Multivariate Behavioral Research, *39* (2), 301-326
Copyright © 2004, Lawrence Erlbaum Associates, Inc.

Two Replicable Suppressor Situations in Personality Research

Delroy L. Paulhus
University of British Columbia

Richard W. Robins, Kali H. Trzesniewski, and Jessica L. Tracy
University of California at Davis

Suppressor situations occur when the addition of a new predictor improves the validity of a predictor variable already in the equation. A common allegation is that suppressor effects rarely replicate and have little substantive import. We present substantive examples from two established research domains to counter this skepticism. In the first domain, we show how measures of guilt and shame act consistently as mutual suppressors: Adding shame into a regression equation increases the negative association between guilt and aggression, whereas adding guilt increases the positive association between shame and aggression. In the second domain, we show how the effects of self-esteem and narcissism operate consistently as mutual suppressors: That is, adding narcissism into a regression equation increases the negative association between self-esteem and antisocial behavior, whereas adding self-esteem increases the positive association between narcissism and antisocial behavior. Discussion addresses the different implications for suppressors in theoretical and variable selection applications.

Introduction

More than 60 years ago, Horst (1941) reported that multiple regression analyses occasionally display a curious, almost paradoxical pattern: A variable that is uncorrelated with the criterion can nonetheless improve prediction — that is, boost the value of R^2 — when it is added to the equation. Horst noted that this pattern occurred in cases where the new (apparently invalid) predictor was correlated with a predictor already in the equation. Horst labeled the new predictor a *suppressor variable*, reflecting the fact that it removes or suppresses criterion-irrelevant variance from the initial predictor.[1] With such irrelevant variance suppressed, the cleansed

The authors would like to thank the editors and the 12 reviewers for their thoughtful comments. Correspondence should be directed to Del Paulhus, Department of Psychology, University of British Columbia, Vancouver, Canada V6T 1Z4.

[1] Although agreeing on the phenomenon, some writers interpret the term "suppressor" somewhat differently as a variable that, before entering the equation, was suppressing the power of the prior predictor (Cohen, Cohen, West, & Aiken, 2003, p. 77; p. 457). In retrospect, the field might have been better off it Horst had used McFatter's (1979) term 'enhancement variable' instead.

variable becomes a more efficient predictor of the criterion. Put another way, the suppressor variable unleashes the latent predictive power of the initial predictor, thereby boosting its beta weight.

In a later article, Horst (1966) provided an example from a study of pilot selection conducted during World War II. To determine their suitability for training, applicants were given a variety of written tests. Not surprisingly, spatial ability predicted successful training but verbal ability did not. Yet verbal ability was correlated with spatial ability — presumably because verbal ability was necessary to follow the test questions measuring spatial ability. As a result, when verbal ability was forced into the regression equation, the total R^2 increased. Horst explained that verbal ability acted as a suppressor: That is,"...spatial test scores were discounted for those who did well on the test simply because of their verbal ability rather than because of abilities required for success in pilot training" (Horst, 1966, p. 355). This example is noteworthy because the theoretical explanation for the suppressor effect makes substantive sense.

Expanding the Definition

Early Extensions. Following Horst (1941), several articles provided empirical examples that satisfied his (rather strict) requirements for a suppressor variable (Lubin, 1957; Wherry, 1946). Eventually, however, the concept was extended beyond Horst's restricted version. Darlington (1968) broadened the definition to include variables whose final beta is opposite in sign to its original validity. The even-more-inclusive definition proposed by Conger (1974) as well as Cohen and Cohen (1975) simply required that the new predictor increase the weight of a predictor already in the equation.

Subsequently, Velicer (1978) suggested that suppressor variables be defined in terms of semi-partial correlations: A suppressor variable is one that that increases the semi-partial correlation of X_1 with Y (p. 314). Using this definition as a starting point, Holling (1983) further extended the concept of suppression, suggesting that it should apply to patterns of relations among multiple predictors and multiple criterion variables. The notion of a suppressor variable is retained, but it now applies to linear combinations rather than single variables. In the most general case, a criterion (Y) is predicted by a number of (standardized) variables that can be partitioned into two (mutually exclusive) sets. X_1 becomes the best linear combination of the current predictor set and X_2 the best linear combination of non-entered predictor variables. For example, X_1 could be a set of job performance predictors while X_2 is a set of response style measures. For simplicity, the

conditions for suppression can then be discussed in terms of X_1 and X_2 as linear combinations rather than as single variables.

Tzelgov and Henik. The most recent extension was provided by Tzelgov and Henik (1985; 1991), who recommended the term *suppression situations*. Their approach integrated the traditional three-variable analysis with Holling's (1983) suggested application of suppressor notions to the general linear model. Tzelgov and Henik noted two problems with the traditional term "suppressor variable." First, it cannot apply to the case of multiple predictors where a complex pattern of changes may ensue. Entering a new predictor may increase the weight of one previous predictor but decrease the weight of another. Simply pluralizing the term to "suppressor variables" does not solve the problem because suppressor effects may come and go during a series of single variable entries. To avoid the implication that several independent suppressors are being entered, Tzelgov and Henik concur with Holling's (1983) suggestion to create a single net suppressor that is a linear combination of variables. The second problem noted by Tselgov and Henik was that two predictors can be mutual suppressors, thereby muddying the distinction between suppressors and other predictors.

Following their lead, we will use the term "suppressor situations" and discuss multivariate cases as if there were only three variables (rather than three linear combinations of variables). Our working definition of a suppressor situation is one where the addition of a new predictor improves the validity of a predictor variable already in the equation.

Intervening Variable (IV) Models. The most recent conceptions of suppressor situations follow Davis (1985) in framing the phenomenon in terms of an implicit causal model set up by the inclusion of a second predictor in a regression equation (MacKinnon, Krull, & Lockwood, 2000; Shrout & Bolger, 2002). The impact of the initial predictor on the criterion is partitioned into a direct effect and an indirect effect via the new predictor. In this causal model, suppression is said to occur when the indirect effect of X_1 is opposite to the direct effect of X_1 (p.430).[2] In other words, the overall X_1 weight was being undermined by the indirect effect via X_2. When the opposing path is controlled by entering X_2 in the equation, then the direct effect of X_1 is seen to rise to a more genuine level. A distinct advantage to this reframing is that suppressor situations can be viewed as one of a set of patterns along with mediation and confounding (Mackinnon et al., 2000; Shrout & Bolger, 2002).

---

[2] The direct effect (c') is the X_1 beta when X_2 is included in the equation. The indirect effect is the product of the X_1 X_2 correlation (a) and the X_2 beta on Y with X_1 included in the equation (b).

Statistical Tests. The arrival of the elaborated intervening variable models has been accompanied by statistical advances. One simplification accompanied the demonstration that statistical tests for suppressor effects are part of a family of tests for intervening variables (e.g., MacKinnon et al., 2000). Specifically, the test for a significant suppressor effect (i.e., a significant increase in a regression weight) is equivalent to tests for a number of other statistical hypotheses, including confounds, spuriousness, and mediating variables (Baron & Kenny, 1986; MacKinnon et al., 2000; Sobel, 1986).

A comprehensive set of statistical tests based on normality assumptions was reviewed and evaluated by Mackinnon and colleagues (2002). Overall, they recommended: (a) testing significance by evaluating the product of the two indirect path weights (e.g., the Sobel test) and (b) testing significance of the change in regression coefficients. Based on simulations, the best balance of power and flexibility was obtained by using the *P* statistic, the product of the two associations (MacKinnon et al., 2002, p. 85). The formula for the standard error is provided by MacKinnon et al. (2000, p. 177).

An alternative approach based on bootstrap methods was recently developed by Shrout and Bolger (2002). Those authors recommend their use with small to moderate samples; for larger samples, the results are unlikely to differ from those reviewed by Mackinnon and colleagues. Given the sufficiently large samples in our datasets and the complexity of the bootstrap methods, we did not use the bootstrap method in our analyses.

The Case Against Suppressors

In his influential book, *Personality and Prediction*, Jerry Wiggins (1973) concluded: "the case for suppressor variables remains to be demonstrated" (p. 38). The impact of the book on a generation of personality researchers led them, quite rightly, to be skeptical about the value of suppressor variables.[3]

A Disappointing History. Wiggins's conclusion was based partly on Ghiselli's (1972) failure to replicate a suppressor effect, leading him to liken suppressor variables to the ephemeral "will-o-the wisp" (p. 270). Wiggins was particularly persuaded by the disappointing results reported by Goldberg, Rorer, and Green (1970). Since those earlier warnings, summary evaluations of the utility of suppressor variables in behavioral science have remained

[3] In many ways, the critque of suppressor variables parallels that of moderator variables (Chaplin, 1991).

guarded at best (e.g., Cohen & Cohen, 1992; Pedhazur, 1982). Even very recently, Maassen and Baker (2001) warned against devoting energy to formulating theoretical explanations for solitary suppression results.

Response Styles as Suppressor Variables. In personality psychology, the potential usefulness of identifying suppressor situations has been taken most seriously in the context of assessing response styles (e.g., Edwards, 1957; Jackson & Messick, 1958; Lubin, 1957; Paulhus, 1991). Response styles are typically defined as habitual tendencies to respond to questionnaire items on the basis of some item characteristic other than their truth value, for example, their desirability, keying direction, or extremity (Paulhus, 1991). Today, most standard personality questionnaires include subscales designed to tap response styles (e.g., the MMPI, the CPI, the 16PF, the MPQ, and the PRF), often labeled "validity scales" (for review, see Baer, Rinaldo, & Berry, 2003).

Of the wide variety of response styles, socially desirable responding (SDR) has elicited the greatest concern. The widely-shared fear is that the predictive power of personality scales will be undermined if some respondents show greater SDR tendencies than others (Butcher, Graham, & Ben-Porac, 1995; Holden, Kroner, Fekken, & Popham, 1992; Lanyon & Goodstein, 1989; Paulhus, 1991). To reduce this source of contamination, researchers have tried adding SDR measures to predictor equations, with the hope that SDR might act as a suppressor. Such attempts, however, have been largely disappointing (e.g., Archer, Fontaine, & McCrae, 1998; Borkenau & Ostendorf, 1992; Goldberg et al., 1970; Ones, Viswesvaran, & Reiss, 1996).[4]

Although other response styles (e.g., acquiescence, extreme responding) have received less attention, the few existing studies have failed to support their utility as suppressor variables (Wiggins, 1973). Indeed, even when significant, the effects are modest in size (Fulkerson, 1957; Watson & Tellegen, 2002). These discouraging findings should not be taken to imply that response biases cannot and should not be controlled before they occur: When possible they should be minimized with a priori methods such as anonymity (Paulhus, 1991). Least effective is the method of allowing the bias to flourish, then measuring it in the hope that it can be controlled later. In any case, our point here is that suppressor variables have a mixed reputation among assessment experts at least partly because of their disappointing history in the study of response styles.

[4] For an exception see Rosse, Stecher, Miller, and Levin (1998).

Three Suppression Configurations

Since Conger (1974), the range of suppressor situations has often been organized in terms of three possible configurations. To discuss them we use the standard three variable framework: That is, a criterion variable (*Y*) is regressed, at first on X_1 alone, and then jointly on X_1 and X_2. For simplicity, the predictors are scored so that they have a positive intercorrelation and so that X_1 has a positive validity. Without loss of generality, this pattern may require reversing the scoring of one or both of the predictors.

When either predictor is entered alone, its beta weight is simply its zero-order correlation with the criterion (i.e., its validity). An improvement in either validity when when the second variable is added to the regression equation indicates a suppression situation. The sign of the initial X_2 validity provides a way of distinguishing among three suppressor configurations. Each configuration requires a specific pattern of bivariate correlations whose impact can be illustrated by reference to Formula 1 for calculating beta weights (reverse the validity coefficients for the case of X_2).

$$\beta_1 = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{1 - r^2_{X_1X_2}} \tag{1}$$

Conceptually, this formula represents the proportion of the non-overlapping predictor variance explained by adding the indirect contribution of X_2 to the zero-order validity. The only way that this first-order validity can exceed its zero-order validity is to have: (a) a negative validity for X_2 and/or (b) a small denominator resulting from a large intercorrelation of the predictors.

Configuration 1: Suppressor Variable (X_2) has a Null Zero-order Validity

The appeal of the original Horst version — *classical suppression* — lay in its paradoxical nurturance effect: The original predictor (X_1) benefits from the entry of new predictor (X_2) that appears to have no predictive power. X_1 may already have attained some success despite its overlap with X_2 — a predictor composed entirely of variance irrelevant to the criterion. Forcing X_2 into the equation alleviates the handicap by removing the irrelevant variance: The 'true' efficacy of the original predictor is thereby unleashed. The removal of irrelevant variance is manifested in its negative beta weight for.

Configuration 2: Suppressor Variable (X_2) has a Negative Initial Validity

In configuration 2, the validities of the two predictors have opposite signs. Because the predictors are positively correlated, however, both validities are being handicapped. Including them together in a regression equation controls for the overlap and their mutual suppression is revealed by boosts in both regression weights. This case has been called *reciprocal suppression* (Conger, 1974) or *cooperative suppression* (Cohen & Cohen, 1975).

Configuration 3: Suppressor Variable (X_2) has a Positive Initial Validity

Here, the three variables are all positively correlated, suggesting a redundancy situation rather than a suppressor situation. Its true identity as a suppressor situation is not apparent until the regression weights are calculated and the smaller beta shows an opposite sign to its validity. This case was labeled *correction for distortion* by Rosenberg (1968), *negative suppression* by Conger (1974) and *net suppression* by Cohen and Cohen (1975). Below, we argue for the term *cross-over suppression* to describe an important subset of such cases.

Illustrative examples of the three configurations are provided in Figure 1. They differ only in the size and direction of the second predictor's validity. Examples (a), (b), and (c) represent the classical, cooperative, and net suppression configurations, respectively. For comparison, example (d) is the more common non-suppressor or *redundancy situation*, where positively correlated predictors have similar validities.

Note the interesting pattern of improvement in R^2 as the second predictor is added. All three suppressor situations account for more variance than the redundancy situation. In fact, the improvement in R^2 increases as the validity of the second predictor becomes more negative, that is, more discrepant from the validity of the first predictor. Together, these examples support Tzelgov and Henik (1991) in disputing the oft-heard recommendation that researchers should seek out orthogonal predictors to achieve maximal prediction (e.g., Thorndike, 1982). The examples also dispute the claim that addition of a new predictor variable is preferable to adding a suppressor (e.g., Jackson & Conger, 1972). In fact, prediction is best improved by adding a predictor positively correlated with a current one but having a validity opposite in sign.

Two Exemplary Substantive Domains

Like Collins and Schmidt (1997), our own research has led us to be more optimistic about the role of suppressors in the tangled nomological network of

personality. In the remainder of this article, we elaborate two substantive domains that involve replicable and theoretically coherent suppression effects. Both effects have been noted in earlier studies, but their theoretical relevance to our own research programs motivated us to evaluate their robustness. These two examples confirm that suppressor situations can be replicated and need to be taken seriously in both statistical and theoretical analyses.

The traditional appeal of suppressor variables lay in their surprising ability to contribute to predictive power. Three configurations were considered separately because, in each case, the suppressor variable began with different initial validity. In our two substantive domains, the distinction between predictor and suppressor is less important than a theoretical analysis of the variables involved in the suppressor situation.

In the first domain, we examined the joint effects of guilt and anxiety on self-reported aggression. Theoretical arguments suggested that they should

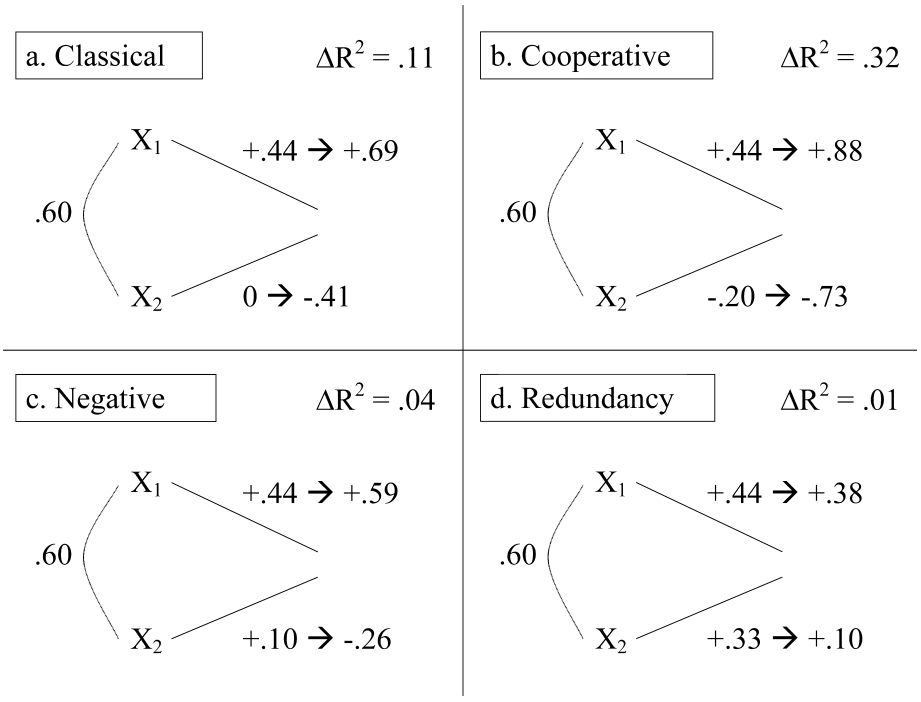

Figure 1
Illustrative Examples of Three Suppressor Situations Plus a Typical Redundancy Situation
The regression weights (Betas) are shown before and after entry of the second predictor.

show cooperative suppression. Here the distinction between predictor and suppressor variable vanishes and the suppression effects are symmetric. In the second domain, we attempted to clarify the association of self-esteem with antisocial behavior. We present datasets where the the association varies from positive to negative to null. The addition of a an overlapping variable, narcissism, helps stabilize the effect of self-esteem and confirms that its association with antisocial behavior is negative. The results suggest an equivalency of the three suppressor situations.

Domain 1: Predicting Aggression From Shame and Guilt

Tangney and her colleagues have conducted an extensive program of research on proneness to shame and guilt — two emotional dispositions that have proved to be less similar than researchers initially assumed. According to Tangney's theoretical model, as well as her empirical research, shame and guilt are distinct emotions with divergent antecedents, correlates, and consequences (Tangney & Dearing, 2002).

To facilitate her research, Tangney and her colleagues developed the Test of Self-Conscious Affect (TOSCA; Tangney, Wagner, & Gramzow, 1989), a scenario-based measure that includes separate scales for proneness to shame and guilt. However, Tangney typically does not use the raw scale scores to test her theoretically derived hypotheses about the distinct correlates of the two dispositions. Instead, she creates a "guilt-free" shame scale and a "shame-free" guilt scale by saving the standardized residuals from regression equations predicting shame from guilt and vice-versa. In many cases, Tangney has found that residualized shame and guilt are better predictors than the original variables. For example, in three different studies, "guilt-free" shame was more strongly positively associated with the SCL-90 Hostility/Anger scale than was the original shame scale. Similarly, "shame-free" guilt was more strongly negatively associated with the Hostility/Anger scale than was the original guilt scale (see Tangney & Dearing, 2002, Tables A.2 to A.4).

These findings are consistent with current theoretical conceptions of the two emotions. Although shame and guilt are both negative "self-conscious" emotions, shame is said to involve negative affect generalized to the global self, whereas guilt involves negative affect about a specific act or behavior (Barrett & Campos, 1987; Lewis, 1971; Tangney, 1990). This distinction is supported by evidence that shame is the more maladaptive of the two emotions (see Tangney & Dearing, 2002, for a review). Nonetheless, both emotions are "self-conscious," meaning that they result from an attentional focus on the self and internal attributions (self-blaming) for a negative event (Tracy & Robins, in press). Thus, despite their distinctive outcomes, shame and guilt often have

similar elicitors, and certainly share considerable variance (correlations typically range from .43 to .48; Tangney, Wagner, & Gramzow, 1992).

Nonetheless, shame and guilt have been found to produce highly divergent, and even opposite, outcomes: For example, guilt is *positively* related to various indicators of psychological well-being, including self-esteem, whereas shame is *negatively* related to such adaptive outcomes (Tangney & Dearing, 2002). One explanation for the contrasting effects of the two emotions can be found in the cognitive processes involved in the two emotions. Because shame involves negative feelings about the entire, global self, it promotes either painful negative views of the self or maladaptive coping to avoid such views. To regulate their feelings, shamed individuals typically turn these feelings outward through externalizations. If a negative event can be attributed to someone other than the self, the shamed individual can feel anger and hostility instead of shame, and become aggressive instead of accepting that his or her self may be flawed. Clinical psychologists have speculated that this process is at the root of the "shame-rage" cycle (Lewis, 1971; Scheff, Retzinger, & Ryan, 1989).

In contrast, guilt may provide an adaptive way of coping with one's bad behavior. By blaming the specific act that caused the problem instead of the global self, the guilty individual can accept responsibility while maintaining an intact self-concept (e.g., "I did a bad thing, but I'm not a bad person."). He or she can honestly acknowledge the mistake, and make efforts to redress the situation. As a result, guilt is associated more with empathy and perspective taking than with hiding, escaping blame, or anger and rage (Tangney & Dearing, 2002). These theoretical distinctions are supported by Tangney's demonstration that the shame-hostility correlation increases when guilt is partialed out, and that the guilt-hostility correlation becomes more negative when shame is partialed out.

If Tangney's arguments are correct, this suppressor pattern should extend to the externalization of hostility. In the present research, we evaluated possible suppressor effects with a somewhat different dependent variable — aggression — to evaluate the robustness of the finding that shame and guilt operate as reciprocal suppressors.

Method

Participants

Sample 1. A total of 238 undergraduate students (58% women) from the University of California (Davis) participated in exchange for course credit.

Sample 2. A total of 4111 undergraduate students (66% women) from the University of California (Davis) participated in exchange for course credit. The data were collected over a one-year period between Spring 2002 and Winter 2003.

Measures

Self-conscious Emotional Dispositions. Shame- and guilt-proneness were assessed with the Test of Self-conscious Affect (TOSCA; Tangney et al., 1989). The TOSCA consists of 15 brief scenarios (10 negative and 5 positive) and associated responses. These scenarios were drawn from written accounts of personal shame and guilt, and their participant-derived nature enhances the ecological validity of the measure. The validity of the measure has been confirmed by other researchers (Woien, Ernst, Patock-Peckham, & Nagoshi, 2003). Alpha reliabilities were .78 for proneness to shame and .61 for proneness to guilt in Sample 1, and .78 for shame and .78 for guilt in Sample 2.

Aggression. Participants in Sample 1 completed the full 75-item Buss-Durkee Aggression Questionnaire (AQ)(1957), and participants in Sample 2 completed a reduced 29-item version. Responses were measured on 5-point scales ranging from 1 "extremely uncharacteristic of me" to 5 "extremely characteristic of me." Both versions of the AQ include four subscales that are aggregated to form the total aggression score (Physical Aggression, Verbal Aggression, Anger, and Hostility). The AQ is the most widely used self-report measure of aggression and has been shown to predict behavioral measures of aggression (e.g., Anderson & Bushman, 1997). The alpha reliability of the AQ was .89 in Sample 1 and .91 in Sample 2.

Results and Discussion

Consistent with Tangney's theoretical model, guilt had significantly negative correlations with aggression: –.13 and –.19, in Samples 1 and 2, respectively. In contrast, shame had significantly positive correlations with aggression: .10 and .14 in Samples 1 and 2, respectively. Yet the two variables had positive intercorrelations in both Sample 1 (.46) and Sample 2 (.43). This pattern suggests a possible suppressor situation.

The hypothesis was tested in both samples by regressing aggression on self-esteem and narcissism — first one at a time and then jointly. In Table 1, the suppressor pattern can be seen clearly in both samples. We used the Sobel test to evaluate the significance of these suppressor effects

Table 1
Regressions of Self-reported Aggression on Guilt and Shame in Two Samples

			Guilt			Shame		
	N	r_{12}	β alone	β with shame	R^2 change	β alone	β with guilt	R^2 change
Sample 1	238	.46	–0.13*	–0.23*	.03	0.10	0.21*	.04
Sample 2	4111	.43	–0.19*	–0.30*	.06	0.14*	0.27*	.07

Note. r_{12} is the intercorrelation of shame and guilt.
*$p < .05$, two-tailed.

(MacKinnon et al., 2000).[5] The effect of shame on aggression increased when guilt was added to the equation (zs = 2.67 and –15.43 for Samples 1 and 2, respectively, ps < .05), and the effect of guilt on aggression increased when shame was added to the equation (zs = –2.95 and 13.92 for Samples 1 and 2, respectively, ps < .05). This pattern exemplifies the cooperative suppression configuration detailed in the introduction.

In short, the mutual suppressor effects of guilt and shame, previously found by Tangney and her colleagues, were replicated in both samples. What underlying cognitive and emotion processes could explain this suppressor relationship? As noted earlier, Tangney's work has focused primarily on disentangling shame and guilt, uncovering their unique outcomes and elicitors (e.g., Niedenthal, Tangney, & Gavanski, 1994). This partitioning is consistent with a functionalist, discrete-emotions perspective: That is, if the two emotions produced the same set of outcomes, there would be little reason for both to exist in the human repertoire. The distinctiveness of their outcomes indicates a distinctiveness of function and refutes the frequent confusion of guilt with shame in common parlance.

However, Tangney's measures share substantial variance — a consistent empirical finding undoubtedly linked to that fact that shame and guilt are both negative self-conscious emotions. Indeed, a recent extension of Tangney's work by Tracy and Robins (in press) proposed a theoretical model of self-conscious emotions holding that self-awareness is at the core of both shame

[5] We also tried several of the other methods listed by Mackinnon et al. (2002) and found similar results.

and guilt, and is the crucial starting point for the elicitation of either emotion. Thus, the shared variance in shame and guilt, which is removed when both dimensions are entered into the regression equation predicting aggression, can be conceived as generalized negative self-consciousness. This general tendency to focus on the self and feel negatively about it may be an important, under-examined variable with its own predictive validity; Alternatively it may be simply a statistical or response style artifact.

More clear is what happens to the two measures when this overlapping variance is removed. Shame-free guilt becomes a highly adaptive, pro-social disposition, negatively related to aggression and anger. This is the "moral" emotion that Tangney and her colleagues have argued should be instilled in children from an early age, and might even provide a solution to the dilemma of recidivism (Tangney & Dearing, 2002). Guilt-free shame, in contrast, is a maladaptive, hostility-inducing disposition, leading to externalization, avoidance of blame, and, eventually, depression. Our results demonstrate how the unique effects of the two emotions can be clarified by exploiting their theoretically-predictable, highly-replicable suppressor relationship whenever their outcomes and correlates are being examined.

Domain 2: Predicting Antisocial Behavior From Self-Esteem and Narcissism

Does high self-esteem promote or discourage antisocial behavior? The direction of the link between self-esteem and antisocial outcomes, such as aggression and delinquency, has recently become the focus of considerable debate among social scientists as well as the popular media. One prominent view posits that low self-esteem promotes antisocial behavior, and therefore increasing the self-esteem of antisocial individuals should reduce their externalizing behaviors (e.g., Branden, 1994; Mecca, Smelzer, & Vasconcellos, 1989; Trzesniewski, Donnellan, Robins, & Paulhus, 2003). The opposing view disputes the alleged benefits of high self-esteem and argues instead that self-esteem has a "dark side,": Individuals with high self-esteem are prone to aggression and other forms of antisocial behavior (Baumeister, Smart, & Boden, 1996).

Rather than arguing for or against self-esteem, other writers have advocated a distinction between its adaptive and maladaptive forms. Among such distinctions are genuine vs. non-genuine forms (Block & Thomas, 1955; Cohen, 1962; Hewitt & Goldman, 1972), self-esteem vs. self-deception (Paulhus, 1984), genuine vs. fragile self-esteem (Baumeister, Campbell, Kreuger, & Vohs, 2003; Campbell, 2000), defensive vs. non-defensive self-esteem (Salmivalli, Kaukianen, Kaistaniemi, & Lagerspetz, 1999), and self-

esteem vs. narcissism (Robins & John, 1997; Tracy & Robins, 2003). These distinctions all have a similar flavor yet show distinctive features (Kurt, Campbell, & Paulhus, 2003). Here, we refer to the adaptive form of high self-esteem as "genuine" and we operationalize it with scores on the Rosenberg Self-Esteem scale controlled for narcissism. We refer to the maladaptive form as "narcissistic" and operationalize it with the scores on the Narcissistic Personality Inventory controlled for self-esteem.

Our reading of this literature suggests that genuine self-esteem is negatively related to antisocial behavior, whereas narcissistic self-esteem is positively related. The notion that the two aspects of self-esteem have such conflicting associations may explain the disconcerting inconsistency found in reviews of the relation between self-esteem and aggression (Baumeister et al., 2003).

Conceptualizing the link between self-esteem and antisocial behavior as a suppressor relationship suggests a testable hypothesis: If measures of self-esteem tap both genuine self-esteem and narcissism, then partialing out variance related to narcissism should decontaminate the measure of self-esteem. The residual should show a clear negative association with antisocial behavior (reflecting the effect of "genuine" self-esteem). In contrast, partialing self-esteem out of narcissism should increase its (already) positive relation with antisocial behavior. This pattern of findings has already been reported in three published studies (Paulhus, 1998; Smalley & Stake, 1996; Twenge & Campbell, 2002), but none of those researchers predicted the finding. Based on that research and the above theoretical arguments, we hypothesize a mutual suppressor effect between measures of self-esteem and narcissism in predicting antisocial behavior.[6]

We tested this hypothesis in three samples. Sample 1 examined self-reported aggressive behaviors in a sample of U.S. college students ($N = 4057$). Sample 2 examined anonymous self-reports of antisocial behaviors (e.g., bullying, crime, drug use) in a sample of Canadian college students ($N = 301$). Finally, Sample 3 examined peer ratings of antisocial behavior in a sample of Canadian college students ($N = 232$).

Method

Participants

Sample 1. A total of 4057 undergraduate students (66% women) from the University of California (Davis) participated in exchange for course credit.

[6] A similar pattern was found in predicting antisocial behavior from anxiety and callousness (Frick, Lilienfeld, Ellis, Loney, & Silverthorn , 1999).

The data were collected over a one-year period between Spring 2002 and Winter 2003. This sample was drawn from the same larger sample used for Sample 2 of the shame and guilt analyses.

Sample 2. A total of 301 undergraduate students (62% women) from the University of British Columbia participated in exchange for course credit.

Study 3. A total of 232 undergraduate students (64% women) from the University of British Columbia participated in exchange for a lottery ticket.

Measures

Self-esteem. Participants completed the Rosenberg Self-Esteem scale (RSE; Rosenberg, 1965). The RSE consists of 10 items assessing global self-esteem (e.g., "On the whole, I am satisfied with myself"); items were rated on a 5-point scale ranging from 1 "not very true of me" to 5 "very true of me." The alpha reliability of the RSE was .90 in Sample 1, .90 in Sample 2, and .86 in Sample 3.

Narcissism. The 40-item forced-choice version of the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1981) was used to assess narcissism. One example item asks participants to select between the options "I am going to be a great person" and "I hope I am going to be successful". The NPI is the best-validated self-report measure of overt narcissism for non-clinical populations, and has been shown to predict psychologists' ratings of narcissism (e.g., John & Robins, 1994). The alpha reliability of the NPI was .84 in Sample 1, .84 in Sample 2, and .80 in Sample 3.[7]

Antisocial Behavior. In Sample 1, antisocial behavior was operationalized by the Buss-Perry Aggression Questionnaire (AQ; Buss & Perry, 1992, modified from Buss & Durkee, 1957). More details on the measure are provided above in the shame-guilt Sample 2. The alpha reliability of the AQ in our sample was .91.

In Sample 2, antisocial behavior was operationalized with an instrument designed to tap a much broader range of misbehavior. Participants completed the Comprehensive Misbehavior Inventory (CMI; Paulhus & Williams, 2002). This 134-item modification of the Elliot-Ageton instrument was designed to measure a variety of antisocial behaviors in non-criminal

[7] Note that differential reliabilities of predictors can have a negative effect on finding and interpreting suppressor effects. Fortunately, the reliabilities of RSE and NPI are comparable as are the shame and guilt scales used in the first substantive domain.

samples. The items cluster into six factors that are aggregated to form the total antisocial score (crime, bullying, drug or alcohol abuse, dangerous driving, and anti-authority behaviors). Although self-report in nature, the validity of the instrument has been supported by convergence with actual behavior and peer-ratings (Williams, McAndrew, Harms, & Paulhus, 2001). The alpha reliability of the total score was .85.

In Sample 3, two informants (mostly friends and family members) rated each participant on four antisocial dimensions, including "Starts fights with others" and "Has physically bullied someone". The ratings were made on a 5-point scale. Raters were recruited by targets, who received bonus course marks by providing names and email addresses of two people familiar with them and willing to evaluate them. Raters received the questions via email and returned them the same way. The inter-rater reliability of the 4-item composite scores (.84) was estimated by alpha (see Shrout & Fleiss, 1979).

Results and Discussion

Across the three studies, narcissism had positive correlations with anti-social behavior. In contrast, self-esteem showed negative or slightly positive correlations. Yet, self-esteem and narcissism were positively correlated in all three samples: In Sample 1; $r = .32$, $p < .05$, Sample 2; $r = .48$, $p < .05$, and Sample 3; $r = .43$, $p < .01$.

In each sample, antisocial behavior was regressed on self-esteem alone, then narcissism alone, and then jointly on both predictors. In all three samples, a suppressor pattern was observed. Again, we used the Sobel test to evaluate the significance of these suppressor effects. In all three samples, we found that the negative effect of self-esteem on antisocial behavior was significantly stronger after narcissism was included in the equation than when it was not (zs = 15.37, 5.20, and 7.61, respectively, ps < .05). In contrast, the positive effect of narcissism on antisocial behavior was significantly stronger after self-esteem was included in the equation than when it was not (zs = 15.92, 2.07, and 7.01, respectively, all ps < .05).

Note that the direction of the original self-esteem validity was positive in one case (+.12, $p < .05$), negative in another case (–.27, $p < .05$), and virtually zero in another case (.02, *ns*). Yet, the final effect was always significantly negative. Together, the regression analyses show a robust *negative* association between genuine self-esteem and anti-social behavior. Several previous studies have found similar patterns, but the pattern was neither predicted nor interpreted (Campbell, Foster, & Finkel, 2002; Paulhus, 1998; Smalley & Stake, 1996; Twenge & Campbell, 2003). The present

results, combined with those previous studies, confirm a robust suppressor situation involving self-esteem, narcissism, and antisocial behavior.

The finding that self-esteem and narcissism suppress some of each other's variance has a variety of theoretical implications. With regard to the debate about the link between self-esteem and antisocial behavior, the results argue strongly that self-esteem in its uncontaminated form is negatively associated with antisocial behaviors. Samples where it appears otherwise are a misleading consequence of the fact that scores on Rosenberg's self-esteem scale tap a heterogeneous amalgam, which includes some contamination from narcissism. Some individuals report high self-esteem for reasons of narcissistic enhancement. Accordingly, researchers should consider removing narcissism when investigating relations between self-esteem and other variables. The interpretation of a purified narcissism measure is less clear. Among the possibilities are that this residual captures pure bias or disagreeableness (Kurt et al., 2003).

The finding of mutual suppression raises another possibility: Perhaps there is a legitimate personality component captured by the shared variance between the two predictors, the removal of which enhances the predictability of both. Following McCrae and Costa (1995) as well as Paulhus and John (1998), the term "self-favorability" is available to refer to either (a) the shared variance between self-esteem and narcissism or (b) the superordinate concept that includes both forms of positive self-views. It is possible that self-favorability predicts important outcomes, such as job application success, where self-confidence is important but the source of positive self-evaluations — whether from an inflated ego or genuine feelings of self-worth — may not matter.[8] It is tempting to speculate that self-esteem is simply the realistic and narcissism the unrealistic component of self-favorability: But recent evidence suggests that both involve some discrepancy with reality (Paulhus, Harms, Bruce, & Lysy, 2003). Finally, it is also possible that the overlap between self-esteem and narcissism is merely artifactual, and cannot be isolated.

Although the two variables overlap in self-favorability, a growing body of literature indicates that their distinctiveness turns on the adaptiveness of their outcomes (e.g., Morf & Rhodewalt, 1999). This divergence may be the source of their independent contributions to antisocial behavior: Self-esteem controlled for self-favorability becomes a pro-social orientation (perhaps it is self-confidence purified of desire to show it off), whereas narcissism controlled for self-favorability becomes anti-social and aggressive (perhaps because underlying feelings of insecurity come closer to the surface).

[8] However, it is also likely that over time the job trajectories of individuals with inflated egos (i.e., narcissists) and those with genuine feelings of self-worth will diverge (Robins & Paulhus, 2001).

Summary. The present findings help clarify the inconsistent relations between self-esteem and antisocial behavior previously reported in the literature (e.g., Baumeister et al., 2003). The present research demonstrates that, when controlled for narcissistic tendencies, self-esteem is negatively related to anti-social tendencies whether judged by the self or by others.

General Discussion

We began by noting the longstanding skepticism about suppressor variables, articulated for a generation of personality researchers by Jerry Wiggins (1973). The traditional grounds for skepticism were two-fold: (a) the failure of suppressors to replicate in new samples and (b) the weak theoretical rationale for predicting or even conceptualizing these effects. We concur with Wiggins that unreplicated suppressor findings do not warrant the attention they are often given. Considerable intellectual energy has been wasted in attempts to explain statistical flukes that failed to reappear in subsequent samples.

Our intent here was to provide some empirical assurance that not all suppressor situations are chance findings of interest only as curiosities: We have highlighted two domains in personality research where suppressor effects are theoretically meaningful and replicable across diverse methods and samples.[9] Theoretical explanations of these suppressor effects can now be addressed without concern they will ultimately disappoint like the "will-o-the-wisp" in Ghiselli's (1972) characterization.

Our introduction reviewed the three configurations typically distinguished in the literature on suppressor variables: (a) classical suppression, (b) cooperative (reciprocal) suppression, and (c) net (negative) suppression. To cover all three configurations and contexts in which they occur, we were persuaded to follow Tzelgov and Henik (1981) in using the term "suppressor situation" rather than "suppressor variable" per se.

Our regressions with guilt and shame as joint predictors clearly fall into the cooperative category because both initial validities were magnified. Our two samples replicated Tangney's suppressor pattern thereby supporting her model of the dynamics of shame and guilt. Our analysis of self-esteem and narcissism was more complex but provided further insight into issues surrounding suppressor situations. In that domain, each of our three samples represented a different configuration of suppressor situations. The over-riding theoretical concern was with the direction of the validity of one predictor, namely, self-esteem. Regression results showed that, whether the

[9] Previously, Collins and Schmidt (1997) were able to cross-validate a number of (non-predicted) suppressors by splitting a single sample.

original validity was positive, negative, or zero, the final validity was always negative. The robustness of that final validity provides special assurance that (genuine) self-esteem bears a negative relation with anti-social behavior.

A Heuristic Definition

Our review of the literature indicated a near consensus that all three configurations should be included under the generic definition of suppressor situation.[10] We agree that all three satisfy the requirement that (at least) one predictor becomes stronger after joint inclusion. But we argue further that all suppressor situations are ultimately of one form, namely, mutual suppression.[11] Whether the added variable has a validity that is positive, negative, or precisely zero is of incidental importance. In all suppression situations, two positively correlated predictors show more discrepant regression weights after they are included together in the same equation. This common pattern can be seen by looking back at Figure 1: In the three suppressor configurations — (a), (b), and (c) — the weights become more discrepant and end up having opposite signs. This mutual repulsion effect makes intuitive sense because, before controlling for their intercorrelation, the overlap of the predictors was forcing the validities to converge more than warranted by their direct, independent effects. Joint inclusion in a regression equation controls for that overlap and yields independent residuals that now have more discrepant associations with the criterion.

Reconsideration of our second substantive domain may help persuade skeptical readers. We measured the same three variables — self-esteem, narcissism and antisocial behavior — in three large samples. Yet the three patterns of associations differed such that each sample represented one of the three suppressor configurations. The differences can be viewed simply as sampling variation in the correlation between self-esteem and narcissism. Because Sample 3 involved the largest correlation of the three samples, the initial validities were especially misleading. This claim is evidenced by the striking similarity of the final results of the three regressions. Clearly the underlying relations varied little: They had simply been distorted in different ways by sampling variation. Moreover, a look at Table 2 reveals that mutual repulsion operates in all three cases.

Our arguments can be couched in the language of the intervening variable models elaborated by MacKinnon and colleagues (2000) as well as

[10] Exceptions to this consensus include Darlington (1962), Velicer (1978) and, most recently, Smith et al. (1992).

[11] Here, we are not referring to 'reciprocal suppression', the synonym for cooperative suppression coined by Cohen and Cohen (1975).

Table 2
Regressions of Antisocial Behavior on Self-esteem and Narcissism in Three Samples

			Self-esteem			Narcissism		
	N	r_{12}	β alone	β with narcissism	R^2 change	β alone	β with self-esteem	R^2 change
Sample 1	4057	.32	–.27*	–.38*	.10	.21*	.32*	.09
Sample 2	301	.44	.02	–.16*	.13	.33*	.40*	.02
Sample 3	232	.50	.12	–.14*	.20	.45*	.52*	.01

Note. r_{12} is the intercorrelation of self-esteem and narcissism.
*$p < .05$, two-tailed.

Shrout and Bolger (2002). For example, our finding that narcissism acts as a suppressor for self-esteem can be seen as the result of inconsistency between the direct and indirect paths from self-esteem to antisocial behavior. At this point in our research on this topic, we hesitate to interpret this pattern as a causal model although a number of such models are possible. For example, some of the individuals reporting high self-esteem are actually engaging in narcissistic self-promotion. Our use of self-report questionnaires may have created an overlap due to common method variance. Alternatively, high self-esteem may have promoted narcissism in a subset of individuals without conferring the requisite modesty or social skills: In other words, this group may possess "unmitigated self-esteem." The data available here are not sufficient to test these competing explanations.

The Case of Cross-Over Suppression

Of particular import is the analysis wherein the addition of the theoretical suppressor variable (narcissism) forced a re-direction in the validity of the original predictor (self-esteem). Clearly the valence of the original validity was misleading: The joint regression ultimately yielded a negative effect of self-esteem that is consistent with the other two samples.

The appeal of such cases of negative suppression derives from the fact that, superficially, they resemble standard redundancy situations. Although all the intercorrelations are positive and therefore algebraically consistent, this structural harmony is illusory. The inconsistent pattern of their betas reveals the true instability or statistical tension. One characterization is that the two

positive validities are not as similar as they should be given their high positive intercorrelation (see Chandler, 1961). Therefore some (reversed) overlapping factor must be pushing them apart. Alternatively, the unstable structure may be described as one where the indirect effect is so strong that it normally overwhelms the direct effect. Only when the indirect path is controlled for does the (more informative) direct effect emerge. Either way, this configuration is the only case where a positive manifold yields suppression.

Previous discussions of suppression of the third kind have focused on the situation where the suppressor validity changes sign (e.g., Darlington, 1968). In our empirical example, it is the initial predictor, self-esteem, that changes sign. The predictor of primary theoretical concern reverses direction thereby supporting a theoretical claim. We recommend the term "cross-over suppression" to describe the case where the addition of a new predictor reverses the direction of a previous predictor.

This cross-over phenomenon may have broader implications for analytic strategy. Recall the Shrout and Bolger (2002) argument that one should proceed with a causal analysis of a theoretical relation even when the association of the initial predictor is zero (p. 438). We push that argument further to argue for a causal analysis even when the initial validity is significant in the wrong direction, that is, it contradicts theoretical assumptions.

Expanding the Search

Researchers may overlook suppressor effects on a regular basis because of the algorithm typically employed for selecting variables in standard regression programs. At each step of stepwise regression, the remaining predictor with the highest validity is added. This forward inclusion procedure will never include a suppressor variable with a small validity. Accordingly, we recommend that backward elimination be used instead. This method will not drop a suppressor variable that contributes significantly to the explained variance.

When a regression program does enter a suppressor variable (perhaps because it was forced in), a unique situation can be observed: The last entered variable may boost R^2 more than did any previously entered variable. The sequence is odd because, normally, the order of entry bears a close correspondence with the strength of the original validity. In a suppressor situation, a surprisingly large increase may be observed despite a weak original validity.

In sum, we concur with Tzelgov and Henik (1991) as well as Collins and Schmidt (1997) that the number of genuine suppressor situations in behavioral science may be far greater than has been assumed and a more vigorous search for such effects is warranted.

Variable-Selection Situations vs. Theoretical Situations

Our arguments might be clarified by making a distinction between exploratory variable-selection situations and theoretical model testing. In selection situations, the analyst often seeks the best set of predictors for a given sample: Here the identification of optimal predictors is paramount.[12] The traditional use of forward selection algorithms is understandable, if not advisable. The goal is an efficient set of predictors, that is, the fewest that can produce an adequate amount of variance in the criterion. In theoretical contexts, suppressors play a rather different role. The focus is on one or two predictors of theoretical interest, not on the search for optimal predictors. Other available variables may have higher validities than the predictor of interest. Even a predictor with a null or an apparently reversed validity can be the theoretical focus when a clear theory of inconsistent mediation is available (Shrout & Bolger, 2002).

Our cross-over example from above is a case in point. The addition of self-esteem to a regression of antisocial behavior on narcissism produced very little benefit to prediction: In fact, the R^2 change was only .01! But the theoretical harvest was substantial. The direction of the self-esteem's validity was thereby reversed and came into accord with previous research including our two other samples.

As a general rule, theoretical contexts are more likely to lead to the examination of interactions and causal relations among predictors. Hence the new analytic tools for analyzing intervening variables are especially relevant and likely to be central to suppressor analyses in the future. We go further to argue that the search for replicable suppressor effects is most likely to be fruitful when guided by well-articulated theoretical models.

References

Anderson, C. A. & Bushman, B. J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. *Review of General Psychology, 1*, 19-41.

Archer, R. P., Fontaine, J., & McCrae, R. R. (1998). Effects of two MMPI-2 validity scales on basic scale relations to external criteria. *Journal of Personality Assessment, 70*, 87-102.

Baer, R. A., Rinaldo, J. C., & Berry, D. T. R. (2003). Response distortions in self-report assessment. In S. Ballesteros (Ed.), *Encyclopedia of Psychological Assessment* (pp. 132-133). London: Sage Publications.

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator distinction in social-psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

[12] Although not always possible, selection experts know that cross-validation of an assessment battery is preferable.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest, 4*, (Whole number 1)

Block, J. & Thomas, H. (1955). Is satisfaction with self a good measure of self-esteem? *Journal of Abnormal and Social Psychology, 51*, 254-259.

Borkenau, P. & Ostendorf, F. (1992). Social desirability scales as moderator and suppressor variables. *European Journal of Personality, 6*, 199-214.

Branden, N. (1994). *The six pillars of self-esteem*. London: Bantam Books.

Buss, A. H. & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 31*, 343-348.

Butcher, J. N., Graham, J. R., & Ben-Porath, Y. S. (1994). Methodological problems and issues in MMPI, MMPI-2, and MMPI-A research. *Psychological Assessment, 7*, 320-329.

Chaplin, W. F. (1991). The next generation of moderator research in personality psychology. *Journal of Personality, 59*, 143-178.

Chandler, R. E. (1961). Two additional formulae for use with suppressor variables. *Educational and Psychological Measurement, 21*, 947-950.

Cohen, J. & Cohen, P. (1975). *Applied multiple correlation/regression analysis for the social sciences.* New York: Wiley.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple correlation/ regression analysis for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Collins, J. M. & Schmidt, F. L. (1997). Can suppressor variables enhance criterion-related validity in the personality domain? *Educational and Psychological Measurement, 57*, 924-936.

Conger, A. J. (1974). A revised definition of suppressor variables: A guide to their identification and interpretation. *Educational and Psychological Measurement, 34*, 35-46.

Conger, A. J. & Jackson, D. J. (1972). Suppressor variables, prediction, and the interpretation of psychological relationships. *Educational and Psychological Measurement, 32*, 579-599.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin, 69*, 161-182.

Davis, J. A. (1985). *The logic of causal order.* Thousand Oaks, CA: Sage Press.

Dicken, C. (1963). Good impression, social desirability, and acquiescence as suppressor variables. *Educational and Psychological Measurement, 23*, 699-720.

Fricke, B. G. (1956). Response set as a suppressor variable in the OAIS and MMPI. *Journal of Consulting Psychology, 20*, 161-169.

Frick, P. J., Lilienfeld, S. O., Ellis, M. L, Loney, B. R., & Silverthorn, P. (1999). The association between anxiety and psychopathy dimensions in children. *Journal of Abnormal Child Psychology, 27*, 381-390.

Ghiselli, E. E. (1972). Comment on the use of moderator variables. *Journal of Applied Psychology, 56*, 270.

Ghiselli, E. E., Campbell, J. P. & Zedeck, S. (1981). *Measurement theory for the behavioral sciences.* San Francisco: Freeman.

Goldberg, L. R., Rorer, L. G. & Green, M. M. (1970). The usefulness of "stylistic" scales as potential suppressor or moderator variables in predictions from the CPI. *Oregon Research Institute Research Bulletin, 10*, (3).

Henik, A. & Tzelgov, J. (1985). Control for halo error: A multiple regression approach. *Journal of Applied Psychology, 70*, 577-580.

Hewitt, J. & Goldman, M. (1974). Self-esteem, need for approval, and reactions to personal evaluations. *Journal of Experimental Social Psychology, 10*, 201-210.

Holden, R. R., Kroner, D. G., Fekken, G. C., & Popham, S. M. (1992). A model of personality test item response dissimulation. *Journal of Personality and Social Psychology, 63*, 272-279.

Holling, H. (1983). Suppressor structures in the general linear model. *Educational and Psychological Measurement, 43*, 1-9.

Holzbach, R. (1978). Rater bias in performance ratings: Superior, self and peer ratings. *Journal of Applied Psychology, 63*, 579-588.

Horst, P. (1941). The role of the predictor variables which are independent of the criterion. *Social Science Research Council, 48*, 431-436.

John, O. P. & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, *66*, 206-219.

Krus, D. J. & Wilkinson, S. M. (1986). Demonstration of properties of a suppressor variable. *Behavior Research Methods, Instruments, and Computers, 18*, 21-24.

Kurt, A., Campbell, J. D., & Paulhus, D. L. (2003). *Distinguishing four theories of insecure self-esteem*. Unpublished manuscript, University of British Columbia.

Lanyon, R. I. & Goodstein, L. D. (1997). *Personality assessment* (3rd ed.). Oxford, England: Wiley.

Lewis, H. B. (1971). *Shame and guilt in neurosis*. New York: International Universities Press.

Lubin, A. (1957). Some formulae for use with suppressor variables. *Educational and Psychological Measurement, 17*, 286-296.

Maassen, G. H. & Bakker, A. B. (2001). Suppressor variables in path models: Definitions and interpretations. *Sociological Methods and Research*, *30*, 241-270.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science, 1*, 173-181.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83-104.

McFatter, R. M. (1979). The use of structural equation models in interpreting regression equations including suppressor and enhancer variables. *Applied Psychological Measurement, 3*, 123-135.

McNemar, Q. (1969). *Psychological statistics.* New York: Wiley.

Mecca, A. M., Smelser, N. J., & Vasconellos, J. (Eds.). (1989). *The social importance of self-esteem*. Berkeley, CA: University of California Press.

Meehl, P. E. (1945). A simple algebraic development of Horst's suppressor variables. *American Journal of Psychology, 58*, 550-554.

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

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp.17-59). New York: Academic Press.

Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing*? Journal of Personality and Social Psychology, 74,* 1197-1208.

Paulhus, D. L. (2001). Narcissism: Two minimalist accounts. *Psychological Inquiry, 11*, 222-224.

Paulhus, D. L. & John, O. P. (1998). Egoistic and moralistic bias in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality, 66*, 1025-1060.

Paulhus, D. L. & Williams, K. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556-568.

Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology, 84*, 681-693.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction.* New York: Holt, Rinehart & Winston.

Raskin, R. R. & Hall, C. S. (1981). The Narcissistic Personality Inventory: Alternative form reliability and further evidence of construct validity. *Journal of Personality Assessment, 45*, 159-162.

Robins, R. W. & John, O. P. (1997). The quest for self-insight: Theory and research on accuracy and bias in self-perception. In R. T. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 649-679). New York: Academic Press.

Robins, R. W. & Paulhus, D. L. (2001). The character of self-enhancers: Implications for organizations. In B. W. Roberts & R. Hogan (Eds.), *Personality psychology in the workplace* (pp. 193-222). Washington, DC: American Psychological Association.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rosenberg, M. (1968). *The logic of survey analysis*. New York: Basic Books.

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology, 83*, 634-644.

Salmivalli, C., Kaukianen, A., Kaistaniemi, L., & Lagerspetz, K. M. J. (1999). Self-evaluated self-esteem, peer-evaluated self-esteem, and defensive egotism as predictors of adolescents' participation in bullying situations. *Personality and Social Psychology Bulletin, 25*, 1268-1278.

Scheff, T. J., Retzinger, S. M., & Ryan, M. T. (1989). Crime, violence, and self-esteem: Review and proposals. In A. M. Mecca, N. J. Smelser, & J. Vasconcellos (Eds), *The social importance of self-esteem* (pp. 165-199). Berkeley, CA: University of California Press.

Shrout, P. E. & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Bulletin, 7*, 422-445.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Smalley, R. L. & Stake, J. E. (1996). Evaluating sources of ego-threatening feedback: Self-esteem and narcissism effects. *Journal of Research in Personality, 30*, 483-495.

Smith, R. L., Ager, J. W., & Williams, D. L. (1992). Suppressor variables in multiple regression/ correlation. *Educational and Psychological Measurement, 52*, 17-28.

Sobel, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. In N. Tuma (Ed.), *Sociological methodology* (pp. 159-186). Washington, DC: American Sociological Association.

Tangney, J. P. & Dearing, R. L. (2002). *Shame and guilt*. New York: Guilford.

Tangney, J. P., Wagner, P., & Gramzow, R. (1989). *The Test of Self-Conscious Affect: TOSCA*. Unpublished measure, George Mason University, Fairfax, VA.

Tangney, J. P., Wagner, P., & Gramzow, R. (1992). Proneness to shame, proneness to guilt, and psychopathology. *Journal of Abnormal Psychology, 101*, 469-478.

Tangney, J. P., Wagner, P. E., Hill-Barlow, D., Marschall, D. E., & Gramzow, R. (1996). Relation of shame and guilt to constructive versus destructive responses to anger across the lifespan. *Journal of Personality and Social Psychology, 62*, 669-675.

Thorndike, R. L. (1982). *Applied psychometrics*. Boston: Houghton Mifflin.

Tracy, J. L. & Robins, R. W. (2003). "Death of a [Narcissistic] Salesman": An integrative model of fragile self-esteem. *Psychological Inquiry, 14*, 57-62.

Tracy, J. L. & Robins, R. W. (in press). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*.

Trzesniewski, K. H., Donnellan, M. B., Robins, R. W., & Paulhus, D. L. (2003). *The antisocial individual: The negative consequences of low self-esteem and narcissism.* Poster presented at meeting of the Society for Personality and Social Psychology.

Twenge, J. M. & Campbell, W. K. (2003). Isn't it fun to get the respect we're going to deserve? *Personality and Social Psychology Bulletin, 29*, 261-272.

Tzelgov, J. & Henik, A. (1985). A definition of suppression situations for the general linear model: A regression weights approach. *Educational and Psychological Measurement, 45*, 281-284.

Tzelgov, J. & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin, 109*, 524-536.

Velicer, W. F. (1978). Suppressor variables and the semipartial correlation coefficient. *Educational and Psychological Measurement, 38*, 953-958.

Watson, D. & Tellegen, A. (2002). Aggregation, acquiescence, and the assessment of trait affectivity. *Journal of Research in Personality, 36*, 589-597.

Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. New York: Addison-Wesley.

Williams, K. M., McAndrew, A., Harms, P. D., & Paulhus, D. L. (1999). *Predicting antisocial behavior from the dark triad of personality*. Presented at the meeting of the American Psychological Association, San Francisco.

Woien, S. L, Ernst, H. A. H., Patock-Peckham, J. A., & Nagoshi, C. T. (2003). Validation of the TOSCA to measure shame and guilt. *Personality and Individual Differences, 35,* 313-326.