lifestyle, mental health, and work-related behavior (a) from

the employee him or herself, and (b) from people who know the employee well (e.g., co-workers), one is in a position to understand and potentially correct for some of the biases associated with each source of data (e.g., employee self-report bias; Donaldson, 1995).

Participants

The Project WORKWELL database consists of two waves of data (six months apart) from 408 (204 co-worker pairs) ethnically diverse non-professional level employees (cf. Donaldson, Ensher, & Grant-Vallone 2000; Ensher, Grant-Vallone, & Donaldson, 2001; Mersman & Donaldson, 2000). Of the 408 participants, 68% of the sample were females, and 45% Latino, 22% European American, 16% African-American, and 10% Asian. Approximately 56% of the employees were single, 26% married, 13% divorced, 3% separated, and 2% widowed. Further, 50% reported a total personal income of less than $20,000 per year and 87% report a total personal income of less than $30,000 a year. More than 87% reported their education level to be less than a 2-year junior college degree. The participants worked an average of 38 hours per week ($SD = 11.2$; *Median* = 40) and had worked for their company for an average of 4.8 years ($SD = 5.25$; *Median* = 3 years). Finally, the participants represent a broad range of companies, industries, and occupations. For example, approximately 20% worked for educational institutions, 17% health care organizations, 11% financial services, 7% retail, 7% county agencies, 6% general service companies (e.g., hair salons, movers, united postal service, etc.), 5% manufacturing, 4% utility companies, 3% municipalities, 3% law offices, 3% grocery stores, 2% nursing homes, 2% churches, 2% restaurants, 2% automotive sales and repair, 2% insurance companies, 2% temporary agencies, and 2% federal agencies (e.g., U.S. Attorney's office, U.S. Census, FEMA). Each participant received $25 for the first visit, $25 for the second visit, a lifestyle assessment, and a summary of preliminary research findings. Only 11% of the original sample did not return to complete the six-month follow-up assessment (i.e., an 89% completion rate).

Procedure

A number of strategies were used to schedule working adults to participate in this study. First, fliers were displayed in highly visible locations at worksites within a 10 mile radius of the data collection site (lunch rooms, hallways, parking structures, and various central meeting places). Second, advertisements were placed in a wide variety of local newspapers, community publications, and company newsletters. Finally, efforts were made to generate referrals from people who were not eligible

for the study themselves, potential participants, and those who actually came to the Research Institute to participate. Some of the referral sources were given a $5 incentive (per co-worker pair) to recruit participants. All of our recruitment materials offered working adults $50 to offset the costs associated with participation, a free lifestyle assessment, and free copies of the research findings in exchange for their participation.

The focus of all these efforts was to get the potential participants to call our study telephone number. Our recruitment telephone line described the eligibility criteria for participating in the study. Those who were eligible were asked to leave their phone number so we could contact them to set up an appointment. All potential participants were screened during a follow-up telephone call. Only employees 1) who worked at least twenty hours per week, 2) who had lived at the same address and had the same phone number for a least six months (this was needed to prevent excessive attrition) or worked in the same company for at least six months, and 3) who did not have a four-year college degree were enrolled in the study. The participants were encouraged (but not required) to bring a co-worker who they knew well; preferably the co-worker who they knew best. The participants completed a battery of questionnaires during the initial visit, and again six months later. The items were counterbalanced to control for order effects at both waves of data collection.

Measures

The Project WORKWELL database contains over 1,400 items per participant. In the present study, we used several self report bias measures to assess social desirability and fear of reprisal. The criterion variables analyzed in this study included self and co-worker reports of job performance, absenteeism, organizational citizenship behavior, loafing at work, employee grievances, physician visits, health care costs, depression, anxiety, vitality, alcohol use, tobacco use, drug use, height, and weight. Data collector ratings of employee height and weight were also recorded, and the actual height and weight of each participants were measured after all questionnaires were completed at wave 2. Finally, we used a variety of measures to assess working conditions. Below, we briefly describe the measurement instruments used.

Bias Measures. The Crowne-Marlowe social desirability scale (Crowne & Marlowe, 1964) was used to assess the employees propensity to give socially desirable responses. Several items administered at the end of the data collection period assessed whether the participants felt they was any possibility that their responses could jeopardize their employment. For example, “Even though this questionnaire is confidential, I was still

concerned that someone might identify my name with my responses," and "Because of the sensitive nature of some of the questions, I did not feel comfortable about giving completely accurate answers to all of the questions."

Criterion Measures. Both employees and co-workers answered a variety of questions about employee job performance, absenteeism, organizational citizenship behavior, loafing at work, employee grievances, physician visits, health care costs, depression, anxiety, vitality, alcohol use, tobacco use, and drug use. These are variables often measured in employee assistance and corporate health promotion research (Donaldson & Klein, 1997). Most of these items were developed using well-established health and work performance measures (e.g., The Wellness Inventory, The Job Content Questionnaire, The General Well-Being Schedule, and The Seriousness of Illness Rating Scale; for health-related lifestyle items including alcohol and drug use see Donaldson, Dent, Sussman, Severson, & Stoddard, 1996; Donaldson, Graham, Piccinin, & Hansen, 1995; Lifestyle Assessment Questionnaire, 1988; for work-related items see Donaldson, 1993; Donaldson & Blanchard, 1995; Karasek, 1985; for health and well-being items see McDowell & Newell, 1987; Wyler, Masuda, & Holmes, 1968).

Predictor Variable Measures. The Job Content Questionnaire (JCQ; see Karasek, 1985; Karasek & Theorell, 1990) was the main instrument used to measure dimensions of the employees' working conditions (e.g., job insecurity, job demands, decision latitude, supervisor support, co-worker support). One advantage of using the JCQ is that it has been found to have sound psychometric properties and national averages are available for many of the items. In addition, the Decontaminated Daily Hassles Scale was used to measure work-related hassles (cf. Smith, 1993).

Background Characteristics. The participants also provided their gender, ethnicity, marital status, the number of children they have, education, personal income, spouses income, history of personal and family health problems, body weight, type of company worked for, job title, length of employment, hours worked per week, family health history, the type of health care providers to which they have access, and their preventive health care practices. However, for the purposes of this paper, we only used gender and ethnicity as control variables in most of our final analyses.

Analyses Strategy

First, reliable standardized scales (alpha > .70) were constructed for the 14 criterion variables. The convergence between the self and co-

worker reports at both wave 1 and wave 2 were examined for each of these constructs. Next, we selected 6 of the criterion constructs of varying levels of convergence between self and co-worker reports. Participants were divided into conditions based on the degree to which they had a propensity to give socially desirable responses (low vs. high) and their level of fear of reprisal (low vs. high). This was analyzed using an ANCOVA (controlling for gender and ethnicity) across the six constructs of interest. Next, a series of partial correlation analyses (controlling for gender and ethnicity) were conducted to illustrate the effects of common method variance on conventional parameter estimates in organizational research. Finally, because we were able to determine the "true state of affairs" for employee height and weight, analyses were conducted to assess the convergence of the various height and weight measures, and absolute differences between the actual measures ("true state of affairs") and the various methods for estimating height and weight.

RESULTS

Agreement Among Co-workers

The results showed that convergence between self and co-worker reports varied considerably across constructs. For example, the following appeared at wave 1: employee tobacco use ($r = .80$, $p < .01$), employee drug use ($r = .40$, $p < .01$), attendance ($r = .36$, $p < .01$), employee vitality ($r = .34$, $p < .01$), loafing at work ($r = .28$, $p < .01$), job performance ($r = .26$, $p < .01$), citizenship behavior ($r = .13$, $p < .01$), and employee grievances ($r = .12$, $p < .01$). While the level of agreement varied considerably between variables, consistent results were found within variables across the wave 1 and wave 2 data collection. The results of these analyses are displayed in Table 1.

Propensity to Give Socially Desirable Responses

Self-reports of five of the six constructs that were examined appeared to be affected by an employee's propensity to give socially desirable responses. At both wave 1 and wave 2, employees who had a high propensity to give socially desirable answers rated themselves higher on job performance ($F = 7.03$, $p < .01$; $F = 6.81$, $p < .05$), citizenship behavior ($F = 10.11$, $p < .01$; $F = 8.69$, $p < .01$), and vitality ($F = 12.10$, $p < .01$; $F = 8.45$, $p < .01$). In addition, they rated themselves significantly lower on drug use (time 1 only; $F = 10.11$, $p < .01$) and on their tendency to loaf at work ($F = 16.11$, $p < .01$; $F = 12.60$, $p < .01$). There were no significant differences regarding reports of attendance.

Table 1
Zero-order Correlations Between Self and Co-Worker Reports at Waves 1 and 2

	Wave 1 Convergence	Wave 2 Convergence
Tobacco Use	.80**	.86**
Alcohol Use	.64**	.75**
Exercise	.52**	.52**
Drug Use	.40**	.39**
Absenteeism	.40**	.50**
Physician Visits	.36**	.53**
Vitality	.34**	.34**
Depression	.34**	.31**
Anxiety	.29**	.30**
Loafing at Work	.29**	.27**
Health Care Costs	.28**	.43**
Job Performance	.26**	.21**
Citizenship Behavior	.13**	.13**
Grievances	.12**	.12**

**$p < .01$.

Fear of Reprisal

Responses also seemed to be affected by an employee's fear of reprisal. Significant differences were found at both wave 1 and wave 2 between those who reported low fear of reprisal and those who reported high fear on loafing at work ($F = 17.36$, $p < .01$; $F = 11.96$, $p < .01$) and employee ratings of absenteeism ($F = 12.84$, $p < .01$; $F = 9.72$, $p < .01$). Those who reported high fear of reprisal rated themselves as more likely to loaf at work and reported lower attendance rates. They also rated themselves lower on vitality at wave 1 ($F = 6.28$, $p < .05$).

Illustration of the Common Method Variance Problem

Analyses contaminated with shared or common method variance are prevalent in organizational behavior research (Sackett & Larson, 1990). Thus, the third issue we considered is how this threat to construct validity can affect conclusions drawn from empirically-based organizational research. Using depression as the criterion, we conducted a series of partial correlation analyses controlling for gender and ethnicity to illustrate this problem.

The most common parameter estimates in organizational behavior research are cross-sectional correlations based solely on self-reports. Our results showed that self-reported depression was significantly correlated with an employee's self-reported decision latitude ($r = .14$, $p < .01$), job insecurity ($r = .20$, $p < .01$), exercise ($r = -.12$, $p < .01$), and daily hassles

at work ($r = .39$, $p < .01$) at wave 1. However, self-reported depression was only significantly correlated with exercise ($r = -.12$, $p < .01$) and daily hassles at work ($r = .16$, $p < .01$) at wave 1 when the co-worker reports were analyzed. This pattern was typical across our criterion variables and illustrates that the common cross-sectional design based solely on self-reports would lead to different conclusions than a design including at least on additional source of information about the behaviors of interest (co-worker reports in this example).

A second set of analyses examined the relationship between decision latitude, job insecurity, exercise, and daily hassles at work at wave 1 and depression at wave 2. The results showed that self-reported depression at wave 2 was only associated with self-reported job insecurity ($r = .14$, $p < .01$), and self-reported ($r = .33$, $p < .01$) and co-worker reported daily hassles at work ($r = .15$, $p < .01$) at time 1. This pattern illustrates that only daily hassles at work remained a significant predictor of depression across the type of measure and time period. The relationships between the other variables were constrained to using a cross-sectional design and/or only one source of data (self-reports).

Which Data Source Is Most Accurate?

Given that it is reasonable to assume that both self-report and co-worker reports contain at least some bias, it is important to understand which perspective is most accurate or least biased. Of course, the answer will most likely depend on which variable is being assessed. Further, in order to answer this question, knowledge of the "true state of affairs" must be known.

While it was impossible to obtain the "truth" about employee depression, job performance, decision latitude, exercise and the like in Project WORKWELL, we were able to determine the "true state of affairs" for employee height and weight. In order to answer the question of accuracy, we collected self, co-worker, and data collector estimates of employee height and weight. At the completion of the study we asked employees to allow us to measure their actual height and weight.

Correlations between the four methods for assessing employee height and weight showed that agreement among sources was quite high for these variables (all correlations were >.75, $p < .01$ for height; all correlations were >.80 for weight). Subsequent analyses were performed to assessed the absolute differences between the actual measures and the various methods for estimating height and weight. It was found that self-reports were significantly more accurate than the co-worker (height $t(317) = 5.75$, $p < .01$; weight $t(316) = 9.36$, $p < .01$) and data collector reports (height $t(317) = 12.18$, $p < .01$; weight $t(316) = 8.99$, $p < .01$).

DISCUSSION

The results of this study suggest that self-report bias tends not to be uniform across constructs assessed in psychological research conducted in business settings. For example, zero-order correlations between self and co-worker reports of the same variables differed considerably across the constructs examined in this research. The various analyses conducted seem to support the notion that the nature of the construct (e.g., level of sensitivity) in combination with employee characteristics (e.g., actual behavior or true score and propensity to give socially desirable responses) and situational pressures must be considered to fully understand and account for self-report bias in organizational behavior research.

These findings have at least two important implications: (1) even the most sophisticated analytic procedures currently available for controlling for response bias, including the various procedures using multitrait-multimethod matrices (see Becker & Vance, 1992; Browne, 1993; Graham et al., 1993; Spector & Brannick, 1995), are not likely to adequately capture this multidimensional complexity, and (2) a minimum of two data sources are needed to help rule out the validity threats of self-report and mono-method bias in business psychology research. Further, the current findings show specific variables likely to be influenced by response bias, and suggest that additional studies are needed to develop a more complete framework that shows which psychological and behavioral measures under specific circumstances are likely to be valid.

On average, employees with a propensity to provide socially desirable responses reported more favorable behaviors than those who were lower on this dispositional characteristic. These findings suggest that regardless of situational factors in a measurement setting, some participants are naturally more prone to bias their answers. Future empirical work might focus on other dimensions that are useful for understanding and controlling for dispositional characteristics that distort the accuracy of organizational behavior measurement.

The fear of reprisal measure was expected to be an indicator of the participants' feelings about the measurement situation (i.e., responses would remain confidential). However, the results suggest that it is more likely operated as a correlate of the participants actual behavior (true score or the "true state of affairs"). That is, for vitality at wave 1 and loafing and attendance at both waves 1 and 2, those who were not fearful reported the most favorable behavior. This suggests that if actual behavior (true score) is undesirable participants fear the repercussions of reporting it to researchers. Again, this highlights the importance of designing studies that are able to isolate and control for both dispositional (actual behavior) and situational method effects.

The illustration of the problem of mono-method bias showed how easily workplace researchers can be mislead when measuring all constructs using the same method or data source. Secondly, it makes clear that self-report bias and mono-method bias are related but different threats to the validity of an empirical study. The analyses of height and weight showed that one should not automatically assume that self-reports are the inferior source of data in workplace research, and that the argument that co-worker or supervisor reports are necessarily better than self reports is dubious (cf. Howard, 1994). Future work is needed to develop a conceptual framework for understanding the biases in co-worker and supervisor reports as well as self-reports.

Another important point to be made from these analyses is that while common method variance appears to substantially inflate parameter estimates, measuring constructs with different but also imperfect methods (e.g., co-worker reports) appeared to underestimate relationships. In any given study, this type of pattern might function as a confidence interval, indicating that the true parameter lies between these two parameter estimates. Finally, it is important to emphasize that using multiple sources of data, in contrast to one (e.g., all self-reports), is a desirable strategy for avoiding the problem of being unable to rule out the problem of mono-method bias in business psychology research (Donaldson, 1995; Shadish, 1993).

Limitations

The empirical analyses conducted in this paper were based on prospective data collected at two points in time. However, all of the constructs were measured (versus manipulated) which limits our ability to sort out cause-effect relationships. For example, the social desirability bias and fear of reprisal analyses are open to problems associated with ruling out reverse causality and unmeasured "third variables" as alternative explanations. Of course, the type of design that would allow unequivocal interpretation of these causal links (a true experiment) is not feasible with the specific variables we chose to analyze in this paper. Nevertheless, the results of this paper must be interpreted cautiously.

The criterion variables selected for the analyses are a subset of variables often analyzed by workplace health promotion researchers. The findings reported may be very different for other constructs of interest to business psychologists. Further, we only used one criterion variable and a restricted number of predictor variables to illustrate the problem of mono-method bias. Although the general principles discussed seem to apply across the range of variables in Project WORKWELL, the robustness of this pattern needs to be determined in future work.

Finally, we acknowledge that it is possible that the external validity

of our findings is somewhat limited. Due to the nature and scope of the proposed study, obtaining a representative sample of ethnically-diverse non-professional level employees was not feasible. It is possible that our recruiting methods have produced a rather unique sample. Further research is needed examine if the findings of this study generalize to other samples and populations.

CONCLUSION

Empirical studies can have long-standing effects on our theoretical understanding of organizational behavior and business psychology. This implies that the popularity or extinction of theoretical perspectives may sometimes be the function of imperfect empirical work rather than sound or faulty theory. The issues presented in this paper suggest that it is feasible that many empirical studies conducted in business settings to date can be very misleading. We agree with Schmitt (1994) that more empirical work needs to be done to develop a better understanding of method effects in organizational research. This paper establishes some initial benchmarks for understanding which measures are likely to be misleading (e.g., some measures of performance and mental health). It is our hope that additional research will be conducted to further develop understanding about the problem of self-report bias in business psychology research. A conceptual framework of methods effects backed up by empirical findings, and credible companion analytic techniques could go a long way toward furthering the progress of psychological theory and research in business settings.

REFERENCES

Becker, T. E., & Vance, R. J. (1992). Construct validity of three types of organizational citizenship behavior using a direct product model. Paper presented at the 1992 Academy of Management Meeting, Las Vegas (August).

Borman, W. C. (1991). Job behavior, performance, and effectiveness. In M. D. Dunnette & L. M. (Eds.), *Handbook of industrial and organizational psychology* (pp. 271–326). Newbury Park, CA: Sage.

Browne, M. W. (1984). The decomposition of multitrait-multimethod matrices. *British Journal of Mathematical and Statistical Psychology*, *37*, 1–21.

Browne, M. W. (1993). Models for multitrait-multimethod matrices. In R. Steyer, K.F. Wender, & K.F. Widamen (Eds.), *Psychometric Methodology* (pp. 570–584). Stuttgart and New York: Gustav Fisher Verlag.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

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

Donaldson, S. I. (1993). Effects of lifestyle and stress on the employee and organization: Implications for promoting health at work. *Anxiety, Stress, and Coping*, *6*, 155–177.

Donaldson, S. I. (1995). Worksite health promotion: A theory-driven, empirically based perspective. In L. R. Murphy, J. J. Hurrel, S. L. Sauter, & G. P. Keita (Eds.), *Job stress interventions* (pp. 73–90). Washington, DC: American Psychological Association.

Donaldson, S. I., & Blanchard, A. L. (1995). The seven health practices, well-being, and performance at work: Evidence for the value of reaching small and underserved worksites. *Preventive Medicine*, *24*, 270–277.

Donaldson, S. I., Dent. C. W., Sussman, S., Stoddard, J. J., & Severson, H. H. (1996). The organizational implications of smokeless tobacco use in the Lumber Mill Industry. *Addictive Behaviors*, *21*, 259–267.

Donaldson, S. I., Ensher, E. A., & Grant-Vallone, E. J. (2000). Longitudinal examination of mentoring relationships on organizational commitment and citizenship behavior. *Journal of Career Development*, *26(4),* 233–249.

Donaldson, S. I., Gooler, L. E., & Weiss, R. (1998). Promoting health and well-being through work: Science and practice. In X. B. Arriaga & S. Oskamp (Eds.), *Addressing community problems: Psychological research and intervention* (pp. 160–194). Newbury Park: Sage.

Donaldson, S. I., Graham, J. W., Piccinin, A. M., & Hansen, W. B. (1995). Resistance-skills training and onset of alcohol use: Evidence for beneficial and potentially harmful effects in public schools and in private catholic schools. *Health Psychology*, *14*, 291–300.

Donaldson, S. I., & Klein, D. (1997). Creating healthful work environments for ethnically diverse employees working in small and medium-sized businesses: A non-profit industry/community/university collaboration model. *Employee Assistance Quarterly*, *13*, 17–32.

Donaldson, S. I., Thomas, C. W., & Graham, J. W. (2002). Understanding self-report bias in prevention research. Manuscript under review.

Donaldson, S. I., Thomas, C. W., Graham, Au, J., & Hansen, W. B. (2000). Verifying drug prevention program effects using reciprocal best friend reports. *Journal of Behavioral Medicine*, *23*, 221–234.

Donaldson, S. I., & Weiss, R. (1998). Health, well-being, and organizational effectiveness in the virtual workplace. In M. Igbaria, & M. Tan, *The virtual workplace* (pp. 24–44). Harrisburg, PA: Idea Group Publishing.

Ensher, E. A., Grant-Vallone, E. J., & Donaldson, S. I. (2001). Effects of perceived discrimination on job satisfaction, organizational commitment, organizational citizenship behavior, and grievances. *Human Resource Development Quarterly*, *12*, 53–72.

Fowler, F. Jr. (1995). *Improving survey questions: Design and evaluation*. Thousand Oaks, CA: Sage.

Graham, J. W., & Collins, N. L. (1991). Controlling correlational bias via confirmatory factor analysis of MTMM data. *Multivariate Behavioral Research*, *26*, 501–523.

Graham, J.W., Collins, N.L., Donaldson, S.I., & Hansen, W.B. (1993). Understanding and controlling for response bias: Confirmatory factor analysis of multitrait-multimethod data. In R. Steyer, K.F. Wender, & K.F. Widamen (Eds.), *Psychometric Methodology* (pp. 585–590). Stuttgart and New York: Gustav Fisher Verlag.

Howard, H. (1994). Why do people say nasty things about self-reports? *Journal of Organizational Behavior*, *15*, 399–404.

Karasek, R. (1985). Job content questionnaire. Department of Industrial systems Engineering, University of Southern California, Los Angeles.

Karasek, R., & Theorell, T. (1990). *Healthy work: Stress, productivity, and the reconstruction of working life*. New York: Basic Books.

Lifestyle Assessment Questionnaire (1988). *LAQ reliability/validity research efforts*. Steven Point, WI: National Wellness Institute.

McCall, M. M., & Bobko, P. (1990). Research methods in the service of discovery. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 381–418). Palo Alto, CA: Consulting Psychologists Press, Inc.

McDowell, I. & Newell, C. (1987). *Measuring health: A guide to rating scales and questionnaires*. New York: Oxford University Press.

Mersman, J. L., & Donaldson, S. I. (2000). Factors affecting the convergence of self-peer ratings on contextual and task performance. *Human Performance*, *13(3)*, 299–322,

Moorman, R. H., & Podsakoff, P. M. (1992). A meta-analytic review and empirical test of

the potential confounding effects of social desirability response sets in organizational behavior research. *Journal of Occupational and Organizational Psychology*, *65*, 131–149.

Podsakoff, P.M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, *12*, 531–544.

Sacket, P. R., & Larson, J. R. (1990). Research strategies and tactics in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 419–489). Palo Alto, CA: Consulting Psychologists Press.

Schmitt, N. (1989). Editorial. *Journal of Applied Psychology*, *74*, 843–845.

Schmitt, N. (1994). Method bias: the importance of theory and measurement. *Journal of Organizational Behavior*, *15*, 393–398.

Schwartz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*, 93–105.

Shadish, W. R. (1993). Critical multiplism: A research strategy and its attendant tactics. *New Directions for Program Evaluation*, *60*, 13–57.

Smith, J. C. (1993). *Understanding stress and coping*. New York: Macmillan.

Spector, P. E. (1987). Method variance as an artifact in self-reported affect and perceptions at work: Myth or significant problem? *Journal of Applied Psychology*, *3*, 438–443.

Spector, P. E. (1994). Using self-report questionnaires in OB research: a comment on the use of a controversial method. *Journal of Organizational Behavior*, *15,* 385–392.

Spector, P. E., & Brannick, M. T. (1995). The nature and effects of method variance in organizational research. In C. L. Cooper & I. T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (pp. 210–245). New York: Wiley & Sons.

Stone, A.A., Turkkan, J.S., Bachrach, C.A., Jobe, J.B., Kurtzman, H.S. & Cain, V.S. (2000). *The science of self-report: Implications for research and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Williams, L. J., & Brown, B. K. (1994). Method variance in organizational behavior and human resources research: Effects on correlations, path coefficients, and hypothesis testing. *Organizational Behavior and Human Decision Processes*, *58*, 185–209.

Williams, L. J., Cote, J. A., & Buckley, M. R. (1989). Lack of method variance in self-reported affect and perceptions at work: Reality or artifact? *Journal of Applied Psychology*, *74*, 462–468.

Wyler, A. R., Masuda, M., & Holmes, T. H. (1968). The Seriousness of Illness Rating Scale. *Journal of Psychosomatic Research*, *11*, 363–374.

Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review, 12*, 250–254.