Defensive Verbal Behavior Assessment

Lisa Feldman Barrett
Boston College

Nathan L. Williams
George Mason University

Geoffrey T. Fong
University of Waterloo

The present investigation introduces a new way to measure the existence of self-protective cognitive strategies: defensive verbal behavior assessment (DVBA). In Study 1, the authors introduce the coding procedure for DVBA and demonstrate evidence for its interrater reliability and construct validity. In Study 2, the authors demonstrate that defensive verbal behavior is influenced both by characteristics of the person and by the situational context. Together, the two studies illustrate that (a) reliable and valid behavioral assessment of defensive processes is possible in nonclinical samples without the need for lengthy assessment times or specialized clinical knowledge and (b) qualities of the person and the situation must be considered to provide a full account of self-protective behavior.

Many psychological theories assume that individuals continuously organize and interpret self-relevant information to construct and maintain a cohesive self-concept. Mental representations of the self in conscious awareness constitute an important part of the self-concept (Campbell, 1990; Greenwald & Pratkanis, 1984; Kihlstrom & Cantor, 1984; Markus & Sentis, 1983; Markus & Wurf, 1987). Emotions, thoughts, behaviors, or information from the environment that are discrepant with an individual's consciously held view of the self can be perceived as a threat, resulting in a negative self-evaluation, lowered self-esteem, and increased negative affect, be it anxiety, depression, or general displeasure (e.g., Brenner, 1982; Doherty, Weigold, & Schlenker, 1990; S. Freud, 1894/1962, p. 47; Gedo, 1980; Higgins, 1987; Horowitz, Markman, Stinson, Fridhandler, & Ghannam, 1990; Klein, 1976; Lockhard & Paulhus, 1988; Moreland & Sweeny, 1984; Perry, 1993; Rogers, 1957; Sackeim & Gur, 1979, 1985; Shapiro, 1989; Swann, Stein-Seroussi, & McNulty, 1992; Vaillant, 1992; Weinberger, 1990; Wylie, 1979). The result is that indi-

viduals often limit their awareness to include only those experiences congruent with, or distort representations of their experience to sustain, their conscious concept of self (S. Freud, 1914/1957; Kernberg, 1967; Rogers, 1951). All people, from time to time, reinterpret information to be more consistent with their self-image, thereby reducing conscious perceptions of threat in the environment and the associated emotional consequences. It should be possible to detect variations in awareness and distortions in people's representations of their conscious experience. Doing so would provide concrete evidence that self-protective processes are afoot. Such is the purpose of this report.

Defense Mechanisms: Cognitive Tools for Self-Protection

In the present report, we relied on the psychodynamic literature to provide a theoretical context for how individuals employ cognitive mechanisms to protect the self (Cooper, 1989; S. Freud, 1914/1957; Horowitz et al., 1990; Sackeim & Gur, 1979; Shapiro, 1989; Vaillant, 1994; Weinberger & Schwartz, 1990; for recent reviews, see Cramer, 2000; Paulhus, Fridhandler, & Hayes, 1997). *Defense mechanism* was Freud's abstract

term for a cognitive strategy employed to transform threatening information into a less threatening form in conscious thought. Defense mechanisms can be thought of as preferred modes of perceiving and thinking (Singer & Sincoff, 1990; Slavin & Kriegman, 1992) that individuals use when they are motivated to maintain self-esteem and regulate affect (A. Freud, 1937, p. 235). They are not special cognitive mechanisms; any of the cognitive functions we use to perceive and interpret our experiences can be employed in the service of self-protection and affect regulation. Conversely, the same cognitive functions that serve the individual's self-protective motivations can serve other purposes (Brenner, 1981; Hartmann, Kris, & Loewenstein, 1964; Stolorow & Lachmann, 1980).

Because defensive functioning occurs outside of awareness, and if successful produces a positive, integrated sense of self, it can be hard to detect and, therefore, difficult to study empirically. Researchers have been stymied for almost a century in their attempts to measure defensive processes empirically (cf. Vaillant, 1992). The purpose of the present article is to introduce a method for detecting traces left by defensive or self-protective processes in the content and structure of speech. Using structured coding procedures, we look for evidence that defensive processing has occurred by examining its output: representations of conscious experience. Our goal was to capture stable individual differences, as well as more momentary fluctuations, in defensive processes.

We begin with the general principle that, all things being equal, the magnitude of defensive functioning will be proportional to the magnitude of perceived threat. Two elements should predict the magnitude of perceived threat: characteristics of the situation and of the individual. Obviously, the degree of threat varies with the situation (Grzegolowska-Klarkowska & Zolnierczvk, 1988). In general, when threatening information is made salient, individuals are more likely to "defend" the self by decreasing their awareness of the threatening stimulus and distancing themselves from their emotional reactions to it to some degree.

All other things being equal, there are also substantial individual differences in the degree to which people perceive the need to protect themselves (Weinberger, 1990). Some individuals perceive large degrees of threat in their environment because they believe themselves vulnerable or that others intend harm (Feldman Barrett & Swim, 1998; Quigley & Feldman Barrett, 1999; Shapiro, 1986; Vaillant, 1992). Some have rigidly defined self-concepts that are easily threatened because a large proportion of their experiences conflict with their consciously held views of themselves (Weinberger, 1990). In either case, such individuals consistently perceive threat in their environments independently of the immediate situation (Horowitz et al., 1990; Weinberger, 1990). We would call such individuals "offensive" rather than "defensive" because they continually perceive threat over and above the real situational contingencies. Their representations of their conscious experience should be quantifiably different from those individuals who are more trusting or who possess more flexible self-concepts such that they can accept thoughts and feelings that challenge their consciously held view of self.

OVERVIEW OF STUDIES

The main objective in the present investigation is to evaluate whether defensive functioning could be (a) meaningfully assessed in representations of conscious experience and (b) understood as an interaction between characteristics of the individual and of the situational context. Study 1 presents a new method for documenting the traces of self-protective processing, called defensive verbal behavior assessment (or DVBA). Our goal was to develop a behavioral assessment strategy that was reliable and valid when used on brief interview data from nonclinical samples and with undergraduate coders. Study 2 uses DVBA to test the hypothesis that an individual's baseline level of defensive functioning (i.e., degree of self-protective behavior that existed in the absence of objective threat and resulted from person characteristics) interacted with increasing amounts of threat contained in the situation to produce measurable changes in defensive verbal behavior.

STUDY 1: INTRODUCING DEFENSIVE
VERBAL BEHAVIOR ASSESSMENT

The theoretical foundation for DVBA was based on a dimensional model of defensive processes. The DVBA coding procedure is different from other theoretically derived, dimensional assessment strategies (e.g., Cooper, Perry, & Arnow, 1988; Vaillant, 1986) in that coders rate the level of defensive verbal behavior according to certain general criteria rather than rating the presence or absence of specific defense mechanisms. Although individuals may use various cognitive strategies (i.e., specific defense mechanisms) to maintain their self-esteem and regulate their affect regulation at any given moment in time, it is the shared consequences of using these strategies (rather than the specific transformations described by specific mechanisms) that are of interest here. Whatever the merits of examining the specific cognitive transformations that occur in someone's mind, we reasoned that it is the outward manifestation shared by the transformations that people will see and react to and that is therefore most important to understand.

DVBA is conceptually rooted in two dimensions. The first dimension, the level of awareness of the precipitating threat, is defined as the extent to which individuals consciously represent conflicts between aspects in their subjective experience or between their subjective experience and their behavior. This dimension is consistent with the view that defense mechanisms differ in the degree to which they prevent threatening contents from entering self-reflective awareness (Horowitz, 1988b; Vaillant, 1990). The second dimension, the degree of distortion created by the use of cognitive strategies, is defined as the modification of feelings, behaviors, or thoughts in representations of conscious experience to preserve a consistent self-concept. All psychodynamic theories and many social psychological theories assume that people hold conscious representations of themselves that diverge from the objective conditions of the situation. The degree of distortion involved with transforming this threatening information is an important property of defense mechanisms because some mechanisms distort conscious experience more than do others (e.g., Haan, 1977; Vaillant, 1977). These two dimensions, awareness and distortion, work together to describe the important aspects of defensive functioning. Because the two are strongly inversely related to one another, we opted for a global judgment strategy.

Our dimensional framework was translated into a coding procedure using clinical theories that detail how to interpret the presence of psychological process from verbal behavior (Shapiro, 1989). According to Shapiro (1989), markers of defense exist not only in the content of verbal responses but also in the coherence and quality of those responses. How something is said can provide insights into the person's motivations. People speak to communicate with others, sort out feelings, and/or bolster or confirm a self-belief. Defense mechanisms are active when the speaker uses speech primarily to influence himself or herself or to confirm a positive self-belief, instead of as a means of communicating with others. In such cases, speech represents a method of modifying, dissipating, or preventing the articulation of thoughts and feelings that will threaten the self, and it is designed to help the speaker try to think or feel something different from what he or she originally thought or felt.

Study 1 was designed to assess the reliability and validity of the DVBA procedure. We assessed the presence of defensive verbal behavior in a standardized, stressful interview where individuals were asked about value-laden experiences that would likely conflict with their consciously held self-concept. Prior to the interview, participants completed several traditional self-report measures associated with self-deception and impression management. These measures were administered to assess the construct validity of the new coding procedure. Participants then took part in an audiotaped interview. Following the interview, participants rated how threatening they found the interview and completed a vocabulary test. The latter measures also were administered to address additional construct validity issues. Trained, undergraduate coders then rated each interview from audiotape.

Reliability. We predicted that DVBA would produce acceptable interrater reliability with a nonclinical sample because it did not attempt to discriminate among specific defense mechanisms.

Construct validity. We expected DVBA would be positively related to preinterview self-report ratings of self-deception but not impression management. Previous research has shown that conscious attempts at positive self-presentation (i.e., impression management) are independent of the individual's attempts to protect the self-concept from perceived threat (Hogan, 1983; Paulhus, 1984, 1988). We predicted that DVBA would be inversely related to conscious appraisals of threat after the interview. Conscious representations of threat are not necessarily congruent with threat appraisals that are made automatically outside of awareness. We hypothesized that appraisals of threat made outside of awareness produce a motivation to protect the self, whereas conscious appraisals of threat derive from a motivation to protect the self. This distinction is supported by previous research (Croyle, 1992). We also predicted that DVBA would be unrelated to vocabulary. We thought it possible that individuals who were less verbally proficient might be seen as more defensive because they might not have the words to explain themselves, and we sought to rule out this possibility.

Incremental validity. We predicted that DVBA would show adequate incremental validity when compared to self-report measures of self-deception. Although we expected a positive relationship between defensive verbal behavior and self-reported self-deception, we expected the relationship to be moderately strong rather than completely overlapping because (a) self-deception is a concept that includes other psychological phenomena in addition to defensive behavior; (b) the self-deception subscale of the Balanced Inventory of Desirable Responding (SDE) measures stable individual differences in self-protective processing, whereas the DVBA is designed to measure both stable individual differences and momentary fluctuations in self-protective processes; and (c) explicit self-reports designed to measure nonconscious psychological processes will not completely overlap with more implicit attempts to measure

those processes (Greenwald & Banaji, 1995; Shedler, Mayman, & Manis, 1993).

Method

PARTICIPANTS

Participants were 51 undergraduate students (14 men, 37 women) who ranged in age from 17 to 52 years ($M = 23$ years). Participants were enrolled in psychology courses at The Pennsylvania State University and received extra credit points toward their final grade in exchange for their participation.

MATERIALS

Standardized stressful interview. We designed a stressful interview consisting of 25 questions. The first 5 questions were designed to assess the participant's baseline style of responding in an interview setting and were neutral in content (e.g., "How do you feel about coming in for the interview today?" or "Tell me a little about your family"). The remaining 20 questions were designed to be threatening by asking participants to discuss value-laden experiences that were likely to conflict with their consciously held self-concepts. The questions spanned four domains: violations of morality (e.g., "Describe a time when you've broken your own moral code"), aggression (e.g., "Describe a time when you have deliberately said something to hurt someone's feelings"), sexuality (e.g., "How satisfied are you with your sex life?"), and negative self-image (e.g., "Tell me about a time when you felt that your parents were really disappointed in you"). The interviews were conducted by one graduate student in clinical psychology or by one of three undergraduate research assistants (including the second author). The graduate student and a Ph.D. clinical psychologist (the first author) trained the undergraduate research assistants to make standardized, neutral queries that were designed to elicit information from a participant without restricting or encouraging the content of the participant's response (e.g., "Can you tell me more about that?"). During training sessions, the research assistants conducted practice interviews that were evaluated by both the graduate student and the psychologist. In addition, all interviewers ran three pilot participants before beginning the actual study and were given feedback on their interviewing techniques.

DVBA, Version 2.0 (Feldman Barrett, Williams, Boekman, & Feeney, 1997). Defense mechanisms allow individuals to reduce their perception of threat in the environment by altering how events are represented in conscious thought. Sometimes people merely distort the meaning of threatening information and other times they may avoid awareness of the information altogether. When people perceive a threat to their self-esteem, they attempt to manage surfacing negative affect by controlling whether the conflict or threat enters consciousness (i.e., awareness) as well as the content of the thoughts and feelings that enter consciousness (i.e., distortion). The result is that, to some extent, they distance themselves from their emotional experience and they avoid thoughts and feelings that are in conflict with their consciously held self-image. The DVBA procedure was designed to assess the extent of awareness and distortion in participants' response to each interview question. DVBA was coded using a 4-point scale (0 to 3). The scale allowed coders to conceptualize a participant's defensive verbal behavior in global terms, yet it preserved the underlying distinctions along our two dimensions. A manual detailing the coding procedure and containing practice examples can be obtained from the first author.

Responses were coded "0" if defensive verbal behavior was absent. With such responses, the participant consciously represents information perceived as threatening to the self and was able to express the associated affects (e.g., regret, embarrassment, vulnerability). There was little or no distortion of potentially threatening information by denying, avoiding, or transforming it. Self-descriptive statements displayed an integration of positive and negative information, and responses were personalized (i.e., the person addressed the question in a personalized manner and takes responsibility for feelings, cognitions, and behaviors). A response also was coded a "0" in cases where participants genuinely did not feel threatened by the interview questions (i.e., gave no verbal indication of feeling threatened by questions; responded openly and directly to interview queries).

Responses were coded "1" if mild defensive behavior was present. Such responses displayed moderate awareness of self-threatening information (i.e., the participant consciously represented some elements of threatening information but countered or downplayed them). The affects generated by this self-threatening information were expressed to a moderate degree. Some distortion of subjective experience was present (i.e., the participant described both positive and negative aspects of experience but minimized or justified the self-threatening aspects). Self-descriptive statements included some negative information, but positive and negative aspects were not integrated. Responses were mildly distanced from the self (e.g., the question was answered in a personal fashion, although there was some reference to social norms or justification).

Responses were coded "2" if moderate defensive behavior was present. Such responses displayed limited awareness of self-threatening information (i.e., there was little conscious representation of negative information). Affects generated by self-threatening information were masked or suppressed to a substantial degree.

There was moderate distortion of subjective experience, such that self-descriptive statements contained little threatening information (because it was actively denied, avoided, or transformed to a substantial degree). In some cases, participants incorrectly attributed negative self-perceptions to others. There was substantial distance from the self (i.e., responses were not personalized, implied no personal responsibility for behavior, and relied heavily on social norms).

Responses were coded "3" if highly defensive behavior was present. Such responses were characterized by little or no awareness of self-threatening information (i.e., no explicit reference to threatening information or their associated affects) combined with evidence of high distortion in subjective experience. For example, only extreme positive aspects of experience are reported, but negative aspects are indirectly acknowledged by attributing their consequence to other people or external circumstances. These responses were characterized by the greatest distance from self.

Responses to each interview question were coded independently of one another. Responses were coded conservatively, meaning that we assigned the lowest possible level of defense to a participant's response based on the verbal material he or she provided for each interview question. In addition, 7.8% of the responses were coded as "non-scorable," meaning that (a) the content of the responses was too ambiguous to be coded using the scoring criteria or (b) the interviewer had queried incorrectly by leading the participant. We included this category to minimize the use of inference by coders, and non-scorable responses were not used in calculating DVBA scores. They were included in estimates of interrater reliability, however.

The interviews were coded by one graduate student in clinical psychology and two of the three undergraduate research assistants who conducted the interviews (one of whom is the second author). Each audiotaped interview was rated independently by two coders. All coders were trained in the use of the coding procedure prior to the study. (Coders rated a series of pilot tapes independently and then discussed their ratings as a group to achieve consensus.)

One defensive verbal behavior score was obtained for each participant by taking the mean of both coders' ratings on 20 psychologically threatening questions (called the DVBA-T score). Another score also was calculated for neutral questions by taking the mean of both coders' ratings for the first five interview questions (called the DVBA-N score). The mean DVBA-T score ranged from .05 to 1.30, with a mean of .51 and a standard deviation of .25. The mean DVBA-N score ranged from .00 to .75, with a mean of .16 and a standard deviation of .17.

Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1988). The BIDR is a commonly used self-report measure of self-deception (20 items) and impression management (20 items). Self-deception (SDE) more closely resembles the psychodynamic concept of defense because it is conceptualized as the unconscious process of deceiving the self to protect against threat; it assesses the stable characteristic of implicitly maintaining a self-protective stance rather than assessing changes or fluctuations in self-protection in response to stimuli in the environment. Impression management is conceptualized as the conscious deception of others (Paulhus, 1988; Sackeim & Gur, 1979). Participants rated their agreement with the 40 BIDR items on a 7-point Likert scale (1 = *not true of me*, 7 = *very true of me*). The scoring key was balanced, and only extreme responses were scored.

Wechsler Adult Intelligence Scale Revised (WAIS-R) Vocabulary subtest (Wechsler, 1981). The WAIS-R vocabulary subtest is a widely validated measure of vocabulary and verbal intelligence. It was administered to participants according to the standardized procedure detailed for the WAIS-R. Participants were asked to explain the meaning of 35 words that become increasingly difficult, and their responses were coded on a 3-point scale (0 = *an incorrect response*, 1 = *a partially correct response*, 2 = *complete response*). Scores were summed to reflect a participant's verbal proficiency.

PROCEDURE

Students from undergraduate psychology courses were asked to participate in a study designed to evaluate the use of interviewing techniques. Upon arriving at the lab, each participant was told that he or she would be interviewed for the purpose of examining the interviewer's ability to use a variety of interviewing techniques. After signing the consent form, the participant completed the BIDR. Next, participants completed the standardized stressful interview. Following the interview, the participant completed the WAIS-R vocabulary subtest. Finally, each participant rated how threatening he or she found the interview experience on a 10-point Likert scale (0 = *not at all threatening*, 10 = *very threatening*) and was debriefed.

Results

INTERRATER RELIABILITY

The intraclass correlation coefficient (Shrout & Fleiss, 1979) for the defensive behavior threat composite (DVBA-T) was adequate, $I_r = .67$, given that an interrater reliability of .60 is considered acceptable for interview-based scoring procedures (Suen, 1988). $I_r = .41$ was low for the neutral composite (DVBA-N).

CONSTRUCT VALIDITY

Table 1 presents the Pearson product correlations between DVBA scores and the other psychological measures. As predicted, individuals scoring high in self-deception (as measured by the SDE scale) evidenced more defensive verbal behavior during the neutral and the threatening portions of the interview than did those lower in self-deception. A hierarchical regression analysis indicated DVBA-T scores were uniquely related to the preinterview self-deception ratings, over and above the level of DVBA-N, $b = 3.79$ ($SE = 1.44$), $B = .33$ ($SE = .12$), $t = 2.61$, $p < .01$. Thus, individuals who were high in self-deception reacted to the threatening interview questions with a larger increase in defensive behavior over baseline than did those who were less self-deceptive. As predicted, defensive verbal behavior was negatively related to conscious appraisals of threat. Individuals who demonstrated lower levels of defensive behavior were more likely to report that they found the interview threatening when compared to those who demonstrated higher levels of defensive verbal behavior. As predicted, impression management (IM) (as measured by the impression management subscale of the BIDR) and vocabulary were not related to defensive verbal behavior.

INCREMENTAL VALIDITY

As predicted, DVBA appeared to have incremental validity when compared to SDE. The correlation between SDE and conscious threat ratings was negative and marginally significant, $r = -.25$, $p < .09$, compared to the stronger negative relationship observed between DVBA-T and threat ratings, $r = -.46$, $p < .01$. Moreover, the relationship between DVBA-T and conscious appraisals of threat remained significant when controlling for individuals' SDE scores, $b = -4.30$ ($SE = 1.49$), $B = -.44$ ($SE = .15$), $t = 2.89$, $p < .01$. The opposite was not true. Thus, it appears that the self-report and behavioral measure both tapped defensive functioning but the behavioral measure seemed to capture variance in criteria related to defensive processes that the self-report measure did not.

Discussion

Study 1 provided evidence that a relatively simple coding procedure for defensive verbal behavior assessment was reliable and valid. The reliability coefficients for the composite defensive verbal behavior scores in the context of threatening questions were acceptable according to established standards for interview-based coding procedures. Although the reliability of defensive verbal behavior scores to neutral questions was weaker, both DVBA-T and DVBA-N were correlated to external measures, indicating that the coding procedure was a valid means of documenting the existence of defensive pro-

TABLE 1: Correlations Between Defensive Verbal Behavior Assessment and Other Psychological Measures: Study 1

Criterion Measures		*Defensive Verbal Behavior*	
Name	M *(SD)*	*DVBA-N*	*DVBA-T*
SDE	5.31 (2.94)	.44***	.40**
IM	5.45 (3.05)	.04	.27†
Perceived threat	4.30 (2.59)	−.31*	−.46***
Vocabulary	51.27 (8.87)	.03	.18

NOTE: $N = 51$. Standard deviations are in parentheses. SDE = the self-deception subscale of the Balanced Inventory of Desirable Responding, IM = the impression management subscale of the Balanced Inventory of Desirable Responding, Perceived Threat = ratings of situational threat, Vocabulary = the vocabulary subtest of the Wechsler Adult Intelligence Scale-Revised, DVB-N = defensive verbal behavior during neutral questions, DVBA-T = defensive verbal behavior during threatening questions, and DVBA-N = defensive verbal behavior during neutral questions.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

cessing. Individuals with higher DVBA scores scored higher on a measure of self-deception, indicating that those who consistently perceive threat independently of the immediate situation gave more evidence of self-protection in their verbal responses to both threatening and neutral questions. Individuals with higher DVBA scores also denied that the interview was threatening. This is consistent with previous research indicating that individuals who are faced with an objectively threatening situation reported that they consciously experienced less threat than those who were not objectively threatened (Croyle, 1990; Croyle & Sande, 1988). DVBA scores were not related to verbal proficiency or to impression management. Taken together, Study 1 provides preliminary evidence consistent with the hypothesis that defensive verbal behavior is an external manifestation of internal, cognitive processes that operate to serve a self-protective function. When individuals psychologically manage conflictual experiences that can threaten their self-concept and produce negative affect, they leave evidence of having done so in the structure of verbal responses.

STUDY 2: A PERSON-SITUATION ANALYSIS
OF DEFENSIVE VERBAL BEHAVIOR

Study 2 was designed to replicate and extend the findings from Study 1 by strengthening both the conceptual and methodological foundations associated with assessing defensive verbal behavior. We conceptualized defensive verbal behavior more formally as an interaction between characteristics of the person and the situation. Evidence for both influences can be found in Study 1. On average, participants were flexible in the amount of defensive verbal behavior that they produced in response to changing situational demands. Participants

appeared to display more defensive verbal behavior in responses to questions that challenged their self-concept (i.e., when interviewed about value-laden topics) than to questions that were more neutral in content. Moreover, some participants displayed more defensive verbal behavior than did others, even when no objective threat was present (during the neutral interview questions). These participants, who described themselves as self-deceptive, were also more reactive to threat as demonstrated by a larger increase in their already higher baseline levels of defense when the threatening interview questions commenced. All other things being equal, then, individuals will differ from one another in their baseline levels of defensive verbal behavior. In addition, all individuals will evidence some increase in defensive verbal behavior when pressed by situational cues, but some (with higher baseline levels of defensive behavior) will be more sensitive to these cues and therefore more guarded in situations that highlight conflictual experiences.

As in Study 1, Study 2 participants began by completing several questionnaire measures. We then assessed their defensive verbal behavior in response to a series of questions that were selected to be psychologically neutral or threatening. The interview in Study 2 was structured using an A-B-A design, such that participants were asked 5 neutral questions, followed by 15 psychologically threatening questions, followed by another 5 neutral questions. Using the self-deception subscale of the BIDR, we established two groups of participants: those who were highly self-deceptive and those who were not. Because self-deception is a characteristic that implies the habitual use of defense mechanisms (Horowitz, 1988b; Paulhus, 1984, 1988; Sackeim, 1983; Weinberger, 1990), we predicted that, as in Study 1, high-self-deceptive individuals would display higher baseline levels of defensive verbal behaviors (during the neutral phases of the interview) than those who were low in self-deception.

During the threatening phase of the interview, participants from each group were randomly assigned to one of two interviewing conditions. In the nonreflective condition, which resembled the interviewing strategy used in Study 1, interviewers used only neutral queries and responses. In the reflection condition, interviewers paraphrased the main idea and feeling in participants' statements to highlight contradictory aspects of participants' experience. These interviewing techniques, called reflection and confrontation (Luborsky, 1984; Shapiro, 1989; Strupp & Binder, 1984), are designed to highlight (a) the emotional information in a verbal report, (b) inconsistencies between verbal reports and nonverbal behavior, or (c) inconsistencies between multiple verbal reports. Both reflections and confrontations are thought to bring information that is outside of conscious

awareness into focus, thereby focusing individuals on aspects of their experience that they may be avoiding or distorting. Individuals who are high in self-deception should be particularly likely to perceive these responses as threatening, engage in defensive processing, and increase their defensive verbal behavior as a result. Thus, we predicted an interaction between self-deception and interview condition would produce increases in defensive verbal behavior during the threatening part of the interview. After the interview, participants completed a measure of vocabulary. Consistent with Study 1, we predicted that defensive behavior scores would be unrelated to the measure of vocabulary.

Method

PARTICIPANTS

Participants were 65 undergraduate students (18 men, 47 women) who ranged in age from 18 to 47 years (M = 20.1 years). Participants were enrolled in psychology courses at The Pennsylvania State University and received extra credit points toward their final grade in exchange for their participation.

MATERIALS

Structured stressful interview. The interview consisted of 25 questions (these are available from the first author on request). The first 5 questions were neutral and designed to assess participants' baseline styles of responding. The next 15 questions were psychologically threatening, designed to elicit information that might conflict with consciously held beliefs about the self, thereby producing an increase in defensive verbal behavior. Once again, the threatening questions tapped four domain areas (violations of morality, aggressive thoughts or behaviors, sexuality, and negative self-image). The remaining 5 questions were neutral and designed to assess the participant's ability to reappraise the degree of threat in the interview situation.

We attempted to improve the interrater reliability of DVBA from Study 1 by modifying the interview questions to maximize their threat value. To choose interview items, 40 potential interview items were rated by an independent sample of 62 undergraduates. These respondents were asked to rate how uncomfortable they would feel, and how uncomfortable other people would feel, answering each question in a face-to-face interview, via a 5-point Likert-type scale (1 = *not at all uncomfortable*, 5 = *very uncomfortable*). The 10 items with the lowest mean ratings were considered neutral, whereas the 15 items with the highest mean ratings were considered psychologically threatening. Nine questions were replaced from those used in Study 1; several threatening items were replaced and 4 neutral items were added.

DVBA, Version 2.0 (Feldman Barrett et al., 1997). The coding procedure used was identical to that described for Study 1. Interviews were coded by four of the seven undergraduate research assistants who conducted the interviews (one of whom is the second author). All coders were explicitly evaluated in the use of the coding procedure during five training sessions prior to rating the interview responses. Training consisted of rating practice tapes as a group and then independently rating a series of three additional practice tapes. Ratings on the three final tapes were compared to criteria ratings for each tape (set by the authors) and then discussed as a group to achieve consensus. Interrater reliability was assessed following the coder's ratings of the first and third practice tape. All coders were trained to .80 reliability on the practice tapes before rating tapes from the study.

Each interview was rated independently by two coders who were blind to participants' self-deception status. As in Study 1, coders assigned ratings to each interview item response independently using the 4-point scale. One composite defensive behavior score was obtained for each participant by taking the mean of both coders' ratings on the 15 psychologically threatening questions (DVBA-T). In addition, defensiveness scores were obtained for each of the two neutral sections of the interview by taking the mean of coders ratings on the first five neutral questions (DVBA-N1) and the last five neutral questions (DVBA-N2). The mean DVBA-T score ranged from .19 to 1.71, with a mean of .76 and a standard deviation of .32. The mean DVBA-N1 score ranged from .00 to 1.50, with a mean of .36 and a standard deviation of .30. The mean DVBA-N1 score ranged from .00 to 1.10, with a mean of .36 and a standard deviation of .26.

Additional measures. As in Study 1, participants completed the BIDR (Paulhus, 1988) and the WAIS-R Vocabulary subtest (Wechsler, 1981).

PROCEDURE

The pretest sample of participants consisted of 199 undergraduate psychology students (52 men, 145 women). Participants were selected from a mass screening in which they completed the BIDR. Based on a tertiary split of the Self-Deception subscale scores ($M = 6.1$, $SD = 3.4$), participants were assigned to the low-self-deceptive group if they scored less than or equal to 4 ($M = 2.8$) or to the high-self-deceptive group if they scored equal to or greater than 8 ($M = 10.1$). The two groups differed significantly in their level of self-deception, $t(64) = 16.6$, $p < .001$, but also in their level of impression management ($M = 4.5$ vs. $M = 6.5$), $t(64) = 2.9$, $p < .01$.

Selected participants were then contacted by telephone and asked to participate in a study of interviewing techniques for partial course credit. As in Study 1, participants were told that they would be interviewed for the purpose of examining the interviewer's ability to use a variety of interviewing techniques. Individuals who agreed to participate were randomly assigned to either the experimental or control condition.

Upon arriving at the lab, each participant was reminded of the cover story. The participant read and signed a consent form and then completed the battery of self-report measures so that we could assess the stability of the groups and obtain pretest measures of the BIDR. Next, the participant took part in the standardized interview. All interviews were audiotaped. Following the interview, the participant completed the WAIS-R vocabulary subtest. Participants were then debriefed.

Standardized interviews were conducted by seven undergraduate research assistants (three women and four men). In the no-reflection condition, interviewer queries were standardized and neutral. The main goal in this condition was to simply elicit information from the participant without influencing the content of the participant's responses. In contrast, the techniques employed in the reflection condition emphasized reflective responding. The interviewer was instructed to identify, clarify, and convey discrepant verbal and nonverbal behavior as well as conflictual verbal statements ("It sounds like you have two feelings. When you first responded, you said that you doubted your ability to succeed in school but also said that you think you have a bright academic future"). To standardize this condition, interviewers were instructed to reflect only what they observed and to refrain from making any inferences. In addition, interviewers were reminded to reflect both the positive and the negative verbal content that they observed.

Results

INTERRATER RELIABILITY

The I_r for the DVBG-T was adequate (.84), as were the I_r for DVBA-N1 and DVBA-N2 scores (.76 and .70, respectively). Of interest, the intraclass correlation coefficients for single items were related to the threat ratings for each question obtained from the independent sample of undergraduates judges ($r = .46$, $p < .01$, for discomfort to self, and $r = .41$, $p < .01$, for discomfort to others). Thus, questions that pressed participants to discuss conflictual experiences attained higher interrater reliabilities than those that elicited more neutral content.

STABILITY AND VALIDITY OF
THE SELF-DECEPTION GROUPS

All participants assigned to the low-self-deception group on the basis of their pretest scores ($N = 31$) remained in that group based on their preinterview self-

deception ratings, and all but one participant assigned to the high-self-deception group ($N = 34$) remained in that group; one participant had a pretest score of 9 and a preinterview score of 4 and was reassigned to the low-self-deception group. As a result, there were 32 participants (24 women, 8 men) in the low-self-deception group and 33 participants (23 women, 10 men) in the high-self-deception group. The frequency of male and female participants was not significantly different in the two groups, $\chi^2(1) = .23$, *ns*. The groups differed in their preinterview SDE scores ($M = 4.65$, $SD = 3.17$ vs. $M = 8.21$, $SD = 3.63$), $F(1, 61) = 17.66$, $p < .001$, but not in their preinterview IM scores ($M = 4.69$, $SD = 2.54$ vs. $M = 5.88$, $SD = 3.55$), $F(1, 61) = 2.40$, $p < .13$.

Next, a 2×2 (Level of Self-Deception × Interview Condition) ANOVA was conducted on the defensive verbal behavior assessment score for the first set of neutral questions (DVBA-N1 composite) to determine baseline differences. The results are presented in the left third portion of Figure 1. As predicted, high-self-deceptive individuals displayed higher baseline levels of defensive verbal behaviors ($M = .46$, $SD = .35$) than those low in self-deception ($M = .25$, $SD = .21$), $F(1, 61) = 7.52$, $p < .01$. Unexpectedly, individuals in the no-reflection condition had lower baseline levels of defensive verbal behavior from those in the reflection condition ($M = .28$, $SD = .23$ vs. $M = .43$, $SD = .34$), $F(1, 61) = 4.02$, $p < .05$. The finding suggests either that random assignment to groups was not successfully achieved or that interviewers were using reflections and confrontations during the neutral questions. A MANOVA on the preinterview BIDR scores with interview condition as the grouping variable was not statistically significant (neither were either of the univariate tests), suggesting that random assignment to conditions was successful and that the later explanation is more likely. The interaction between self-deception and condition was not significant, $F(1, 61) = 2.32$, $p < .13$.

DIFFERENTIAL RESPONSIVENESS TO INTERVIEW QUESTIONS

Threatening questions. A 2×2 (Level of Self-Deception × Interview Condition) ANOVA was conducted on the composite of 15 threatening questions (DVBA-T). The middle portion of Figure 1 presents the level of defensive verbal behavior for each condition. As predicted, there was a main effect for self-deception: high-self-deceptive individuals displayed more defensive verbal behavior than low-self-deceptive individuals ($M = .89$, $SD = .35$ vs. $M = .63$, $SD = .23$), $F(1, 61) = 12.83$, $p < .001$. There was also a main effect for interview condition, with individuals in the reflective interview condition displaying more defensive verbal behavior than those in the no-reflection condition ($M = .88$, $SD = .33$ vs. $M = .61$, $SD = .24$), $F(1, 61) = 16.10$, $p < .001$. Both main effects were qualified by a
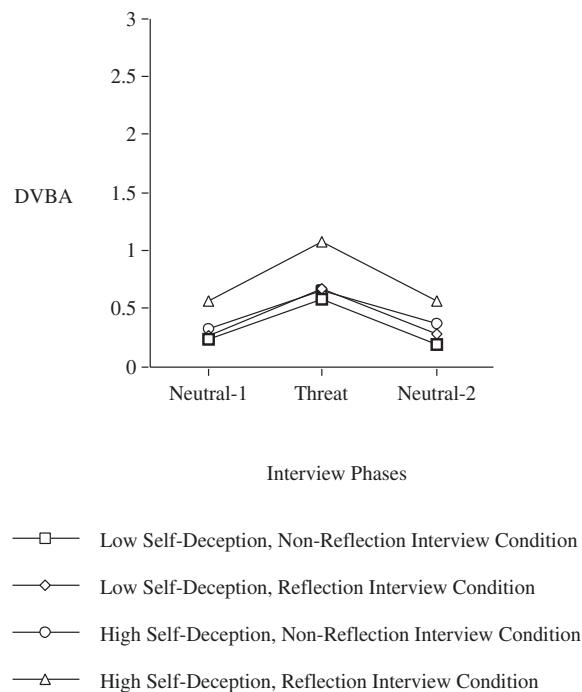

Figure 1   Group differences in defensive verbal behavior across three phases of interview.
NOTE: DVBA = defensive verbal behavior assessment.

significant interaction, $F(1, 61) = 6.40$, $p < .01$. As predicted, high-self-deceptive individuals in the reflective interview condition displayed more defensive verbal behavior than did all other groups of participants. The results of analyses using the DVBA-G scores were identical to those just presented.

Next, a hierarchical regression analysis was conducted to test the hypothesis that high-self-deceptive participants in the reflective interview condition demonstrated the largest increase in their DVBA-T scores, over and above their level of defensive behavior during the first neutral portion of the interview. The first neutral composite (DVBA-N1) was centered and entered as the predictor in Step 1 of the regression. Participants' level of defensive verbal behavior during the neutral phase of the interview was significantly related to that in the threatening part of the interview, $b = .55$ ($SE = .11$), $B = 53$ ($SE = .11$), $R^2 = .28$, $F(1, 63) = 24.18$, $p < .001$. Participants' level of self-deception and their interview condition were entered as predictors in Step 2 of the regression; effect codes were used to indicate group membership and condition. The incremental change in R^2 from Step 1 to Step 2 was significant, R^2 change $= .18$, $F(2, 61) = 8.24$, $p < .001$. Self-deception and interview condition had an additive effect on the level of defensive verbal behavior expressed during the threatening phase of interview. As predicted, individuals who were high in self-deception demonstrated greater increases in their

level of defensive verbal behavior in response to threat, over and above their relatively higher level during the first neutral interview phase, than did those lower in self-deception, $b = .08$ ($SE = .03$), $B = .27$ ($SE = .10$), $t = 2.61$, $p < .01$. Similarly, individuals who were assigned to the reflection interview condition displayed more defensive verbal behavior than did those in the nonreflective condition, $b = .10$ ($SE = .03$), $B = .32$ ($SE = .10$), $t = 3.22$, $p < .01$.

These effects were moderated by a significant interaction between level of self-deception and interview condition, indicated by entering their cross-product term into the regression at Step 3, $b = .06$ ($SE = .03$), $B = .20$ ($SE = .10$), R^2 change $= .04$, $F(1, 60) = 4.28$, $p < .05$. As predicted, high-self-deceptive individuals in the reflection interview condition displayed a larger increase in defensive verbal behavior during the threatening part of the interview, even after controlling for baseline levels, as did low-self-deceptive individuals in the nonreflective condition. Thus, as predicted, reflections and confrontations seemed to heighten the defensive behavior of highly self-deceptive individuals but decreased the defensive behavior of individuals who were low in self-deception.

Second set of neutral questions. A 2×2 (Level of Self-Deception × Interview Condition) ANOVA was conducted on the second composite of neutral questions (DVBA-N2). The results are presented in the right portion of Figure 1. As predicted, high-self-deceptive individuals displayed more defensive verbal behavior ($M = .48$, $SD = .26$) than did those low in self-deception ($M = .24$, $SD = .19$), $F(1, 61) = 18.10$, $p < .001$. Furthermore, those in the reflection interview condition continued to display more defensive verbal behavior ($M = .43$, $SD = .27$) than did those in the nonreflective condition ($M = .28$, $SD = .20$), $F(1, 61) = 6.77$, $p < .01$. The interaction was not significant, $F(1, 61) = .89$, *ns.*

Next, a hierarchical regression analysis, similar to that reported above, was conducted for DVBA-N2. DVBA-T was centered and entered as the predictor in Step 1 of the regression to establish participants' response to the threatening interview items. Participants' level of defensive verbal behavior during the threatening phase of the interview was significantly related to that in the neutral conclusion of the interview, $b = .59$ ($SE = .07$), $B = .73$ ($SE = .09$), $R^2 = .54$, $F(1, 63) = 72.67$, $p < .001$. Participants' level of self-deception and their interview condition were entered as predictors in Step 2. The incremental change in R^2 from Step 1 to Step 2 was marginally significant, R^2 change $= .04$, $F(2, 61) = 2.82$, $p < .07$. As predicted, individuals high in self-deception continued to demonstrate greater amounts of defensive verbal behavior at the conclusion of the interview, even when controlling for their relatively higher level during the threatening interview phase, than did those lower in self-

deception, $b = .06$ ($SE = .02$), $B = .22$ ($SE = .09$), $t = 2.37$, $p < .02$. Interview condition did not continue to influence defensive verbal behavior, however, $b = .01$ ($SE = .02$), $B = .02$ ($SE = .09$), $t = .22$, *ns.* Furthermore, the interaction term, entered in Step 3 of the regression, was not significant, $b = -.02$ ($SE = .02$), $B = -.07$ ($SE = .09$), R^2 change $= .0$, $F(1, 60) = .57$, *ns.* Finally, paired t tests suggested that there was no difference between the DVBA-N1 and DVBA-N2 scores across groups of participants.

COMPLETING EXPLANATIONS

Neither sex differences nor vocabulary scores provided a viable alternative explanation for the pattern of observed results. Individuals high and low in self-deception did not differ on the vocabulary measure. High-self-deceptive participants in the nonreflective interview condition scored lower on the vocabulary test that did participants in all of the other conditions (including high-self-deceptive individuals in the reflection condition who scored highest in DVBA). In addition, all of the analyses were rerun with sex of participant, sex of interviewer, and their cross-product term as covariates. In all analyses, the results were practically identical to those already reported. Therefore, sex effects did not provide a viable alternative explanation for the differences in verbal behavior that we observed.

GENERAL DISCUSSION

For decades, social psychologists have focused a considerable amount of energy studying how people negotiate the competing motivations to know, but to protect, the self (e.g., Heider, 1958; Sedikides & Strube, 1995). In recent years, theorists have integrated psychodynamic theories with social psychology (e.g., Andersen, Reznik, & Manzella, 1996; Barron, Eagle, & Wolitzky, 1992; Erdelyi, 1985; Hinkley & Andersen, 1996; Horowitz, 1988a; Westen, 1990, 1991, 1992) to understand how individuals select, process, and remember information about themselves. According to current social psychology and psychodynamic theories, experiences that conflict with one's self-concept are often perceived as psychologically threatening. In an effort to protect the self, individuals employ cognitive strategies (in social psychology) or defense mechanisms (in psychodynamic theories) to avoid awareness of conflicting aspects of their experience. The present investigation introduced evidence that the products of these processes are observable and quantifiable.

In Study 1, we presented evidence that defensive verbal behavior increased from baseline when participants were asked to discuss value-laden topics in a face-to-face interview, and this change was measured reliably. Defensive verbal behavior in response to the threatening interview questions was associated with a variety of self-report

measures that helped to establish both the convergent, discriminant, and incremental validity of the coding procedure.

In Study 2, we demonstrated that defensive behavior, similar to any other human behavior, was a product of both the person and situation. Highly self-deceptive individuals produced more defensive verbal behavior at baseline than did those lower in self-deception. Those high in self-deception were more responsive to the threat than were those lower in self-deception, and this was magnified by reflections and confrontations. When the objective degree of threat in the interview decreased during the final set of neutral questions, all participants produced less evidence of defensive processing, although self-deceptive individuals continued to display more defensive verbal behavior than those lower in self-deception. Finally, vocabulary differences and sex effects did not provide a viable alternative explanation for the differences in verbal behavior that were observed.

IMPLICATIONS

Taken together, these studies have several implications. First, experimental evidence indicated that environmental conditions, in combination with person factors, can be manipulated to produce predictable increases or decreases in defensive verbal behavior and presumably the use of the motivated cognitive strategies from which they derive. Thus, defensive behavior itself does not necessarily tell us anything about chronic motivations of the individual—it might tell us something about the situation, or a person's response to the situation. Consider Figure 1: It is interesting to note three of the four conditions (low-self-deceptive control, low-self-deceptive reflective, and high-self-deceptive control) looked very similar in their average levels of defensive behavior. In particular, the means for the low-self-deceptive reflective and high-self-deceptive control groups are almost identical, suggesting that whether threat stems from a situational context alone or from a person characteristic alone, the magnitude of defensive behavior that results may be the same. Although defensive verbal behavior may not represent chronic motivations, it does tell us when self-protection and the need for affect regulation is occurring. In this way, the coded verbal behavior is unlikely to represent other characteristics, such as need for cognition (Cacioppo & Petty, 1982) or cognitive complexity (Singer, 1984).

Second, our findings have the potential to contribute a new perspective to the literature on self-protective processes. For example, take the debate over the adaptive nature of distorted beliefs about the self (e.g., Colvin & Block, 1994; Shedler et al., 1993; Taylor, 1989; Taylor & Brown, 1988, 1994a, 1994b). From our perspective, defensive processing is not synonymous with mental ill-

ness (as suggested by Taylor, 1989), and it is not identical to illusory mental health (as suggested by Shedler et al., 1993). Rather, defense mechanisms are a form of motivated information processing that may be adaptive but that can leave individuals vulnerable to maladaptive behavior under certain conditions. Typically, there may be little or no cost to the distortion and lack of awareness associated with motivated processing of self-relevant information. Using cognitive strategies to augment one's self-esteem and positive affect, even if it means distorting reality to do so, may not be problematic when there is little cost to doing so and may even be adaptive under those circumstances. It is also adaptive, however, to adjust cognitive and behavioral responses to the demands of the situation. As was evident in Study 2, individuals can calibrate to objective changes in the level of threat presented to them. Some individuals may be less able to adapt to the changing environmental conditions that they are faced with in their everyday lives, however, and such inflexibility constitutes a vulnerability. Future research might profit from a detailed analysis of the costs and benefits associated with motivated processing of self-relevant information in particular situational contexts so as to identify the environmental conditions that transform a vulnerability into a real problem.

Finally, in the most practical sense, we hope that we have demonstrated that it is possible to measure defensive functioning in a relatively simple way. Although other coding systems for defensive behaviors exist, ours is based on a relatively elemental dimensional theory that preserves the underlying conceptual distinctions implied by more complicated models of defense, but without the need to discriminate between specific defense mechanisms or clusters of defense mechanisms. In addition, the coding procedure is reliable and valid when used with brief interview data with nonclinical samples and undergraduate coders, thereby avoiding difficulties associated with assessment time, sampling issues, and specialized clinical knowledge. The idea that individuals use cognitive strategies to protect the self is not novel to most social or personality psychologists. What is novel is that methods informed by psychodynamic principles are effective in detecting evidence that motivated information processing about the self is occurring.

REFERENCES

Andersen, S. M., Reznik, I., & Manzella, L. M. (1996). Eliciting facial affect, motivation, and expectancies in transference: Significant other representations in social relations. *Journal of Personality and Social Psychology, 71*, 1108-1129.
Barron, J. W., Eagle, M. N., & Wolitzky, D. L. (1992). *Interface of psychology and psychoanalysis.* Washington, DC: American Psychological Association.
Brenner, C. (1981). Defense and defense mechanisms. *Psychoanalytic Quarterly, 50*, 557-569.

Brenner, C. (1982). *The mind in conflict.* New York: International Universities Press.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42,* 116-131.

Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology, 59*(3), 538-549.

Colvin, R. C., & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown Formulation. *Psychological Bulletin, 116,* 3-20.

Cooper, S. (1989). Recent contributions to the theory of defense mechanisms. *Journal of the American Psychoanalytic Association, 37,* 865-891.

Cooper, S. H., Perry, J. C., & Arnow, D. (1988). An empirical approach to the study of defense mechanisms, I: Reliability and preliminary validity of the Rorschach defense scales. *Journal of Personality Assessment, 52,* 187-203.

Cramer, P. (2000). Defense mechanisms in psychology today: Further processes for adaptation. *American Psychologist, 55,* 637-646.

Croyle, R. T. (1990). Biased appraisal of high blood pressure. *Preventive Medicine, 19,* 40-44.

Croyle, R. T. (1992). Appraisal of health threats: Cognition, motivation, and social comparison. *Cognitive Therapy and Research, 16*(2), 165-182.

Croyle, R. T., & Sande, G. N. (1988). Denial and confirmatory search: Paradoxical consequences of medical diagnosis. *Journal of Applied Social Psychology, 18,* 473-490.

Doherty, K., Weigold, M. F., & Schlenker, B. R. (1990). Self-serving interpretations of motives. *Personality and Social Psychology Bulletin, 16*(3), 485-495.

Erdelyi, M. H. (1985). *Psychoanalysis: Freud's cognitive psychology.* New York: Freeman.

Feldman Barrett, L., & Swim, J. (1998). Appraisals of prejudice: A signal detection framework. In J. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 11-36). Academic Press.

Feldman Barrett, L. A., Williams, N. L., Boekman, L., & Feeney, K. (1997). Manual for the Defensive Verbal Behavior Assessment, Version 2.0. Unpublished Coding Manual, Boston College.

Freud, A. (1937). *The writings of Anna Freud: Vol. 2. The ego and the mechanisms of defense.* New York: International Universities Press.

Freud, S. (1957). On the history of the psycho-analytic movement. In J. Strachey (Ed.), *The standard edition of the complete works of Sigmund Freud, Volume 14* (pp. 7-66). London: Hogarth. (Original work published 1914)

Freud, S. (1962). The neuro-psychoses of defense. In J. Strachey (Ed.), *The standard edition of the complete works of Sigmund Freud, Volume 3.* London: Hogarth. (Original work published 1894)

Gedo, J. E. (1980). Reflections on some current controversies in psychoanalysis. *Journal of the American Psychoanalytic Association, 28,* 363-383.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4-27.

Greenwald, A. G., & Pratkanis, A. R. (1984). The self. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 13, pp. 129-178). Hillsdale, NJ: Lawrence Erlbaum.

Grzegolowska-Klarkowska, H., & Zolnierczvk, D. (1988). Defense of self-esteem, defense of self-consistency: A new voice in an old controversy. *Journal of Social and Clinical Psychology, 6,* 171-179.

Haan, N. (1977). *Coping and defending: Processes of self-environment organization.* New York: Academic Press.

Hartmann, H., Kris, E., & Loewenstein, R. M. (1964). Papers on psychoanalytic psychology. *Psychological Issues,* Monograph 14.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: John Wiley.

Higgins, T. E. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*(3), 319-340.

Hinkley, K., & Andersen, S. M. (1996). The working self-concept in transference: Significant-other activation and self change. *Journal of Personality and Social Psychology, 71*(6), 1279-1295.

Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), *Nebraska symposium on motivation* (Vol. 29, pp. 55-89). Lincoln: University of Nebraska Press.

Horowitz, M. J. (1988a). *Introduction to psychodynamics.* New York: Basic Books.

Horowitz, M. J. (1988b). Psychodynamic phenomena and their explanation. In M. J. Horowitz (Ed.), *Psychodynamics and cognition* (pp. 3-20). Chicago: University of Chicago Press.

Horowitz, M. J., Markman, H. C., Stinson, C., Fridhandler, B. M., & Ghannam, J. H. (1990). A classification theory of defense. In J. L. Singer (Ed.), *Repression and dissociation: Implications for personality theory, psychopathology, and health* (pp. 61-84). Chicago: University of Chicago Press.

Kernberg, O. (1967). Borderline personality organization. *Journal of the American Psychological Association, 15,* 641-685.

Kihlstrom, J. F., & Cantor, N. (1984). Mental representations of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17). New York: Academic Press.

Klein, G. S. (1976). *Psychoanalytic theory: An exploration of essentials.* New York: International Universities Press.

Lockhard, J. S., & Paulhus, D. L. (Eds.). (1988). *Self-deception: An adaptive mechanism?* New York: Prentice Hall.

Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy: A manual for supportive-expressive treatment.* New York: Basic Books.

Markus, H., & Sentis, K. (1983). The self in social information processing. In H. Markus & J. Cross (Eds.), *Psychological perspectives on the self* (Vol. 1, pp. 41-70). Hillsdale, NJ: Lawrence Erlbaum.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38,* 299-337.

Moreland, R. L., & Sweeny, P. (1984). Self-expectancies and reactions to evaluations of personal performance. *Journal of Personality, 52*(2), 156-176.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598-609.

Paulhus, D. L. (1988). *Assessing self-deception and impression management in self-reports: The Balanced Inventory of Desirable Responding.* (Manual available from the author at the Department of Psychology, University of British Columbia, Vancouver, BC, Canada V6T 1Y7)

Paulhus, D. L., Fridhandler, B., & Hayes, S. (1997). Psychological defense: Contemporary theory and research. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality* (pp. 544-580). New York: Academic Press.

Perry, J. C. (1993). The study of defense mechanisms and their effects. In N. Miller, L. Luborsky, J. Barber, & J. Docherty (Eds.), *Psychodynamic treatment research: A handbook for clinical practice.* New York: Basic Books.

Quigley, K. S., & Feldman Barrett, L. (1999). Emotional learning and mechanisms of intentional psychological change. In J. Brandtstadter & R. M. Lerner (Eds.), *Action and development: Origins and functions of intentional self-development* (pp. 435-464). Thousand Oaks, CA: Sage.

Rogers, C. R. (1951). *Client-centered therapy, its current practice, implications, and theory.* Boston: Houghton-Mifflin.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21,* 95-103.

Sackeim, H. A. (1983). Self-deception, self-esteem, and depression: The adaptive value of lying to oneself. In J. Masling (Ed.), *Empirical studies of psychoanalytic theories.* Hillsdale, NJ: Lawrence Erlbaum.

Sackeim, H. A., & Gur, R. C. (1979). Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology, 47,* 213-215.

Sackeim, H. A., & Gur, R. C. (1985). Voice recognition and the ontological status of self-deception. *Journal of Personality and Social Psychology, 48*(5), 1365-1368.

Sedikides, C., & Strube, M. J. (1995). The multiply motivated self. *Personality and Social Psychology Bulletin, 21*(12), 1330-1335.

Shapiro, D. (1986). *Neurotic styles.* New York: Basic Books.

Shapiro, D. (1989). *Psychotherapy of neurotic character.* New York: Basic Books.

Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist, 48,* 1117-1131.

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

Singer, J. L. (1984). The private personality. *Personality and Social Psychology Bulletin, 10*, 7-30.

Singer, J. L., & Sincoff, J. B. (1990). Beyond repression and the defenses. In J. L. Singer (Ed.), *Repression and dissociation: Implications for personality theory, psychopathology, and health* (pp. 471-496). Chicago: University of Chicago Press.

Slavin, M. O., & Kriegman, D. (1992). Psychoanalysis as a Darwinian depth psychology: Evolutionary biology and the classical-relational dialectic in psychoanalytic theory. In J. W. Barron, M. N. Eagle, & D. L. Wolitzky (Eds.), *Interface of psychology and psychoanalysis* (pp. 37-76). Washington, DC: American Psychological Association.

Stolorow, R., & Lachmann, F. (1980). *The psychoanalysis of developmental arrest.* New York: International Universities Press.

Strupp, H. H., & Binder, J. L. (1984). *Psychotherapy in a new key.* New York: Basic Books.

Suen, H. K. (1988). Agreement, reliability, accuracy, and validity: Toward a clarification. *Behavioral Assessment, 10*, 343-366.

Swann, W. B., Stein-Seroussi, A., & McNulty, S. E. (1992). Outcasts in a white lie society: The enigmatic worlds of people with negative self-conceptions. *Journal of Personality and Social Psychology, 62*(4), 618-624.

Taylor, S. E. (1989). *Positive illusions: Creative self-deception and the healthy mind.* New York: Basic Books.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.

Taylor, S. E., & Brown, J. D. (1994a). "Illusion" of mental health does not explain positive illusion. *American Psychologist, 49*, 972-973.

Taylor, S. E., & Brown, J. D. (1994b). Positive illusions and well-being revisited: Separating fact from fiction. *Psychological Bulletin, 116*, 21-27.

Vaillant, G. E. (1977). *Adaptation to life.* Boston: Little, Brown.

Vaillant, G. E. (1986). *Empirical assessment of ego mechanisms of defense.* Washington, DC: American Psychiatric Press.

Vaillant, G. E. (1990). Repression in college men followed for half a century. In J. L. Singer (Ed.), *Repression and dissociation: Implications for personality theory, psychopathology, and health* (pp. 259-274). Chicago: University of Chicago Press.

Vaillant, G. E. (1992). *Ego mechanisms of defense.* Washington, DC: American Psychiatric Press.

Vaillant, G. E. (1994). Ego mechanisms of defense and personality psychopathology. *Journal of Abnormal Psychology, 103*, 44-50.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* San Diego, CA: Harcourt Brace Jovanovich.

Weinberger, D. A. (1990). The construct validity of the repressive coping style. In J. L. Singer (Ed.), *Repression and dissociation: Implications for personality theory, psychopathology, and health* (pp. 180-211). Chicago: University of Chicago Press.

Weinberger, D. A., & Schwartz G. E. (1990). Distress and restraint as superordinate dimensions of self-reported adjustment: A typological perspective. *Journal of Personality, 58*(2), 381-417.

Westen, D. (1990). Psychoanalytic approaches to personality. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 21-65). New York: Guilford.

Westen, D. (1991). Cultural, emotional, and unconscious aspects of self. In R. C. Curtis (Ed.), *The relational self: Theoretical convergences in psychoanalysis and social psychology* (pp. 181-210). New York: Guilford.

Westen, D. (1992). Social cognition and social affect in psychoanalysis and cognitive psychology: From regression analysis to analysis of regression. In J. W. Barron, M. N. Eagle, & D. L. Wolitzky (Eds.), *Interface of psychology and psychoanalysis* (pp. 375-388). Washington, DC: American Psychological Association.

Wylie, R. (1979). *The self-concept, Volume 2.* Lincoln: University of Nebraska Press.