Desirable Responding Triggered by Affect: Automatic Egotism?

Delroy L. Paulhus and Karen Levitt
University of British Columbia, Vancouver, Canada

Two experiments demonstrated an increase in socially desirable responding in the presence of affect-laden stimuli. Subjects responded "me" or "not me" to trait adjectives presented on a microcomputer. Affect was manipulated by pairing each trait adjective with a distractor word presented nearby. Some distractors were affect-laden (e.g., *sex, blood*); others were innocuous (e.g., *station, lake*). In Study 1, some trait adjectives were positive traits and others were neutral. Results showed that endorsements of positive traits were increased and speeded up by the affective distractors; denials of positive traits were reduced and slowed down by affective distractors. Both claims and denials of neutral traits were slowed by the affective distractors. In Study 2, positive, neutral, and negative traits were presented. The Study 1 results were replicated with parallel results for negative traits: Denials of negative traits were increased and speeded by the affective distractors, whereas claims of negative traits were reduced and slowed. This overall pattern of results was interpreted as a response-potentiation effect; that is, dominant responses were facilitated and subordinate responses were inhibited. Thus the net reaction to the presence of affective distractors was increased desirable responding. The high speed of this process suggests mediation by a fast-rising arousal or an attentional mechanism. The latter model suggests that self-perception automatically becomes more egotistical. This automatic egotism may underlie a variety of self-presentation phenomena, including certain defense mechanisms.

Many social interactions involve the processing of social information (e.g., categorization, decision making) while under the influence of affective states (e.g., fear, love, anxiety, sexual arousal). Indeed, one might argue that all important social judgments involve some simultaneous processing of affect and cognition. Only recently, however, has much research been directed toward the interplay of affect and social cognition (for a review, see Isen, 1984).

There is, of course, a classic literature on the effects of anxiety on task performance (Spence & Spence, 1966; Yerkes & Dodson, 1908). The anxiety effects appear to be a subset of the more general link between arousal and task performance (for a review, see Eysenck, 1982). The general finding is a response-potentiation effect; that is, arousal facilitates dominant responses and inhibits subordinate responses. For instance, Pallak, Pittman, Heller, and Munson (1975) found that when subjects were threatened with shock, a response-potentiation effect appeared on a subsequent Stroop task.

This research was supported in part by a grant from the Social Sciences and Humanities Research Council of Canada.

We gratefully acknowledge the helpful comments provided by Jennifer Campbell, Ray Corteen, Eric Eich, Jim Johnson, Dan Kahneman, Gord Logan, Carol Martin, Dale Miller, Janet Metcalfe, Jim Russell, Abe Tesser, and two anonymous reviewers.

Correspondence concerning this article should be addressed to Delroy L. Paulhus, Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5.

In the social psychology literature, such response-potentiation effects have been used as evidence that arousal ensues from states such as social facilitation (Zajonc, 1965), self-awareness (Liebling, Seiler, & Shaver, 1974), and cognitive dissonance (Pallak & Pittman, 1972). Other researchers have manipulated arousal to demonstrate that arousal facilitates such effects as aggression (Zillmann, 1971), prosocial behavior (Mueller & Donnerstein, 1977), cognitive dissonance (Cooper, Zanna, & Taves, 1978), and overeating (Slochower, 1976). Most of this work is based on a traditional activation concept of arousal where physiological arousal is said to energize behavior (Duffy, 1962; Malmo, 1959; Zajonc, 1965).

Affect and Social Cognition

Research on the influence of affective states on social cognition is now growing (see Isen, 1984). We do know a fair amount about the influence of mood states on cognition, primarily from the work of Isen and Bower. For example, inducing a positive mood was found to increase the use of cognitive heuristics (Isen, Means, Patrick, & Nowicki, 1982), to decrease complexity of decisions as well as decision time (Isen & Means, 1983), and to increase the use of inclusive categories (Isen & Daubman, 1984). Other researchers have shown that a positive mood induction improved self-regulation of problem solving (Kirschenbaum, Tomarken, & Humphrey, 1985).

Arousal effects on social cognition have also been found. Arousal has been shown to mediate such effects as just-world attributions (Thornton, 1984) and complexity of social judg-

Journal of Personality and Social Psychology, 1987, Vol. 52, No. 2, 245-259
 0022-3514/87/$00.75

ments (Paulhus & Lim, 1985). Clark, Milberg, and Erber (1984) demonstrated that altering subjects' arousal level affected their judgments of emotions in others' facial expressions and written statements. Only a couple of known studies have examined arousal effects on self-perception. Wegner and Giuliano (1981) showed that a neutral source of arousal (exercise) increased self-focus. Gollwitzer, Earle, and Stephan (1982) found that arousal increased beneffectance, that is, taking credit for success and denying responsibility for failure.

All of these studies involved creating a prolonged affective state before the primary task is performed and comparing the behavior of groups with different affective states. Many social interactions, however, involve rapidly fluctuating affective and nonaffective information from the same source. For example, the processing of an ordinary conversation entails a complex analysis of rapidly changing affect and cognition (Mandler, 1975). Person perception, too, is a dynamic process involving the joint influence of affective and nonaffective information on encoding social targets (Fiske & Taylor, 1984). Certainly, the understanding of defense mechanisms requires a clarification of the rapid interplay of affect and cognition (Blum & Barbour, 1979; Erdelyi, 1974; Hamilton, 1983).

Most relevant to this dynamic focus are the few studies that have manipulated affect within-subjects to examine effects on processing nonaffective information. For instance, Erdelyi and Appelbaum (1973) found that emotional distractors debilitated the encoding of neutral words. Using a dichotic-listening task, Nielsen and Sarason (1981) showed that affective words in the unattended channel tended to intrude and interfere with a shadowing task, particularly for anxious subjects. Corteen and colleagues, however, found that affect-laden words can produce physiological responses without causing errors in shadowing (Corteen & Wood, 1972). The lack of interference in the latter study is probably traceable to evidence that the affective words did not reach awareness (Corteen & Dunn, 1974).

In a visual parallel, our studies examine the effects of incidental affective words on a primary task. In contrast to the studies just cited, the primary task (self-ratings) is highly self-relevant. Thus a powerful paradigm is brought to bear on the issue of affect and self-perception.

Distractor Paradigm

The experimental technique used here permits an examination of the dynamic effects of affect on information processing. Subjects react to traits presented on a microcomputer by responding "me" or "not me" according to how self-descriptive they judge the traits to be. Subjects' responses and reaction times are recorded for subsequent analysis. Affective state is manipulated by simultaneously presenting, off to the right, a distractor word that subjects are told to ignore. The distractor word is then manipulated to be affect laden (related to sex or violence) or innocuous. The subject's performance on the primary task (self-ratings) may then be examined as a function of affective and innocuous distractors.

The first study presented here explored the effects of affective distractors on endorsements of positive and neutral traits. Four competing models (full interference, partial interference, arousal, and mood repair) make distinguishable predictions about the effects of affective distractors[1] on (a) endorsement rates of positive and neutral traits and (b) reaction times to endorsing and denying traits.

Arousal Model

Assume that physiological arousal has an energizing effect on behavior (e.g., Duffy, 1962). This assumption is often used to explain why arousal facilitates dominant (high probability) responses and debilitates subordinate (low probability) responses (for a review, see Eysenck, 1982). In the case of trait endorsements, a range of evidence suggests that the dominant response is to respond in a socially desirable fashion[2] (Edwards, 1957; Osgood, Suci, & Tannenbaum, 1957; Paulhus, 1981; Voyce & Jackson, 1977). That is, desirable traits have a high probability of being endorsed. In large samples of traits, the correlation between the desirability rating of the trait and its probability of endorsement is above .90 (Edwards, 1957).

The arousal model rests on the proposition that stimuli related to sex or violence evoke arousal. This arousal might be cortical, autonomic, or both (Lacey, 1967). Thus for Study 1 this model predicts that affective distractors will trigger arousal, which will then (a) enhance the rate and speed of claiming desirable traits, (b) debilitate the rate and speed of denying desirable traits, (c) have little effect on the choice of responses to neutral traits because neither response is dominant, and (d) increase response times to neutral traits because such decisions represent difficult tasks (Spence & Spence, 1966).

Mood-Repair Model

This model suggests a more controlled, purposive reaction: A negative affective state provokes attempts at mood repair with whatever response options are available (Clark & Isen, 1982). Here the trait-rating task gives an opportunity for mood repair via the claiming of desirable attributes. Thus the model predicts that distractors laden with negative affect[3] will provoke socially desirable responding: Positive traits will be endorsed more often with affective than with nonaffective distractors, and neutral traits will be unaffected. Because a controlled compensation should take time, the model predicts slower reaction times in the presence of affective compared to nonaffective distractors.

---

[1] It was difficult to come up with terms for the two types of distractors that are independent of the model being discussed. For one type, the label *affective* seems to be the most general and noncommittal term covering emotion, arousal, threat, anxiety, and so forth. For the other type, *innocuous* is preferable to *neutral* because we already use *neutral* to describe one category of traits. Our terms seem to be appropriate in discussing any of the four competing models.

[2] Socially desirable responding is actually a response unit comprising two S-R pairs: responding "me" to desirable traits and responding "not me" to undesirable traits.

[3] We do not know whether our affective distractors (sex and violence) provoke positive affect or negative affect, or simply orienting responses. The latter may evoke arousal (Scherer, 1982). However, given their incongruence with the laboratory setting, we might presume that students appraised the affective distractors as threats, that is, negative affect.

Attention Models

Attention is usually defined as a limited cognitive resource operating in serial fashion (e.g., Shiffrin, in press). In our distractor paradigm, distractor words should consume some attention, thereby leaving less for processing the primary task. Affective distractors, being more vivid and unexpected, should consume more attention than innocuous distractors. Hence, in the presence of affective distractors, less attention will be paid to the self-rating task. A variety of predictions then follow without any assumption that the distractors evoke an emotional response, which then influences the processing of the trait ratings.[4] For instance, affective distractors, by drawing attention, should be better encoded and therefore more easily recalled than innocuous distractors. Other predictions depend on the specific model of attention.

Full Interference Model

Several studies have shown a deterioration of current task performance in the presence of incidental affective stimuli (e.g., Erdelyi & Appelbaum, 1973; Nielsen & Sarason, 1981; Suedfeld, Erdelyi, & Corcoran, 1975). For our purposes, such studies suggest that affect-laden distractors will interfere with the trait-rating task. The most gross form of interference would be such that the trait term would not be encoded at all. The effect would be as if a blank trait term were presented, but the subject responded anyway because the instructions demanded it. No memory search for self-information would occur. Not even the desirability value of that trait would be processed. Therefore the usual high endorsement rate for desirable traits should drop to chance levels. The only systematic response pattern here would result from an overt response bias, for example, a right-hand bias or a "me" bias (acquiescence).

Partial Interference Model

A less radical form of interference may debilitate only the complex, more controlled processes, such as searching long-term memory for information about the self. The faster, more automatic processes, such as semantic activation, may operate intact (Deutsch & Deutsch, 1963; Treisman, 1960). Thus, a semantic property such as the desirability of the trait is more likely to be available than is specific information about whether the trait applies to the self. We know, for instance, that desirability ratings of traits are performed faster than actual self-ratings (Ferguson, Rule, & Carlson, 1983). Indeed, ratings of the likability of a variety of targets are performed faster than deciding whether one has even seen the target before (Zajonc, 1980).

This relatively high availability of trait desirability should favor its increased use in the trait-rating decision as the more controlled processes deteriorate. Thus, the partial interference model predicts an increased rate of socially desirable responding under affect-laden distractors. Moreover, reaction times to clearcut desirable or undesirable traits should be faster because they are unencumbered by a search for self-information. Neutral traits, however, may take more time under interference because, without social desirability as a guide, there is no easy basis for claiming or disclaiming them (Edwards, 1962). These predictions are similar to those derived from the arousal model.

Individual Differences

The most relevant personality variable here is the tendency to bias self-reports with socially desirable responding (SDR). Many variations of this concept appear in the literature and almost as many measures are available. Fortunately, the myriad measures have been shown to cluster around two factors: self-deception and impression management (Paulhus, 1984a). The first factor represents a nonconscious, honest form of bias (Lockard & Paulhus, in press; Sackeim & Gur, 1978). Individual differences in this self-deceptive style of processing information are assessed with the Self-Deception Questionnaire (SDQ) developed by Sackeim and Gur (1978). The construct validity of the SDQ is supported by a number of studies (Gur & Sackeim, 1979; Paulhus, 1982, 1984a; Sackeim & Gur, 1979; Winters & Neale, 1985). The concept is closely related to the idea of a repressive style: the tendency to avoid anxiety-arousing thoughts and stimuli (Byrne, 1964). In fact, Byrne's (1961) Repression-Sensitization (R-S) Scale loads highly on the self-deception factor (Paulhus, 1986).

The second factor of socially desirable responding, impression management, represents the more conscious, strategic form of self-presentation (Meehl & Hathaway, 1946; Paulhus, 1986; Sackeim & Gur, 1978). Paulhus (1984a) concluded that the best single measure of impression management is Sackeim and Gur's (1978) Other-Deception Questionnaire (ODQ).

In Study 1, both the R-S scale and the SDQ are used to determine individual differences in self-deceptive style (SDS); the ODQ is used to assess impression management. SDS is expected to moderate the effects of threat on self-descriptions. Previous work indicates that subjects high in SDS claim more desirable traits, avoid more negative information (Mischel, Ebbessen, & Zeiss, 1973), and are less disturbed by mild threats than are low-SDS subjects (Bell & Byrne, 1978; Paulhus, 1986). This tendency to shut out bothersome stimuli may explain why, under ego threat, high-SDS subjects show a severe decrement in their recall of distractor words (Markowitz, 1969). In our study, high-SDS subjects should be better able than their counterparts to ignore the affective distractors, with two results: The distractor manipulation will have little effect on them, and they will have poor memory for the affective-distractor words. Therefore, we predict that high-SDS relative to low-SDS subjects will claim more positive traits, be less affected by the affective distractors, and recall fewer affective distractors.

Study 1

Method

Subjects

Forty undergraduates, 20 women and 20 men, participated in this experiment. All subjects were recruited from the subject pool in a third-year psychology methods course.

[4] Some attention models hold that arousal itself reduces attention (Easterbrook, 1959; Kahneman, 1973). These more complicated models will be addressed later.

Overview

After completing three personality questionnaires, subjects rated 40 traits presented to them on a computer monitor as either descriptive or not descriptive ("me" or "not me"). Next to each presented trait was a distractor word, which they were instructed to ignore. The subjects' responses and reaction times were recorded for each of the 40 experimental trials. Finally, subjects were given an unexpected free-recall test for the distractor words that they had been initially instructed to ignore.

Apparatus

An Apple II-Plus microcomputer was used to present the trait–distractor word pairs. Subjects responded on the keyboard. For each of the 40 trials, the computer recorded the subject's button-press response ("me" or "not me") to the stimulus presented, as well as the reaction time.

Materials

Individual difference measures. Subjects completed three personality scales: the SDQ and the ODQ, both developed by Sackeim and Gur (1978), and an abridged version of Byrne's (1961) R-S scale (Paulhus & Levitt, 1983).

The SDQ is designed to measure individual differences in self-deceptive style. The rationale behind this scale is that if one adamantly denies a series of psychologically threatening statements (e.g., "I have sometimes hated one or both of my parents"), then one tends to process self-related information with an honest but defensive bias.

The ODQ measures a person's tendency to engage in other-deception (conscious lying). The items on this scale concern overt behavior. It is assumed that if one answers consistently in a socially desirable way about clear-cut events (e.g., "I always declare everything at customs"), then one is engaging in other-deception. The specific versions of the SDQ and ODQ administered here were the balanced versions constructed by Paulhus (1984a).

To help reduce the time subjects spent completing questionnaires, we used the abridged R-S scale (Paulhus & Levitt, 1983).[5] In this study, we have reversed the usual scoring direction so that a low score on the R-S scale indicates a tendency toward sensitization (anxiety expression), and a high score indicates a repressive style. Thus the R-S scale and the SDQ are now pointed in the same direction as measures of self-deceptive style.

Trait adjectives. The 40 traits used for the self-rating task on the computer were chosen from the Interpersonal Adjective Check List (Wiggins, 1979). By selecting Anderson (1968) likableness ratings above 5 (maximum = 7), three quarters of these were classified as positive traits (e.g., cheerful, sincere, likable, reliable). Using likableness ratings between 3 and 5, one quarter of the traits were neutral (e.g., secretive, dominant, defensive, childlike).

Distractors. The distractor word list consisted of 30 innocuous and 10 affect-laden words. Most of the distractor words were chosen from Heise's (1965) list of 1,000 words rated on their semantic profiles. Four affective distractors were added to the list. In a pretest, all 10 affective distractors were rated by three judges as being more emotional, threatening, and arousing than any of the 30 innocuous distractors. Examples of the innocuous distractor words are *lake, cloud, station, value, outside, limit, scene, step, public,* and *spirit.* The affective distractor words were *torture, penis, death, coffin, vagina, guts, suicide, blood, shut,* and *breast.* Affective and innocuous distractor words were equated for frequency and word length. During presentation, a total of 7 of 27 positive traits (and 3 of 13 neutral traits) were accompanied by affective distractors.

Procedure

Subjects filled out the three individual difference scales in a group session. They later reported individually to a laboratory to complete the experiment. Each subject was asked to sit down at the microcomputer. The instructions on the screen identified the experiment as a self-description study and directed the subject to respond to a series of traits, presented one at a time on the screen, by pressing the *M* key ("me") if the trait described them or the *N* key ("not me") if the trait did not describe them. Subjects were further instructed that they had to chose one or the other, even if the decision was difficult, and that they were to respond as quickly as possible.

The experimenter then told the subject to place one index finger on the *M* key and the other on the *N* key. She added that another word (the distractor word) would appear on the right of the trait. This was said to be part of the next study, and subjects were instructed to ignore it for this series of presentations (cf. Markowitz, 1969).

After all the instructions were given, the subject went through a practice run of three trait–distractor word presentations with the experimenter watching. The trait to be rated appeared in the center of the screen, and the distractor word appeared approximately 3 cm to the right of the trait (subtending a foveal angle of approximately 5° for the typical subject). As soon as the subject pressed either the *M* or the *N* button, the pair of words vanished. After 10 s the next pair appeared.

After making sure that the subject understood all of the instructions, the experimenter left the subject to complete the 40 experimental trials. Immediately afterward, the subject was given an unexpected free-recall test, which involved writing down in 3 min as many of the distractor words as could be remembered. Finally, the subject was completely debriefed and thanked for participating in the study.

Results

The major dependent variables were trait endorsement (proportion of "me" responses), reaction time, and recall of distractor words. Two independent variables were within-subjects variables: trait type (positive vs. neutral) and distractor type (affective vs. innocuous). The third independent variable was a between-subjects variable: level of SDS (high vs. low) as measured by median splits on the R-S scale and the SDQ. Unless otherwise indicated, the reported statistics are based on the R-S split, and the SDQ results were consistent.

Sex of subject (women, high; men, low) correlated −.19 and −.24 with the SDQ and the (reversed) R-S scale, respectively. However, when included with either personality scale in the analyses of variance (ANOVAs) reported here, no sex effects were significant. Therefore, to simplify reporting, sex was removed from the ANOVAs.

Reaction Time

For purposes of analysis, we calculated each subject's mean reaction time across all trials (trait presentations) within each condition.

Analyses of variance. The factors in a 2 × 2 × 2 ANOVA on reaction time were Trait Type × Distractor Type × SDS. There were only two significant effects on reaction time. One was a strong main effect for trait type, $F(1, 36) = 14.32$, $p < .001$.

[5] In a sample of 306 students, this shortened version of the R-S scale showed Cronbach's alpha of .84 and correlated .90 with the full R-S scale.

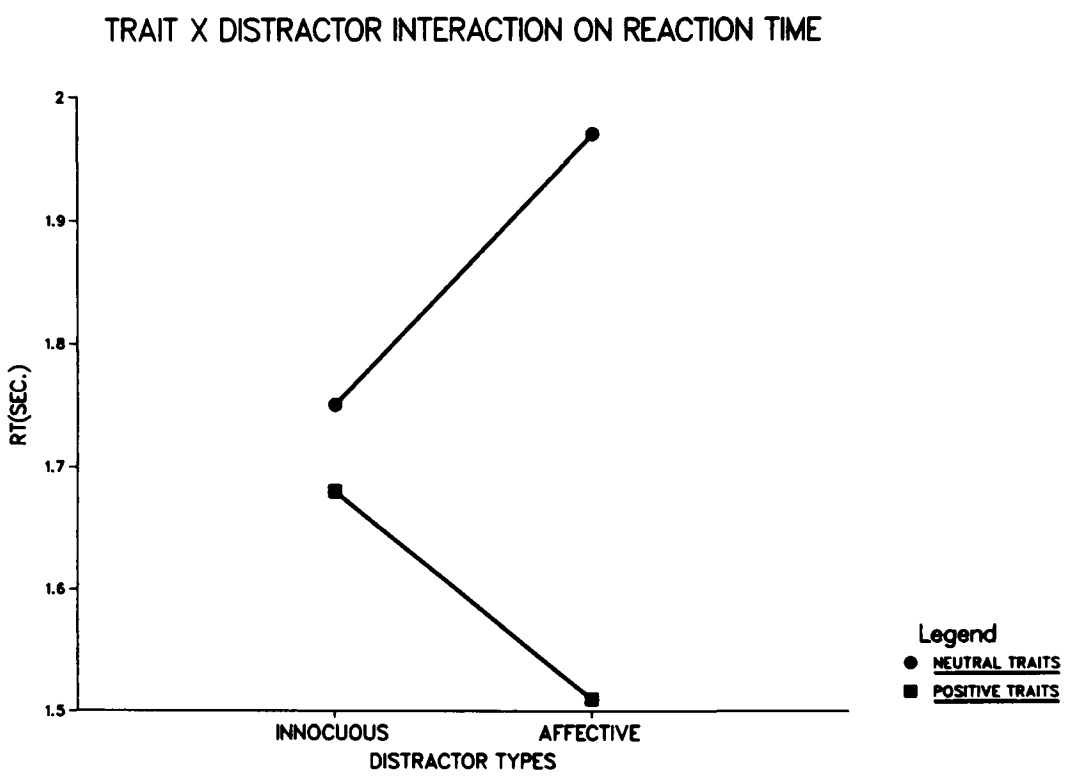

Figure 1. The effects of trait type and distractor type on time to make self-ratings ("me" or "not me") in Study 1.

Reaction times were faster for positive traits (M = 1.60 s) than for neutral traits (M = 1.86 s). As is clear from Figure 1, there was a significant Trait Type × Distractor Type interaction, $F(1, 36) = 6.93$, $p < .01$. For positive traits, the effect of the affective distractor was to speed up reaction times from 1.68 to 1.51 s, $t(39) = 1.96$, $p < .06$. In contrast, for neutral traits the affect slowed reaction times from 1.75 to 1.97 s, $t(39) = 2.07$, $p < .05$.

We then subdivided the responses to positive traits into "me" and "not me," that is, desirable and undesirable responses. Only the "me" responses sped up under affect (from M = 1.65 s to M = 1.40 s), $t(38) = 3.40$, $p < .01$. In contrast, "not me" responses showed a trend to slow down (from M = 1.73 s to M = 1.80 s), $t(38) = 1.21$, *ns*. In a similar subdivision of the neutral traits, the "me" responses slowed down (from 1.71 s to 1.94 s), $t(38) = 3.01$, $p < .01$, but the "not me" responses also slowed down (from 1.77 s to 1.99 s), $t(38) = 1.81$, $p < .08$.

Correlations. Reaction time was significantly correlated with SDS ($r = -.36$, $p < .01$) but not with impression management ($r = .02$, *ns*). That is, subjects with a self-deceptive style responded more quickly; this pattern was sustained across positive and neutral traits and affective and innocuous distractors. The corresponding main effect was not found after dichotomizing the SDS scales for use in the ANOVAs. A weaker result is not surprising given the diminished power after dichotomizing a personality variable (Humphreys, 1978).

Trait Endorsement

Trait endorsement was defined as the proportion of "me" responses out of all trait presentations within a condition. For example, the mean trait endorsement rate was .86 for positive traits and .45 for neutral traits. To minimize heterogeneity of variance, each proportion was transformed by 2 × arcsin $P^{1/2}$, following Winer (1971, p. 400). The proportions reported in the text, however, are untransformed.

Analyses of variance. The three factors in a 2 × 2 × 2 ANOVA on trait endorsement were, again, Trait Type × Distractor Type × SDS. There were only two significant main effects. As expected, there was a strong main effect for trait type: Subjects endorsed positive traits (M = .86) significantly more often than neutral traits (M = .45), $F(1, 36) = 147.3$, $p < .001$. There was also a strong main effect for distractor type: Subjects endorsed traits more often when the distractor was threatening (M = .71) than when it was innocuous (M = .61), $F(1, 36) = 16.8$, $p < .01$.

There were also two significant interactions. Self-deceptive style interacted with trait type, $F(1, 36) = 5.68$, $p < .03$, such that high-SDS individuals claimed more positive traits than did low-SDS individuals (M = .91 vs. .82), whereas high-SDS people claimed slightly fewer neutral traits than did low-SDS people (M = .42 vs. .47).

Most important was the second interaction, that between trait type and distractor type, $F(1, 36) = 6.65$, $p < .02$. Note in Figure 2 that for positive traits, the effect of the affective distractor was to increase significantly the mean trait endorsement from .80 to .93, $t(38) = 2.61$, $p < .01$. For neutral traits the affective distractor effected a nonsignificant rise in mean trait endorsement from .42 to .48, $t(38) = 1.31$, *ns*.

Correlations. Endorsement of positive traits was significantly correlated with SDS ($r = .35$, $p < .05$) but not with im-

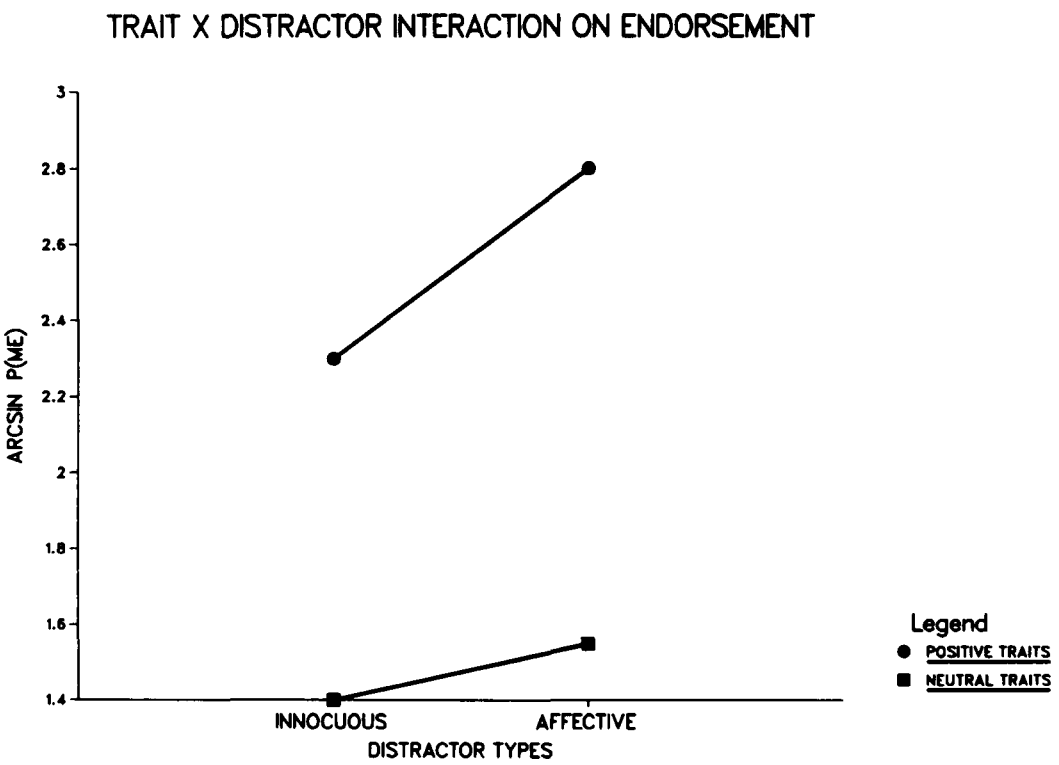

Figure 2. The effects of trait type and distractor type on proportion of "me" responses (i.e., the probability of claiming traits) in Study 1.

pression management ($r = .10$, *ns*). Thus subjects with a self-deceptive style claimed more positive traits than did low-SDS subjects.

Recall of Distractors

Analyses of variance. Free recall was indexed by the proportion of distractors correctly recalled. For each subject, separate indexes were calculated for affective and nonaffective distractors. To reduce heterogeneity of variance, each proportion was transformed by 2 arcsin $p^{1/2}$. The two factors in a 2 × 2 ANOVA were distractor type and SDS. Only the main effect for distractor type emerged significant, $F(1, 36) = 42.39$, $p < .001$. A much higher proportion of affective distractors ($M = .16$) than innocuous distractors ($M = .02$) was recalled.

Correlations. None of the personality variables predicted free recall of innocuous distractors. Recall of affective distractors was correlated with the ODQ, the impression management scale ($r = -.32$, $p < .03$), but not with the SDS scales, SDQ and R-S.

Discussion

Despite instructions to ignore the distractors, it is clear that some subjects were processing them to some degree. First, the free-recall results showed some memory for the distractors, mostly the affective ones. This superior retention of affective words suggests that the processing of distractors had reached the semantic level during presentation (Deutsch & Deutsch, 1963; Treisman, 1960). Second, and more important, the type of distractor significantly influenced subjects' responses and reaction times on the trait-rating task. Again, it is difficult to see how such effects could occur without semantic activation.[6]

Competing Models

The results were not consistent with the full interference model, which predicted that affective distractors would produce general response debilitation, that is, slower reaction times and chance levels of desirable responding.[7] Nor were the results predictable from the mood-repair model. Although the increased desirable responding is consistent, the polarization of reaction times does not follow from the model. Rather, the model incorrectly predicted a general slowing of reaction times under affective distractors. To salvage this model, one would have to allow that mood repair is automatic.

The results do support the arousal model. Traits having an

[6] There is certainly no guarantee that all subjects read all the distractors. First, not all semantic priming is automatic (Posner & Snyder, 1975). Moreover, semantic priming effects could result from subjects "reading" the prime on as few as 5% or 10% of the trials (Kahneman & Treisman, 1984).

[7] One might argue that the increased claim rate for positive traits resulted from an increase in the overt-response bias to choose the right-hand key (or choose the "me" response). However, the proportion of "me" responses to neutral traits actually dropped slightly under affective distractors.

initially high claim rate (positive traits) were facilitated both in terms of endorsement rates and reaction times. In contrast, traits with an indifferent claim rate (neutral traits) were inhibited in the presence of affective distractors. To compare reaction times of dominant with those of subordinate responses, we broke down the positive trait responses into "me" (dominant responses) and "not me" (subordinate responses). The reaction times were clearly in accord with the arousal model. "Me" responses sped up and "not me" responses slowed down with emotional distractors.

The results are also consistent with the partial interference model. Reduced attention is said to debilitate the more complex, controlled processes, leaving the less vulnerable automatic processes to predominate. In the distractor paradigm, salient distractors disturb memory-search processes, leaving semantic properties like trait desirability to dominate the choice of response. Thus, affective distractors should enhance socially desirable responding, as found.

Individual Differences

Subjects scoring high on the ODQ, a measure of conscious impression management, reported seeing fewer affective distractors than did those scoring low on the scale. Apparently some concerned subjects were reluctant to report to the experimenter that they had seen words associated with sex and violence. This result brings into question the value of the free-recall measure in this context. Preferable would be a recognition-memory test, which permits the assessment of memory accuracy free of any response bias to claim or deny recognition (Swets, 1964). When untangled from response bias, the memory measure may then show the expected correlations with the SDS measures.

The two measures of self-deceptive style (R-S scale and SDQ) showed virtually identical patterns. The results were consistent with previous work in showing that high-SDS subjects claimed more positive traits than did low-SDS subjects (Byrne, 1964; Sackeim & Gur, 1979). Another finding was that on both measures, high-SDS subjects were faster than low-SDS subjects on the self-rating task. Consistent with these results are previous studies showing slow reaction times in subjects with traits linked to low SDS: socially anxious subjects (Turner, 1978) and depressed subjects (Friedman, 1964; Hall & Stride, 1954; Martin & Rees, 1966). Finally, high-SDS subjects did not show the predicted immunity from the distractor manipulation. Perhaps they could not ignore the distractors as easily as they did in the Markowitz (1969) study. Use of recognition memory, as recommended earlier, may tell us whether high- and low-SDS subjects attended equally to the distractors.

Study 2

Study 2 was designed to replicate and extend the results of Study 1. Two major changes were made. First, to extend the trait domain, negative traits were added to the list of positive and neutral traits presented. Second, to more validly assess memory for distractors, a recognition test was used instead of free recall. On recognition tests, subjects rate their recognition of a list of items, only some of which have previously been presented. Subjects' recognition ratings are then subjected to signal-detection analysis wherein memory accuracy is distinguished from a response bias to claim or disclaim recognition (Swets, 1964).

Given the pattern of results found in Study 1, the following predictions were made: The presence of affective distractors would facilitate endorsements of positive traits, inhibit denials of positive traits, and not alter the endorsement rates of neutral traits but increase reaction times. In addition, on the basis of the arousal and partial interference models, we made the following predictions for negative traits: Affective distractors would facilitate denials and inhibit endorsements. Finally, given the Markowitz (1969) results, we predicted that high-SDS subjects would show poorer memory than low-SDS subjects would for affective distractors now that the memory measure was uncontaminated.

Method

Subjects

Seventy subjects, 35 women and 35 men, were recruited from first-year psychology classes. All subjects received class credit for participation.

Materials

The 48 traits used for the self-rating task were divided equally into three categories: socially desirable (positive) traits, socially undesirable (negative) traits, and neutral traits. These traits were chosen from a list of 208 adjectives based on their social-desirability rating (Kirby & Gardner, 1972) and from a list of 555 personality-trait words that had been rated on likableness (Anderson, 1968).

Sixteen of the 48 distractor words were classified as affective, and the remaining 32 were considered innocuous. The same distractor words as those used in the first experiment were used here, with additional ones being chosen from Heise's (1965) list of words and their semantic profiles. Five of the affective distractors were associated with sex (penis, sex, vagina, slut, and breast), 4 with violence (guts, blood, torture, knife), 5 with death (cancer, corpse, coffin, suicide, death), and 2 were miscellaneous (hate, fail).

During presentation the traits and distractors were paired as follows. The 48 distractors were presented in a fixed order with affective distractors (ADs) appearing on Trials 3, 8, 10, 14, 15, 18, 21, 23, 25, 26, 29, 31, 34, 38, 42, and 46. Innocuous distractors (IDs) appeared on the remaining trials. Sixteen trios of trait adjectives were presented in a fixed order, starting with a random trio. Within each trait trio (one positive, one neutral, one negative), the order of presentation was random. Thus, subjects could vary somewhat in the exact number of positive, neutral, and negative traits paired with ADs and IDs. The final percentages of trials with combinations positive trait–ID, neutral trait–ID, negative trait–ID, positive trait–AD, neutral trait–AD, and negative trait–AD were 10.8, 11.0, 11.5, 22.4, 22.3, and 21.9, respectively.

Procedure

Subjects were given the SDQ, ODQ, and the Short R-S scale to complete when they arrived at the lab. The computer presentation of the trait–distractor word pairs was the same as in the first study, with two exceptions. First, there were 48, instead of 40, experimental trials. Second, the distractor words were presented about 1 cm directly above the trait (foveal angle = 1.6°) in the center of the screen, as opposed to their being horizontal to each other. The latter change was made to extend the generality of the distractor effects.

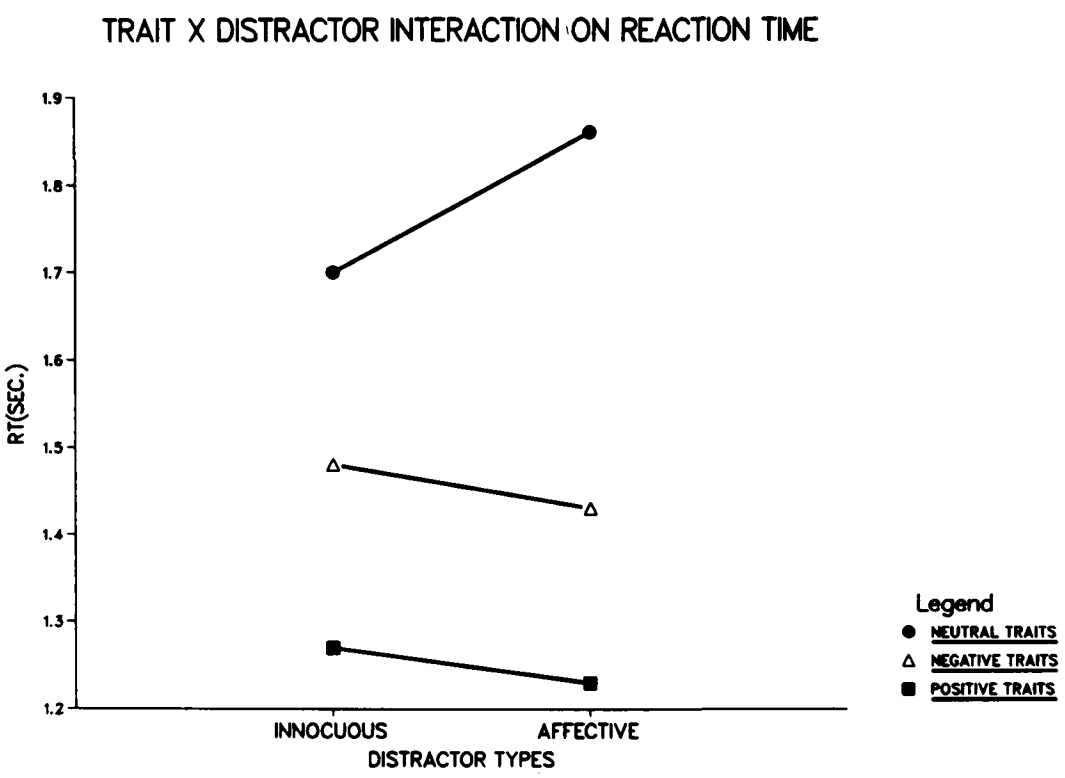

Figure 3. The effects of trait type and distractor type on time to make self-ratings ("me" or "not me") in Study 2.

After the 48 experimental trials, 50 of the 70 subjects were given a recognition test for the distractors. The recognition test consisted of a list of 60 words, 40 of which had been presented as distractor words and 20 of which had not. Fifteen of the words that had been presented and 10 of the words that had not been presented as distractor words were affective; the rest of the words in the recognition task were innocuous. Subjects were asked to rate on a scale of 1 (*definitely didn't see*) to 5 (*definitely saw*) how confident they were that they had seen the word on the screen during the trait-rating session. In all other respects, the methodology of Study 2 was the same as that of Study 1.

Results

As in Study 1, the major dependent variables were trait endorsement (proportion of "me" responses), reaction time, and memory for distractor words. Two of the independent variables were within-subjects variables: trait type (positive, neutral, negative) and distractor type (affective vs. innocuous). The third variable was a between-subjects variable: self-deceptive style (high vs. low) as measured by median splits on the R-S scale and the SDQ. The F ratios and probability levels refer to R-S results, but the SDQ gave similar results except where noted. When sex was included, no sex differences emerged. Therefore, for simplicity, sex was dropped from all ANOVAs. The correlations of the R-S, SDQ, and ODQ with sex (men, low; women, high) were .28, .04, and −.04, respectively.

Reaction Time

A subject's reaction time for each condition was the subject's median reaction time across all trait presentations within that condition. The mean of the median reaction times across all subjects was 1.47 s.

The three factors in a 3 × 2 × 2 ANOVA were trait type, distractor type, and SDS. There were only three significant effects. First, the main effect for trait type was strong, $F(2, 132) = 33.11$, $p < .001$. Reaction times were fastest for positive traits ($M = 1.25$ s) and negative traits ($M = 1.46$ s) and slowest for neutral traits ($M = 1.78$ s). The interaction Trait Type × SDS was also significant, $F(2, 132) = 2.99$, $p < .03$. For positive and negative traits, high-SDS subjects were faster than low-SDS subjects, whereas for neutral traits, high-SDS subjects were slower than low-SDS subjects.

Finally, as predicted, the Trait Type × Distractor Type interaction was significant, $F(2, 132) = 2.80$, $p = .03$, one-tailed. Reaction times to neutral traits slowed down with the affective distractor, whereas reaction times to positive and negative traits sped up slightly. The pattern of mean reaction times is depicted in Figure 3.

To test directly for potentiation effects, dominant responses must be compared with subordinate responses. We pooled "me" responses to positive traits with "not me" responses to negative traits and compared them with "not me" responses to positive traits pooled with "me" responses to negative traits. The results, depicted in Figure 4, support the potentiation hypothesis. Dominant responses sped up significantly under affect, $t(39) = 3.12$, $p < .01$, and subordinate responses slowed down under affect, $t(39) = 2.01$, $p < .05$.

The speeding up of dominant responses held for positive traits, $t(39) = 2.01$, $p < .05$, and negative traits, $t(39) = 1.70$,

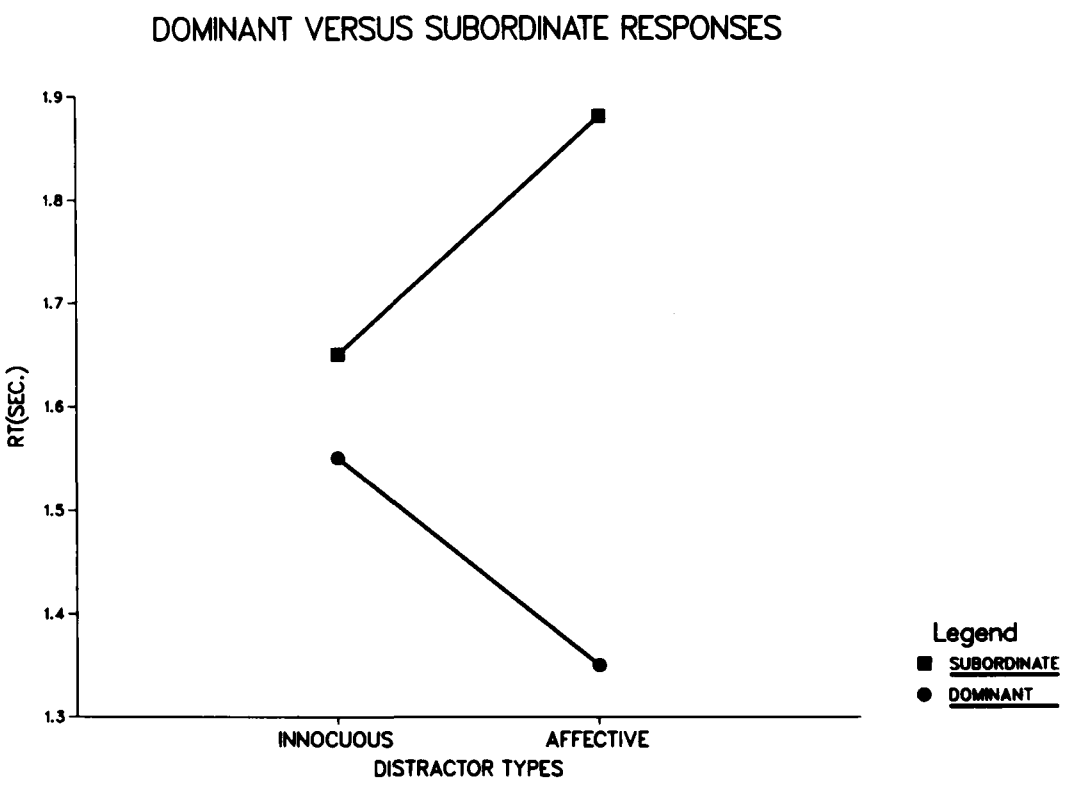

Figure 4. The effects on response time of distractor type and type of response in Study 2. (Dominant responses include "me" responses to positive traits and "not me" responses to negative traits; subordinate responses include "me" responses to negative traits and "not me" responses to positive traits.)

$p < .08$. The slowing of subordinate responses held for positive traits, $t(39) = 1.93$, $p < .06$, and for negative traits, $t(39) = 2.22$, $p < .05$.

Trait Endorsement

Trait endorsement was defined as the proportion of "me" responses out of all trait presentations within a condition. The overall mean endorsement rate was .53. To minimize heterogeneity of variance, all proportions were transformed by 2 arcsin $p^{1/2}$, following Winer (1971, p. 400).

The three factors in the 3 × 2 × 2 ANOVA were the same as for the reaction time ANOVA: trait type (within), distractor type (within), and SDS (between). As expected from our selection of traits, the main effect for trait type was strong, $F(2, 132) = 318.8$, $p < .001$. Positive traits were endorsed ($M = .92$) more than neutral traits ($M = .56$), which were endorsed more than negative traits ($M = .17$). The main effect for SDS was significant, $F(1, 66) = 7.59$, $p < .01$. High-SDS subjects ($M = .50$) endorsed fewer traits than did low-SDS subjects ($M = .60$).

Again, the Trait Type × SDS interaction was significant, $F(2, 132) = 6.27$, $p < .003$. The means are depicted in Figure 5. High-SDS subjects claimed fewer negative ($M = .09$) and neutral traits ($M = .50$) than did low-SDS subjects ($Ms = .25$ and .63, respectively). The two groups claimed an equally high number of positive traits ($M = .92$).

Finally, the Trait Type × Distractor interaction was significant, $F(2, 132) = 7.6$, $p < .001$. As seen in Figure 5, with the affective distractors, subjects claimed more positive traits, $t(39) = 4.14$, $p < .01$, fewer negative traits, $t(39) = 2.72$, $p < .01$, and about the same number of neutral traits, $t(39) = .22$, *ns*.

Recognition of Distractors

In contrast to free recall, recognition tests of memory permit a signal-detection analysis (Swets, 1964), which allows the separate assessment of sensitivity and response bias. In memory studies, measures of sensitivity (e.g., d') index how accurate the individual is in distinguishing previously presented material (signal) from material that was never presented (noise). We used the formula $d' = Z$ (false alarms) $- Z$ (hits), as detailed in McNicol (1972).[8]

Measures of response bias (e.g., β) index the subjects' tendency to say no ("I don't recognize the target") independent of the subject's sensitivity in recognition. We used the formula $\beta = I$ (hits)/I (false alarms), where $I(z)$ represents the ordinate on the unit normal curve at z. Response-bias scores tend to reflect the various payoffs and punishments accruing from hits, false alarms, misses, and correct rejections. A high response bias sug-

[8] The overall d' was the average of four values, a d' for cutoffs at $c = 2, 3, 4, 5$. A d' at $c = 1$ is inappropriate because P(Hits) and P(FA) are necessarily both unity.

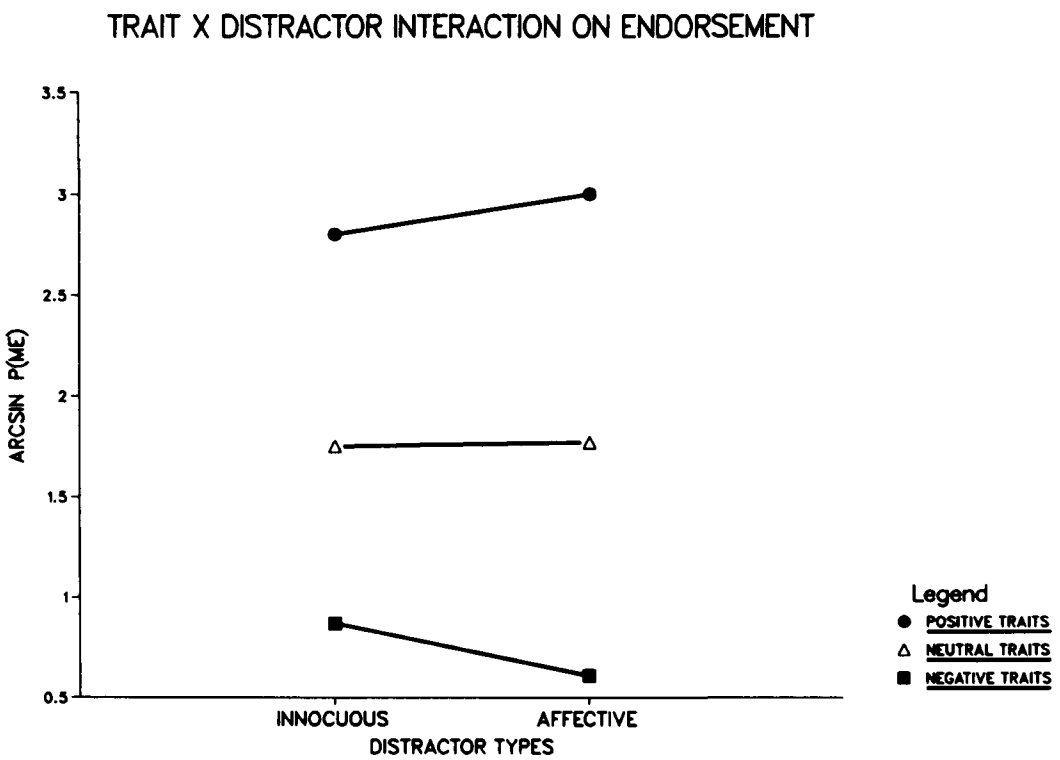

Figure 5. The effects of trait type and distractor type on proportion of "me" responses (i.e., the probability of claiming traits) in Study 2.

gests a greater fear of making false alarms (saying yes too often) than of making misses (saying no too often). In this study the correlation of β with d' was .41 ($p < .01$). That is, subjects with accurate recognition memory showed more bias to say *no* than did subjects showing inaccurate memory.

Analysis of variance on d'. The 2 × 2 ANOVA on recall sensitivity involved distractor type and SDS. The main effect for distractor type was significant, $F(1, 46) = 4.45$, $p = .04$. Subjects were more accurate at recognizing affective distractors ($d' = 2.3$) than innocuous distractors ($d' = 1.9$).

Correlations with d'. Accurate recognition of distractors was significantly lower for high-SDS subjects ($r = -.27$, $p < .03$). Marginal results appeared for affective distractors ($r = -.20$, $p < .09$). Surprisingly, recognition of innocuous distractors showed slightly stronger correlations with SDS ($r = -.27$, $p < .04$).[9]

Analysis of variance on β. The Distractor Type × SDS ANOVA on β showed no main effects, but a significant interaction emerged, $F(1, 46) = 5.39$, $p < .03$. For affective distractors, high-SDS subjects showed a higher response bias ($\beta = 8.51$) than did low-SDS subjects ($\beta = 6.77$); that is, the former tended to say *no* ("I didn't see it") more often than the latter. For innocuous distractors, high-SDS subjects showed a lower bias ($\beta = 7.10$) than low-SDS subjects ($\beta = 9.65$) to deny recall. The reported results occurred with SDQ as the measure of SDS; a similar but nonsignificant pattern occurred using the R-S scale.

Correlations with β. Despite the Distractor Type × SDS interaction, response bias to deny recognition showed no significant correlations with measures of SDS. However, for affective distractors the correlation between β and sex emerged significant ($r = -.26$, $p < .04$): Men showed more denial than women did.

Discussion

The results of Study 2 sustained our interpretation of Study 1 and clarified some ambiguities. The predictions of the arousal and partial interference models again found support. The presence of affective distractors was found to facilitate desirable responses and inhibit undesirable responses.[10] The effect appeared for negative traits and replicated the Study 1 results for positive and neutral traits. The potentiation effects appeared on reaction times as well as on endorsement rates.

The arousal explanation remains compelling. Because desirable traits have a high claim rate without affective distractors ($M = .86$), the dominant response is to claim them. Similarly, the low claim rate for undesirable traits ($M = .21$) makes denial the dominant response. This close correspondence between the endorsement rate of traits and their social desirability was estab-

[9] One might expect the correlations of SDS with d' to be larger for affective than for innocuous distractors. One explanation is that high-SDS subjects have a superior ability to shut out any distraction, affective or otherwise. Another possibility is that the high-SDS subjects, once noticing that the distractors were sometimes threatening, shut them all out.

[10] Detailed analyses suggested that the sex- and violence-related distractors showed the strongest effects.

lished some time ago (Edwards, 1953). Over a broad range of traits the correlation between desirability and endorsement rate is above .90 (Edwards, 1957). This natural confounding is likely a result of desirable responding being a well-learned, highly practiced response (Heilbrun, 1964).

Consequently, the net effect of the arousal process was inflated desirable responding: Subjects claimed more of the positive and fewer of the negative self-descriptors. Nonetheless, we cannot tell from this experiment whether arousal facilitates desirable responding directly, or only indirectly as a consequence of desirable responses typically being dominant (having high probability). Future work must attempt to separate desirability from frequency.

Note that this arousal must be fast to rise: The affective distractors channeled responses that averaged less than 1500 ms. We therefore propose the term *fast-rising arousal* to label the arousal-like process that mediates the observed pattern. This arousal resembles the category of cortical arousal, which is fast acting and has been linked to orienting responses and evoked potentials (Kimmel, Van Olst, & Orlebeke, 1979; Lacey, 1967). In comparison, autonomic or visceral arousal is slower to respond and associated with deep emotions (Lacey, 1967; Mandler, 1975).

Corroborative evidence to support an arousal model comes from other studies conducted in our laboratory. A more direct manipulation of arousal, loud white noise, was shown to increase the importance of evaluation in subjects' judgments (Paulhus & Lim, 1985). Similar effects resulted when exercise was the source of arousal (Paulhus, Lim, Reid, & Murphy, 1986). This converging evidence from studies using a variety of sources of arousal leads us to favor an arousal model (Paulhus & Suedfeld, in press). Nevertheless, we cannot say at this point why such different forms of arousal should yield parallel effects.

Unlike the arousal model, attentional models have much to say about recall for distractors. Therefore, before we evaluate the attentional models, let us review the memory results.

Memory for Distractors

The recognition-memory technique used in Study 2, followed by signal-detection analysis, permitted the independent assessment of memory accuracy and response bias. Using the sensitivity statistic d', we found a higher accuracy for recall of affective distractors. Considering both studies, we now have some confidence about this hypersensitivity for affective distractors. However, this study cannot pinpoint the stage of processing where this sensitivity is regulated. The effect is presumably some unknown combination of perceptual, encoding, consolidation, and retrieval effects (Erdelyi, 1974).

In previous studies, sensitization to incidental stimuli was typically associated with deterioration on the primary task (Blum, Geiwitz, & Stewart, 1967; Lewis, 1970; Suedfeld et al., 1975; Treisman, 1960). For instance, Erdelyi and Appelbaum (1973) found that an affective distractor debilitated the subsequent recognition of neutral pictures. Most comparable to our study is that by Nielsen and Sarason (1981), who also found a high recall sensitivity for affective words. Distractors were presented in the unattended channel of a dichotic-listening task. The recall accuracy was higher for sexually explicit distractors than for other distractors. That study and ours agree in that affective words can intrude on a primary task.[11] On Nielsen and Sarason's shadowing task, however, these intrusions degraded the primary task, suggesting that affective distractors simply hamper other processes. We would argue that for all previous studies, the dependent measures (e.g., shadowing) are simply not sensitive to response-potentiation effects such as those observed in our study.

Attentional Models

The attentional models introduced earlier required no affective state such as arousal. The affective distractors simply distract, without triggering arousal.[12] Our failure to find interference effects on the primary task did rule out the full interference model; however, the partial interference model is still viable.[13]

The model assumes that affective distractors inhibit the memory search process but not semantic activation. Thus trait desirability comes to dominate judgments, yielding the observed response-potentiation effect. Although this model is not unreasonable, there is one piece of evidence that fails to support it. The attentional model predicts that subjects who attended most to the affective distractors (as indicated by a high recognition rate) should show the potentiation effect more clearly. However, high-SDS subjects, who showed faster reaction times and poorer recognition of distractors, exhibited distractor effects to the same degree as did low-SDS subjects.[14] To obtain further evidence, we calculated a more direct measure of how much subjects had attended to the distractors: We assessed recognition level as an individual difference variable by a median split on d' for affective distractors. Recognition level did not interact with the enhancement effect for either dependent measure. This suggests that attention to the distractors did not mediate the response-potentiation process: They had an impact on processing and memory without extra attention.[15] Some affective process such as arousal seems necessary for a complete explanation. Perhaps a measure of individual differences in arousal to affective distractors, had we collected it, would have moderated the distractor effects.

---

[11] At least one study has shown affective responses without interference on shadowing (Corteen & Wood, 1972): Presumably the affective response was insufficient to attract attention (Corteen & Dunn, 1974). Other studies have shown interference on shadowing owing to synonyms on the unattended channel (Lewis, 1970; Treisman, Squire, & Green, 1974).

[12] An uncontrolled possibility is that under affective distractors, subjects are adjusting their speed–accuracy trade-off criterion. Because we have no measure of trait accuracy (whatever that means), we cannot evaluate this possibility.

[13] Yet another model is suggested by Mandler (1975) and Sanders (1981): Attentional conflict may induce arousal.

[14] It is intriguing to speculate on why high-SDS subjects cannot remember distractors that had manipulated their behavior. Such subjects may not consolidate the trace of affective information during the encoding opportunity. Or, these subjects may tag this information such that subsequent retrieval is impaired.

[15] We also examined possible moderator effects of differential recognition of affective and innocuous distractors. Again, no interaction was found.

Perhaps the best model is a hybrid based on the proposition that arousal increases attentional selectivity (Cohen, 1978; Easterbrook, 1959; Kahneman, 1973). In this model, even without a loss of attention to affective distractors, these distractors may alter the way the trait task is processed. Such a model has recently been applied to social judgments and defense mechanisms (Paulhus, 1984b; Paulhus & Suedfeld, in press). The gist is that arousal reduces cognitive complexity, that is, the number of dimensions used in decisions. In social judgments, including self-descriptions, evaluation is the primary dimension (Osgood et al., 1957). Wallsten and Barton (1982) showed that the primary dimension is the last to drop out under increasing time pressure. Thus, when attention to the primary task is reduced, the importance of evaluation will increase. The normally positive self-evaluations of most subjects will then loom even larger in self-perceptions. In short, attentional selectivity in rating the self should increase socially desirable responding, as observed.

Automatic Egotism?

The various definitions of automaticity all emphasize the lack of intention involved in initiating certain processes (Logan, 1980; Posner & Snyder, 1975; Shiffrin & Schneider, 1977). Automatic processes can operate in parallel with other tasks and are not affected by a reduction in attention (for a review emphasizing social cognition, see Bargh, 1984). Recent analyses, however, suggest a continuum rather than a sharp division between automatic and controlled processes (Kahneman & Treisman, 1984; Shiffrin, in press).

There is already evidence that affective stimuli can evoke automatic effects (see Bargh, 1984). There is also evidence that trait attributions from behavioral information are automatic (Winter, Uleman, & Cunniff, 1985). Let us therefore consider the possibility that the inflated egotism we observed is an automatic rather than a controlled process. It has been claimed that affect can automatically trigger an egotistic style of information processing (Gollwitzer et al., 1982; Paulhus, 1984b; Paulhus & Suedfeld, in press). Here we argue that a more positive self-perception is induced, at least in subjects with a positive self-esteem. (The automaticity may not include the actual overt response, which may itself be controlled.) Given that our subjects happened to be giving self-descriptions, this egotism was manifested as self-enhancement on positive traits and defensive denial on negative traits.

Several findings are consistent with automaticity. First, given the instructions to ignore the distractors, any effects owing to distractors occurred in spite of subjects' good intentions. Second, given the high speed of responses, it is impossible for subjects to have deliberated about the implications of particular distractors for the claim or denial of a trait. Third, the egotistic response was rather indiscriminate, if not inappropriate: Subjects reported themselves to be nicer people in the presence of words like *penis* and *blood*. This egotism may be the response to a general category of threatening events (Bargh & Bond, 1983, as cited in Bargh, 1984). Finally, as noted earlier, our reanalyses showed no influence on the egotism effect of individual differences in attention to the affective distractors.

General Discussion

Earlier research on the influence of affect on cognition has emphasized the prolonged effects of mood states (Isen, 1984). Our work suggests a more dynamic relation between affect and cognition. Affective states can rise quickly to influence concurrent cognitive processes. Thus, the immediate affective properties of a stimulus can influence the way that stimulus is processed. The effects observed here are consistent with the claim that affective responses may sometimes precede cognitive responses to a stimulus (Izard, 1977; Leventhal; 1974; Wilson, 1979; Zajonc, 1980). The results also have wide implications for social judgment.

Most social interactions require making judgments about others under the influence of affect. Consider the speed at which social discussions and arguments can move. If the affect associated with a speaker's word choice or phrasing automatically inflates egotism, then the rest of the speaker's sentence may be processed under polarized evaluation. Moreover, the listener's response may be generated within a polarized evaluative system. Similar effects on interactions might result from affective reactions to the speaker's appearance or a brief nonverbal behavior like an attentive gaze or a grimace.

The claim for automatic egotism has implications for the nature of impression management, which is typically viewed as a deliberate strategy (e.g., Jones & Pittman, 1982; Snyder, 1974; Tedeschi, 1981). Rather than being strategic, some of what passes for impression management may be automatic and unintentional defensiveness elicited by threat. Such effects might be manifested as short-lived outbursts of self-aggrandizement or defensive overreactions. Thus we can account for egotistic behaviors that, far from impressing people, foster a negative image. Bragging and defensiveness, for instance, are frequently observed despite the fact that they are seldom rewarded.

Defense Mechanisms

If, as we have argued, the impact of affective stimuli is automatic, then there are also implications for defense mechanisms. One of the traditional arguments against the logical possibility of defense mechanisms is the following: An individual must accurately recognize a threat to defend against it, but given the accurate recognition, it is too late. This argument has been applied to perceptual defense (Howie, 1952), repression (Ichheiser, 1960), and self-deception (Demos, 1960; Sartre, 1943/1956, p. 52). As noted some time ago (Dollard & Miller, 1950), this argument is bypassed by a threat-signal analysis. If the organism preattentively perceives that some threat (any threat) is present, then the cognitive machinery adjusts automatically and the ensuing processing of information is conducted in a fashion appropriate for meeting threats, that is, defensively. Our data support the claim that the defensive response to threat is indiscriminant egotism. A threat from one domain (sex, violence) triggered a defensive response that seemed appropriate for ego threats, that is, protecting self-esteem.

The observed relation between threat and egotism is explicable by a model of dynamic cognitive complexity (Paulhus, 1984b; Paulhus & Suedfeld, in press). Paulhus and Suedfeld (in press) argue more specifically that upon threat, any cognitive

analysis of the stimulus is reduced in complexity. There is evidence that reduced complexity exaggerates the importance of evaluation (Driver, 1962, as cited in Schroder, Driver, & Streufert, 1967; Paulhus & Lim, 1985). In effect, evaluation comes to engulf the cognitive field. The defensive benefits are detailed in Paulhus and Suedfeld (in press).

Future Directions

The experimental paradigm used in this study lends itself to the pursuit of a variety of issues concerning affect and cognition. Clearly our results leave a number of questions begging to be answered. First, is the potentiation of desirable responding simply a by-product of the confounding of trait desirability with trait endorsement rate? The next study of this issue requires an independent manipulation of desirability and endorsement probability. One solution is to add the two missing categories of traits, namely high desirability–low endorsement and low desirability–high endorsement. If desirability, but not endorsement, was found to predict response potentiation, then the pure arousal model would be undermined. This independent manipulation might be effected within subjects by pretesting subjects in a separate setting and selecting positive, neutral, and negative traits that were rated as descriptive or not descriptive.

A second issue is whether the arousal effects can occur even when subjects are unaware of seeing the affective distractors. Such a finding would add strength to the argument for automaticity. To test this hypothesis, the presentation time of the distractors could be shortened to yield chance recognition. Alternatively, the effects of arousal could be compared for recognized versus unrecognized affective distractors (our data-collection format precluded this analysis). Finally, by assessing later recognition of traits as well as distractors, we could examine the allocation of attention to the two stimuli.

Another critical issue is whether the arousal effects depend on the fact that subjects were rating themselves. Many differences between self- and other-processing have already been documented (e.g., Greenwald & Pratkanis, 1984). To address this issue, subsequent studies might have subjects rate neutral, liked, and disliked strangers. Finally, is there something special about the trait-rating task? A true arousal mechanism should show comparable effects on mental arithmetic, verbal categorization, and even a Stroop test. Such studies would help determine the critical factors underlying this first demonstration of an automatic egotism.

References

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology, 9,* 272–279.

Bargh, J. A. (1984). Automatic and conscious processing of social information. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 1–43). Hillsdale, NJ: Erlbaum.

Bell, P. A., & Byrne, D. (1978). Repression-sensitization. In H. London & J. E. Exner (Eds.), *Dimensions of personality* (pp. 449–486). New York: Wiley.

Blum, G. S., & Barbour, J. S. (1979). Selective inattention to anxiety-linked stimuli. *Journal of Experimental Psychology: General, 108,* 182–224.

Blum, G. S., Geiwitz, P. J., & Stewart, C. G. (1967). Cognitive arousal: The evolution of a model. *Journal of Personality and Social Psychology, 5,* 138–151.

Byrne, D. (1961). The Repression-Sensitization scale: Rationale, reliability and validity. *Journal of Personality, 29,* 334–349.

Byrne, D. (1964). Repression-sensitization as a dimension of personality. In B. Maher (Ed.), *Progress in experimental personality research* (Vol. 1, pp. 169–220). New York: Academic Press.

Clark, M. S., & Isen, A. M. (1982). Toward understanding the relationship between feeling states and social behavior. In A. H. Hastorf & A. M. Isen (Eds.), *Cognitive social psychology* (pp. 73–108). New York: Elsevier.

Clark, M. S., Milberg, S., & Erber, R. (1984). Effects of arousal on judgments of others' emotions. *Journal of Personality and Social Psychology, 46,* 551–560.

Cohen, S. (1978). Environmental load and the allocation of attention. In A. Baum, J. Singer, & S. Valins (Eds.), *Advances in environmental psychology* (Vol. 1, pp. 1–20). Hillsdale, NJ: Erlbaum.

Cooper, J. C., Zanna, M. P., & Taves, P. A. (1978). Arousal as a necessary condition for attitude change following induced compliance. *Journal of Personality and Social Psychology, 36,* 1101–1106.

Corteen, R. S., & Dunn, D. (1974). Shock-associated words in a nonattended message: A test for momentary awareness. *Journal of Experimental Psychology, 102,* 1143–1144.

Corteen, R. S., & Wood, B. (1972). Autonomic responses for shock-associated words in an unattended channel. *Journal of Experimental Psychology, 94,* 308–313.

Demos, R. (1960). Lying to oneself. *Journal of Philosophy, 57,* 588–595.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70,* 80–90.

Dollard, J., & Miller, N. E. (1950). *Personality and psychotherapy: An analysis in terms of learning, thinking and culture.* New York: McGraw-Hill.

Duffy, E. (1962). *Activation and behavior.* New York: Wiley.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66,* 183–201.

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

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* New York: Dryden Press.

Edwards, A. L. (1962). The social desirability hypothesis: Theoretical implications for personality measurement. In S. Messick & J. Ross (Eds.), *Measurement in personality and cognition* (pp. 91–108). New York: Wiley.

Erdelyi, M. H. (1974). A new look at the new look: Perceptual defense and vigilance. *Psychological Review, 81,* 1–25.

Erdelyi, M. H., & Appelbaum, G. A. (1973). Cognitive masking: The disruptive effect of an emotional stimulus upon the perception of contiguous neutral items. *Bulletin of the Psychonomic Society, 1,* 59–61.

Eysenck, M. W. (1982). *Attention and arousal.* New York: Springer-Verlag.

Ferguson, T. J., Rule, B. G., & Carlson, D. (1983). Memory for personally relevant information. *Journal of Personality and Social Psychology, 44,* 251–261.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition.* Reading, MA: Addison-Wesley.

Friedman, A. S. (1964). Minimal effects of severe depression on cognitive functioning. *Journal of Abnormal and Social Psychology, 69,* 237–243.

Gollwitzer, P. M., Earle, W. B., & Stephan, W. G. (1982). Affect as a determinant of egotism: Residual excitation and performance attributions. *Journal of Personality and Social Psychology, 43,* 702–709.

Greenwald, A. G., & Pratkanis, A. R. (1984). The self. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 129–178). Hillsdale, NJ: Erlbaum.

Gur, R. C., & Sackeim, H. A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology, 39,* 1190–1199.

Hall, K. R. L., & Stride, E. (1954). The varying response to pain in psychiatric disorders: A study in abnormal psychology. *British Journal of Medical Psychology, 27,* 48–60.

Hamilton, V. (1983). *The cognitive structures and processes of human motivation and personality.* New York: Wiley.

Heilbrun, A. B. (1964). Social learning theory, social desirability, and the MMPI. *Psychological Bulletin, 61,* 377–387.

Heise, D. R. (1965). Semantic differential profiles for 1,000 most frequent English words. *Psychological Monographs, 79*(8, Whole No. 601).

Howie, D. (1952). Perceptual defense. *Psychological Review, 59,* 308–315.

Humphreys, L. G. (1978). Doing research the hard way: Substituting analysis of variance for a problem in correlational analysis. *Journal of Educational Psychology, 70,* 873–876.

Ichheiser, G. (1960). On Freud's blind spots concerning some obvious facts. *Journal of Individual Psychology, 16,* 45–55.

Isen, A. M. (1984). Toward understanding the role of affect in cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 179–236). Hillsdale, NJ: Erlbaum.

Isen, A. M., & Daubman, K. A. (1984). The influence of affect on categorization. *Journal of Personality and Social Psychology, 47,* 1206–1217.

Isen, A. M., & Means, B. (1983). The influence of positive affect on decision-making strategy. *Social Cognition, 2,* 18–31.

Isen, A. M., Means, B., Patrick, R., & Nowicki, G. (1982). Some factors influencing decision-making strategy and risk-taking. In M. S. Clark & S. T. Fiske (Eds.), *Affect and cognition: The 17th Annual Carnegie Symposium on Cognition* (pp. 243–261). Hillsdale, NJ: Erlbaum.

Izard, C. E. (1977). *Human emotions.* New York: Plenum Press.

Jones, E. E., & Pittman, T. S. (1982). Toward a general theory of strategic self-presentation. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 10, pp. 231–262). Hillsdale, NJ: Erlbaum.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 29–61). New York: Academic Press.

Kimmel, H. D., Van Olst, E. H., & Orlebeke J. F. (Eds.). (1979). *The orienting reflex in humans.* Hillsdale, NJ: Erlbaum.

Kirby, D. M., & Gardner, R. C. (1972). Ethnic stereotypes: Norms on 208 words typically used in their assessment. *Canadian Journal of Psychology, 26,* 140–154.

Kirschenbaum, D. S., Tomarken, A. J., & Humphrey, L. L. (1985). Affect and self-regulation. *Journal of Personality and Social Psychology, 48,* 509–524.

Lacey, J. I. (1967). Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & R. Trumbull (Eds.), *Psychological stress: Issues in research* (pp. 14–42). Englewood Cliffs, NJ: Prentice-Hall.

Leventhal, H. (1974). Emotions: A basic problem for social psychology. In C. Nemeth (Ed.), *Social psychology: Classic and contemporary integrations.* Chicago: Rand McNally.

Lewis, J. (1970). Semantic processing of unattended mirages using dichotic listening. *Journal of Experimental Psychology, 85,* 225–228.

Liebling, B. A., Seiler, M., & Shaver, P. (1974). Self-awareness and cigarette smoking behavior. *Journal of Experimental Social Psychology, 10,* 325–332.

Lockard, J. S., & Paulhus, D. L. (in press). *Self-deception: An adaptive mechanism?* New York: Prentice-Hall.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology, 12,* 523–553.

Malmo, R. B. (1959). Activation: A neuropsychological dimension. *Psychological Review, 66,* 367–386.

Mandler, G. (1975). *Mind and emotions.* New York: Wiley.

Markowitz, A. (1969). Influence of the repression-sensitization dimension, affect value, and ego threat on incidental learning. *Journal of Personality and Social Psychology, 11,* 374–380.

Martin, I., & Rees, L. (1966). Reaction times and somatic reactivity in depressed patients. *Journal of Psychosomatic Research, 9,* 375–382.

McNicol, D. (1972). *A primer of signal detection theory.* London: George, Allen, & Urwin.

Meehl, P. E., & Hathaway, S. R. (1946). The K-factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology, 30,* 525–564.

Mischel, W., Ebbessen, E. B., & Zeiss, A. R. (1973). Selective attention to the self: Situational and dispositional determinants. *Journal of Personality and Social Psychology, 27,* 129–142.

Mueller, C. W., & Donnerstein, E. (1977). The effects of humor-induced arousal upon aggressive behavior. *Journal of Research in Personality, 11,* 73–82.

Nielsen, S. L., & Sarason, I. G. (1981). Emotion, personality, and selective attention. *Journal of Personality and Social Psychology, 41,* 945–960.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

Pallak, M. S., & Pittman, T. S. (1972). General motivational effects of dissonance arousal. *Journal of Personality and Social Psychology, 21,* 349–358.

Pallak, M. S., Pittman, T. S., Heller, J. S., & Munson, P. (1975). The effect of arousal on Stroop color-word task performance. *Bulletin of the Psychonomic Society, 6,* 248–250.

Paulhus, D. L. (1981). Control of social desirability in personality inventories: Principal-factor deletion. *Journal of Research in Personality, 15,* 383–388.

Paulhus, D. L. (1982). Individual differences, self-presentation, and cognitive dissonance: Their concurrent operation in forced compliance. *Journal of Personality and Social Psychology, 43,* 838–852.

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

Paulhus, D. L. (1984b, August). *Self-deception: An information processing approach.* Paper presented at the annual convention of the American Psychological Association, Toronto.

Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaire* (pp. 143–165). New York: Springer-Verlag.

Paulhus, D. L., & Levitt, K. (1983). *A revised Repression-Sensitization scale: The mini-R-S.* Unpublished manuscript, University of British Columbia, Vancouver, Canada.

Paulhus, D. L., & Lim, D. T. K. (1985, August). *The effects of arousal on cognitive complexity and evaluation.* Paper presented at the annual convention of American Psychological Association, Los Angeles.

Paulhus, D. L., Lim, D. T. K., Reid, D. B., & Murphy, G. (1986, August). *Exercise arousal, and social judgments.* Paper presented at the annual convention of the American Psychological Association, Washington, DC.

Paulhus, D. L., & Suedfeld, P. (in press). Self-deception: A dynamic-complexity model. In J. S. Lockard & D. L. Paulhus (Eds.), *Self-deception: An adaptive mechanism?* New York: Prentice-Hall.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive con-

trol. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium.* Hillsdale, NJ: Erlbaum.

Sackeim, H. A., & Gur, R. C. (1978). Self-deception, self-confrontation and consciousness. In G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research* (Vol. 2, pp. 139–197). New York: Plenum Press.

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

Sanders, G. S. (1981). Driven by distraction: An integrative review of social facilitation theory and research. *Journal of Experimental Social Psychology, 17,* 227–251.

Sartre, J.-P. (1956). *Being and nothingness* (H. Barnes, Trans.). New York: Washington Square Press. (Original work published 1943)

Scherer, K. R. (1982). Emotion as process: Function, origin and regulation. *Social Science Information, 21,* 555–570.

Schroder, H. M., Driver, M. J., & Streufert, S. (Eds.). (1967). *Human information processing.* New York: Holt, Rinehart & Winston.

Shiffrin, R. M. (in press). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens's handbook of experimental psychology* (2nd ed.). New York: Wiley.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Slochower, J. (1976). Emotional labelling and overeating in obese and normal weight individuals. *Psychomatic Medicine, 38,* 131–139.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30,* 526–537.

Spence, J. T., & Spence, K. W. (1966). The motivational components of manifest anxiety: Drive and drive stimuli. In C. D. Spielberger (Ed.), *Anxiety and behavior* (pp. 291–326). New York: Academic Press.

Suedfeld, P., Erdelyi, M. H., & Corcoran, C. R. (1975). Rejection of input in the processing of an emotional film. *Bulletin of the Psychonomic Society, 5,* 30–32.

Swets, J. A. (1964). *Signal detection and recognition by human observers.* New York: Wiley.

Tedeschi, J. T. (Ed.). (1981). *Impression management theory and social psychology research.* New York: Academic Press.

Thornton, B. (1984). Defensive attribution of responsibility: Evidence for an arousal-based motivational bias. *Journal of Personality and Social Psychology, 46,* 721–734.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12,* 242–248.

Treisman, A. M., Squire, R., & Green, J. (1974). Semantic processing in dichotic listening? A replication. *Memory and Cognition, 2,* 641–646.

Turner, R. G. (1978). Self-consciousness and speed of information processing. *Personality and Social Psychology Bulletin, 4,* 456–460.

Voyce, C. D., & Jackson, D. N. (1977). An evaluation of a threshold theory for personality assessment. *Educational and Psychological Measurement, 37,* 383–408.

Wallsten, T. S., & Barton, C. (1982). Processing probabilistic multidimensional information for decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 361–384.

Wegner, D. M., & Giuliano, T. (1981). Arousal-induced attention to self. *Journal of Personality and Social Psychology, 38,* 719–726.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37,* 395–412.

Wilson, W. R. (1979). Feeling more than we can know: Exposure effects without learning. *Journal of Personality and Social Psychology, 37,* 811–821.

Winer, B. J. (1971). *Statistical principals in experimental design* (2nd ed.). New York: McGraw-Hill.

Winter, L., Uleman, J. S., & Cunniff, C. (1985). How automatic are social judgments? *Journal of Personality and Social Psychology, 49,* 904–917.

Winters, K. C., & Neale, J. M. (1985). Mania and low self-esteem. *Journal of Abnormal Behavior, 94,* 282–290.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative and Neurological Psychology, 18,* 459–482.

Zajonc, R. B. (1965). Social facilitation. *Science, 149,* 260–274.

Zajonc, R. B. (1980). Preferences need no inferences. *American Psychologist, 35,* 160–171.

Zillmann, D. (1971). Excitation transfer in communication-mediated aggressive behavior. *Journal of Experimental Social Psychology, 7,* 419–434.

Received August 1, 1985
Revision received June 18, 1986 ■